

 Castledown



This work is licensed
under a Creative
Commons Attribution
4.0 International
License.

Chinese language learners evaluating machine translation accuracy

Li-Ching Chang

li-ching.chang@mymail.unisa.edu.au

UniSA Education Futures, University of South Australia

With increasingly rapid advances in machine translation (MT) technology, such as Google Translate, MT has become an indispensable learning resource for second or additional language learners. Many studies indicate that MT or post-editing of MT (PEMT) can be an effective tool for L2 learning and teaching. Nevertheless, little research illustrates how language learners judge or evaluate the accuracy of MT output. The judgement of MT accuracy is essential because MT is not yet error-free. Therefore, the aim of this research is to explore how L2 learners attempt to judge the accuracy of MT output when using MT or PEMT. This study was undertaken through a teaching intervention in an online English-Chinese translation course. Student participants included L2 learners of Chinese studying at a university in Taiwan. Findings from observations of screen recordings and focus group discussions reveal that students use different MT tools and additional online-based resources as complementary strategies to better judge MT accuracy. These include: 1) using more than one MT tool to cross-check MT output; 2) using dictionaries to check word meaning and word usage by looking at example sentences; 3) using search engines to check word definitions, translations, and collocations.

Keywords: Google Translate; machine translation (MT) in language learning; post-editing of MT (PEMT); MT in L2 written production; MT accuracy

Introduction

In the 21st century, one of the practical issues that may attract attention is artificial intelligence (AI) used in all walks of life, including language education. In the era of AI, one of its subfields, called Human Language Technology (HLT), serves to enable machines to comprehend human languages (Varile *et al.*, 1997). Bodomo (2006) indicates that the applications of HLT such as web-based tools (e.g., web-based translators, search engines and multilingual dictionaries) and MT software have been constantly upgraded and used in different areas and

for different purposes, for example, language learning. In this study, MT refers to free online software used for the translation of spoken and written text, such as Google Translate. Online or web-based MT is not only free to access but also easy to use for translating between different pairs of languages immediately on electronic devices, such as computers, tablets, and smartphones. These features and its ongoing improvement make MT a prevalent tool for most language learners (Case, 2015; Groves & Mundt, 2021; Mundt & Groves, 2016). MT has been adopted as a “revision” tool to facilitate language learners’ writing in L2¹ (Lee, 2020; Tsai, 2020). Due to the increased demand for PEMT in language service industries, language instructors have introduced PEMT to their advanced language learners (Niño, 2008, 2009).

Regardless of how MT is applied to language learning, two important considerations are the capabilities of language students 1) to judge the accuracy of MT output; and 2) to ensure greater accuracy in translation for maximal language learning outcomes. For these reasons, a small qualitative study was designed to explore the possibilities of using MT/PEMT for language learning. This may help language educators and practitioners gain a better understanding of L2 learners’ strategies to judge and mediate their use of MT for the purposes of achieving or seeking accuracy. The findings show that using more than one MT tool and additional online-based resources as complementary strategies may help students evaluate MT accuracy more effectively. They indicate that a larger-scale follow-up study would be worthwhile to explore the efficacy of this study. Potentially, a follow-up study might also explore how the findings might contribute to the development of future curricula and pedagogies for language teaching, particularly in computer assisted language learning.

Literature review

State-of-the-art MT technology

Recently, neural machine translation (NMT) has emerged at the cutting edge of MT technology (Briggs, 2018; Bowker & Ciro, 2019; Forcada, 2017). Numerous studies have paid attention to the investigation of the translation quality of state-of-the-art NMT, such as Google NMT (e.g., Wu *et al.*, 2016) and Microsoft NMT (e.g., Hassan *et al.*, 2018). Wu and colleagues (2016) selected five-hundred English sentences from Wikipedia and website news to compare translation quality between Google phrase-based statistical MT (PBMT), Google NMT, and human translation. The selected human translators in their research are fluent in the chosen language pairs (English-French, English-Spanish, English-Chinese). Their study shows that firstly Google NMT can reduce 60% of translation errors compared to Google PBMT. Secondly, the translation quality of Google NMT is comparable to that of human translation. Apart from Google Translate and Microsoft Translator, Baidu and Youdao are two of prominent MT tools and widely used in Mainland China (Wang, 2017). As the largest search engine in China, Baidu has offered MT service since 2011 and adopted NMT systems for Chinese-English translation since 2015 (He, 2015). More recently,

a study by Sun *et al.* (2019) shows that the translation quality of Baidu NMT in English-Chinese news texts has significantly improved. Zhu and Zhang (2021), however, argue that comparing the translation quality of Google, Baidu, Youdao, and official translations using error analysis, Youdao translation performs better in terms of pragmatics.

MT assisted language learning

An increasing number of studies has focused on the investigation of MT in facilitating L2 writing (Cancino & Panes, 2021; Chon, Shin, & Kim, 2020; Lee, 2020; Tsai, 2020). Due to the improvement of MT quality, Lee (2020) and Tsai (2020) instructed students of English to compare their own L2 English writing with MT output for revision. More specifically, Lee (2020) guided Korean students majoring in English to compose their English writing using translation as an approach. Students produced their first draft in Korean and then translated it into English by themselves (without using MT) and also by MT (Google Translate or a Korean MT tool called Papago). Students subsequently compared their own translation with MT output for revision. Lee found that this process helped students reduce lexical and grammatical inaccuracies while improving revisions in L2 writing. In the same vein, Tsai (2020) adopted similar processes to instruct Chinese-speaking students of English, both English and non-English majors, to write an essay by means of translation. Students translated their texts from Chinese to English without consulting any resources, followed by comparing their own English translation with Google English translation for modification. Tsai's study shows that MT as a 'revision' tool can improve L2 writing, and enable students to provide more detailed content, more advanced word choice and fewer inaccuracies in spelling and grammar, particularly for non-English majors.

Moreover, MT has been used to teach L2 learners to post-edit MT errors, or to process PEMT (La Torre, 1999; Niño, 2008, 2009, 2020). La Torre (1999) instructed learners of Spanish at intermediate-high and advanced levels processing PEMT into L1 in order to introduce translators' tools (e.g., MT and online dictionaries) to students. La Torre suggests that the process of assessing MT output helps to develop students' critical thinking and problem-solving skills. Similarly, Niño (2008) introduced PEMT to her advanced Spanish students due to the increased demand for post-editing in the translation industry. The author argues that PEMT not only "triggers error awareness and provides opportunities for error correction through focus on form and negotiation of meaning," but it also supplies students with "language autonomy and vocational training" (p.44–45). Recently, Niño (2020) has conducted a survey with thirty-seven students learning different languages at a university in the UK, and the survey results show that the majority of the students find post-editing necessary when consulting MT.

Despite the ongoing advancement in MT technology and improvement in translation quality, MT output is not yet error-free (Chon *et al.*, 2020; Niño, 2020). Several factors will influence MT output quality (Jolley & Maimone,

2015); therefore, proofreading and post-editing serve as essential elements when MT is consulted (Chon *et al.*, 2020; Niño, 2020). The capability of students in judging the accuracy of MT output, however, is the key to determine whether MT can be used appropriately and whether PEMT can be undertaken effectively (Briggs, 2018; Case, 2015; Jolley & Maimone, 2015). Some language teachers raised their concern regarding L2 learners' judgement of MT accuracy and pointed out that L2 learners might be overconfident in MT output and overlook MT errors, which may lead to the issue of academic integrity (Jolley & Maimone, 2015). Therefore, many studies have suggested the need to guide students in the appropriate use of MT (e.g., Clifford, Merschel & Munné, 2013; Ducar & Schocket, 2018; Jolley & Maimone, 2015; Jiménez-Crespo, 2017). Although there have been some studies on using online resources (e.g., corpora and concordancers) to help with L2 writing (Conroy, 2010; Han & Shin, 2017; Lai & Chen, 2015; Yoon, 2016), very few on how students use MT to solve linguistic problems in L2 written production. It is clear that L2 learners may benefit from using MT output for revising written texts, and that PEMT as a teaching tool may encourage students to make corrections to MT errors while learning language by paying attention to linguistic form and meaning. However, questions about how students evaluate MT accuracy remain unclear. Hence, the aim of this study is to investigate how L2 learners attempt to judge the accuracy of MT output when using MT or PEMT. This research is driven by the following research question: Which strategies and how do L2 learners use these to make effective judgements of MT output?

Methodology

Translation course outline

This study was undertaken through a teaching intervention in an online English-Chinese translation course in the context of L2 learning. The purpose of this translation course was to teach students how to use human translation (HT) and PEMT to engage in translation activities while improving their L2 Chinese capabilities. HT in this study refers to a translation method, in which manual translation is produced by students, potentially with consultation of dictionaries. PEMT, too, denotes a translation method, in which translation first produced through MT is followed by students engaging in a process of post-editing the MT text.

The course was originally designed for face-to-face teaching with a cohort of approximately fifteen students at a university in Taiwan in May 2020. Owing to the unexpected occurrence of COVID-19, the course had to be changed to online and the number of students dropped to four. This specialised translation course took place over eight weeks via Zoom, with a total of fifteen sessions consisting of thirty learning and teaching hours (15 sessions × 2hrs = 30hrs). The five main teaching phases of this course comprised 1) a brief introduction to translation, including HT, MT and PEMT; 2) translation practice through HT and PEMT; 3) comparisons of translations produced by HT and PEMT; 4) translation exercise

through HT, PEMT or a mix of both; and 5) review of translation methods. More importantly, this translation course involved using multiple MT tools to process PEMT to cross-check MT output. The recommended MT included Google Translate, Microsoft Translator, and the translation tools on Baidu and Youdao. The rationale for choosing the four MT tools was because of the improvement in their translation quality as mentioned in the literature review. Besides, the researcher compared them before the intervention began and found that the translation quality of the four tools was somewhat comparable and complementary to some degree in terms of source text meaning, the appropriate use of language (Incl. word choice, phrasing, collocation, and grammar), and style and cohesion (Angelelli, 2009). More details about the course schedule and the focus of each teaching phase are shown in Appendix A.

Participants

Student participants were recruited in a Chinese language program at a university in Taiwan. In addition to the nature of the English-Chinese translation course, previous research suggests that advanced level students are more suitable for the practice of PEMT (Case, 2015; Niño, 2008, 2009). As a result of PEMT requiring rather advanced language abilities in two languages, the requirements for students' language proficiency included fluency in English, and an advanced Chinese level based on the grading of the Chinese language program, which is equivalent to B2–C1 level of Common European Framework of Reference (CEFR) for languages. The recruited four participants included three native speakers of English from the US and Australia, and one native speaker of Dutch from Belgium. Regarding their Chinese proficiency, three (Student 2, 3 & 4) were at CEFR B2 levels and one (Student 1) with a C1 level.

Data collection and analysis

Observations of screen recordings of pre- and post-tasks and focus group discussions were used to triangulate data from different dimensions for gaining a more comprehensive understanding of this study. A triangulation of more than one research method aims at 'knowing more' on a social phenomenon (Moran-Ellis *et al.*, 2006, p. 45). Students were requested to complete pre- and post-tasks on the Zoom platform using the screen sharing feature to record their translation processes and behaviours. Observation sheets were prepared to document key observation points pertaining to the research purpose (see Appendix B). The prepared observation sheets, in fact, were reviewed and revised iteratively several times, based on the research question and the collected screen recordings, prior to finalisation. In brief, the observations focused on students' interaction with MT and their strategies to evaluate MT output. In addition to the use of observation sheets, screenshots were also adopted to capture digital images to demonstrate students' use of MT. As Burns (2009) suggests that images can supplement observation notes and provide a more immediate impression than words.

With regard to the pre- and post-tasks, students were required to translate two Chinese texts into English and two English texts into Chinese within a period of 90 minutes. They were allowed to use the internet to access any online resources during the pre- and post-tasks, although no emphasis was placed on the use of MT (see Appendix C for task instructions and the four texts). Each text was a 100-word paragraph excerpted from authentic texts,² including a Chinese magazine article (Text1), Chinese idiom story (Text 2), English biography (Text 3), and English short novel (Text 4). The selected texts for the pre- and post-tasks were not identical but were extracted from the same articles to maintain the same level of lexical and syntactical difficulty, except for Text 2. The length of Text 2 for the pre-task was already around 100 words, therefore, another version of the same story was chosen for the post-task. At the end of the translation course, a semi-structured focus group discussion was carried out to understand in more depth how students judged the accuracy of MT output and their views on PEMT during the intervention. The discussion questions were: 1) How do you ensure MT is accurate? 2) Have your views of using PEMT changed during the course? 3) Has this course been useful? How? In summary, the procedure of data collection is shown in Figure 1 below.

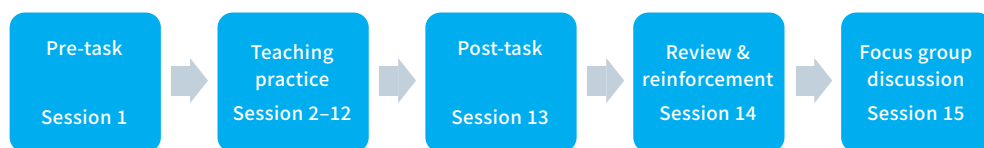


Figure 1. Procedure for data collection

As shown in Figure 1, a pre-task was set in Session 1 before the teaching component of the intervention began and a post-task in Session 13. A course review and reinforcement were offered in Session 14 ahead of a focus group discussion in Session 15. Regarding data analysis, an inductive approach was applied for coding and categorising. The prepared observation sheets had already included initial categories that were created through an iterative revision process. The preliminary categories included numbers of MT tools used, other resources used, and the process of translation and problem solving. Moreover, the focus group discussion was transcribed and analysed using thematic analysis (Braun & Clarke, 2012) to engender themes relating to the research question. Simply put, the two datasets were analysed separately in the first place, followed by an integrated analysis (Moran-Ellis *et al.*, 2006) to interweave findings that can produce a fuller picture of students' interaction with MT.

Results

Three of the students (Student 1, 2, and 4) used Google Translate to help with translation in the pre-task, even though the use of MT was not emphasised in the task instructions. When using Google Translate, they usually employed additional resources, such as dictionaries, to double-check MT output. In the post-task, the three students changed their approach to adopt more than one

MT tool together with their commonly used resources to evaluate MT accuracy. Student 3 did not use MT in the pre-task, although she adopted Google Translate to help with English to Chinese translation in Text 3 and Text 4 in the post-task. Therefore, the presentation of screen observation results only focuses on Student 1, 2, and 4 to illustrate the changes in their use of MT between the pre- and post-tasks. What follows is to 1) demonstrate observation results on how the three students evaluate the accuracy of MT output; and 2) combine screen observation and focus group discussion results to gain a more rounded understanding of what strategies students use to make judgments of MT accuracy.

Screen observations

Student 1 (S1) – L1 English & C1 Chinese level

Table 1. S1 using MT in the pre- and post-tasks

S1 Pre-task				
Text types	C to E	C to E	E to C	E to C
	1. Magazine			
MT adoption	article	2. Idiom story	3. Biography	4. Short novel
Number of MT tools used	1	1	1	1
Using other resources to cross-check MT accuracy	No	No	Yes	Yes
Process of translation and problem solving	HT with Google Translate assistance for words/phrases (using other resources to double-check MT Chinese output) → Comparing translations made by HT and MT for revision when needed			
S1 Post-task				
Text types	C to E	C to E	E to C	E to C
	1. Magazine			
MT adoption	article	2. Idiom story	3. Biography	4. Short novel
Number of MT tools used	4	4	4	4
Using other resources to cross-check MT accuracy	Yes	Yes	Yes	Yes
Process of translation and problem solving	Using multiple MT tools to get different MT outputs → Choosing one MT output as a draft to do PEMT (using other MT outputs and resources to cross-check MT accuracy) → Comparing translations made by PEMT and MT for revision when needed			

As shown in Table 1, S1 used HT to engage in translation while adopting Google Translate to look up words and phrases in the pre-task. This student did not double-check MT output when using it for Chinese to English translation in Text 1 and Text 2, although he proofread MT output using the Pleco dictionary when translating from English to Chinese with Text 3 and Text 4. After completing translation through HT, S1 compared his own translation with MT output for revision, particularly for translation he was not so sure about. In the post-task, S1 adopted four MT tools to obtain four versions of MT output and subsequently made a judgment about which of the four appeared to produce the

best version to serve as a draft for processing PEMT. To judge MT accuracy, he referred to the other three MT outputs and adopted other resources for cross-checking. Like in the pre-task, S1 would compare translation made by PEMT with MT output when needed.

Figure 2 shows screenshots of S1 using Google Translate output to compare with his own translation of Text 1 in the pre-task. S1 spent most of the time on the underlined Chinese sentence in Figure 2. The reference translation for the underlined Chinese sentence is as follows: *Research shows that current education treats people too much like ‘hardened clay’ that cannot change its shape after it has been moulded and fired, unlike ‘wet clay’ that can be reshaped.*



Figure 2. S1 using one MT tool in the pre-task

As shown in Figure 2, S1 was uncertain about if his translation through HT (marked as HT) was accurate, especially the underlined sentence, so he referred to the output of Google Translate (marked as GT) for confirmation. By checking with GT, S1 found that he misinterpreted the underlined Chinese source sentence; therefore, he made modifications to his original translation (marked as revision version). Analysed from the revised version, S1 significantly improved his translation by checking with GT, nevertheless, the underlined sentence remained slightly inaccurate compared with the source text. One possible explanation might be that because S1 only consulted Google Translate, he could only double-check his HT against one MT output. One hypothesis is that the use of a second or third MT tool might offer increased opportunities for accuracy. In the post-task, S1 changed his approach to use multiple MT tools to perform PEMT as shown in Table 2.

Table 2. S1 using multiple MT tools in the post-task

Text 1: Chinese to English translation			
<p>Part of the source text: 因應科技發展,教育方式亦隨之改變...AI崛起亦有可能帶來同樣的效應,讓教育變革成為必要之事,並以適性教育來促成教育變革。</p> <p>Reference translation: <i>In response to the development of technology, education has also changed accordingly...The rise of AI may also bring about the same effect, making education reform a necessity, with 'adaptive teaching' driving education reform.</i></p>			
<p>Google translation: In response to the development of science and technology, the education method has also changed... The rise of AI may also bring about the same effect, making educational reform a necessity, and using *appropriate education to promote educational reform.</p>	<p>Baidu translation: In response to the development of science and technology, the way of education has also changed...The rise of AI may also bring about the same effect, making education reform necessary, and promoting education reform with *appropriate education.</p>	<p>Youdao translation: Education has changed in response to technological development...The rise of AI is likely to have the same effect, making education change necessary and bringing about change in *sexuality education.</p>	<p>Microsoft translation: In response to technological developments, the way education has changed has been followed...The rise of AI may have the same effect, making education change necessary and promoting educational change with *appropriate sex education.</p>
<p>S1 final translation made by PEMT with the use of multiple MT tools: Education methods change in response to technological developments...The rise of AI is likely to have a similar effect, making the transformation of education an unavoidable necessity, with '<i>adaptive learning</i>' forming the basis of education reform.</p>			

As demonstrated in Table 2, S1 used multiple MT tools to process PEMT in the post-task. He skimmed through the translation of the four MT tools and selected the best MT version in his view as a draft to perform PEMT. In this text, S1 initially chose Baidu translation as a draft but soon after changed to the output of Microsoft translator. Still not satisfied, he eventually selected Youdao translation as a basis to do post-editing. When coming to the italic Chinese term ‘適性教育 (adaptive teaching/learning)’, this student consulted Pleco and referred to the output of the other three MT tools, as indicated by asterisks. To further understand this Chinese term, he used Google Search to look for its definition and translation to solve this translation problem and obtain an English correspondence (see Figure 3).

As shown in Figure 3, S1 typed the Chinese term on the Google search engine and obtained relevant information and useful translations of this term, such as ‘adaptive education’. He then looked for the meaning of ‘adaptive education’ for further information. Analyses of this example suggest that employing multiple MT tools and dictionaries for cross-checking and Google Search for seeking definitions of terms seems to help S1 post-edit the inaccurate MT output (e.g., sexuality education) to an appropriate translation (e.g., adaptive learning).

This strategy in turn enables him to render a more adequate final translation as shown at the bottom of Table 2.



Figure 3. S1 solving translation problems using Google Search

Student 2 (S2) – L1 Dutch, fluency in English & B2 Chinese level

Table 3. S2 using MT in the pre- and post-tasks

S2 Pre-task				
Text types	C to E	C to E	E to C	E to C
	1. Magazine			
MT adoption	article	2. Idiom story	3. Biography	4. Short novel
Number of MT tools used	1	1	1	1
Using other resources to cross-check MT accuracy	Yes	Yes	Yes	Yes
Process of translation and problem solving	Pleco, Nciku (LineDict), Google search engine			
Process of translation and problem solving	HT with Google Translate assistance for words/phrases (using other resources to check MT accuracy → Translating sentences using Google Translate for doing PEMT (using other resources to check MT accuracy)			
S2 Post-task				
Text types	C to E	C to E	E to C	E to C
	1. Magazine			
MT adoption	article	2. Idiom story	3. Biography	4. Short novel
Number of MT tools used	2	2	3	3
Using other resources to cross-check MT accuracy	Yes	Yes	Yes	Yes
Process of translation and problem solving	Pleco, MOE dictionary, Google search engine			
Process of translation and problem solving	For Text 2: HT with Google and Baidu assistance for words/phrases → Translating sentences using Google Translate for doing PEMT (using other resources to check MT accuracy) For Text 1, 3 & 4: Using different MT tools to translate a paragraph → Using different versions of MT output as references to produce translations (adopting other resources to cross-check MT accuracy)			

As shown in Table 3, in the pre-task S2 completed translations through HT with the use of Google Translate to look up words and phrases. She would adopt MT to translate a sentence to perform PEMT when encountering difficulties in translation. When consulting MT, she would use other resources listed in the table to double-check MT accuracy. S2 changed her approach to employ varied MT tools (Incl. Google, Baidu and Youdao) to complete translation in the post-task. This student's translation process in Text 2 was similar in both pre- and post-tasks, but she used two MT tools (Google and Baidu) to look up words/phrases to double-check MT output. For the other three texts, S2 adopted different MT tools to obtain varied translation versions as references to produce final translations. Additional resources were used to cross-check MT accuracy.

Figure 4 shows screenshots of S2 dealing with translation difficulties using Google Translate in Text 4 English to Chinese translation in the pre-task.

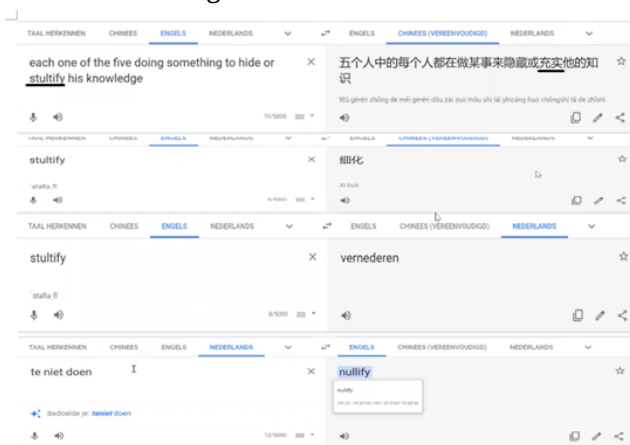


Figure 4. S2 using one MT tool in the pre-task

As shown in Figure 4, S2 accessed Google Translate to obtain a Chinese sentence translation to process PEMT. MT mistranslated the underlined English word 'stultify' to Chinese³ '充實 (enrich)'. To understand the meaning of 'stultify', S2 used Google Translate to make a three-directional translation, from English to Chinese, from English to Dutch, and from Dutch to English. As indicated in Figure 4, this Google Chinese translation '細化 (refine)' remained inaccurate so S2 resorted to L1 Dutch to better understand this English word. In addition, S2 consulted Pleco to determine a suitable Chinese translation. In so doing, she eventually post-edited this inaccurate translation '充實 (enrich)' to a more adequate translation '貶低 (belittle)'. Despite her effort of correcting this MT error, S2 did not accept this accurate translation '隱藏 (to hide)', but instead she changed it to an incorrect translation '隱瞞 (to hide the truth)', which was unsuitable in terms of Chinese collocation. Accordingly, S2 was unable to correctly judge the accuracy of MT Chinese output using only one MT tool and dictionaries.

In the post-task, S2 adopted multiple MT tools and critically compared between the source text and different versions of MT output. She paid close attention to the varied translations of MT tools as indicated by the blue underlined text in Figure 5.

In June 2012 the U.S. Anti-Doping Agency (USADA) alleged that Armstrong and five of his associates had been part of a decadelong doping conspiracy beginning in the late 1990s. According to USADA, Armstrong used performance-enhancing drugs. The allegations

google

2012年6月,美国反兴奋剂机构(USADA)指控阿姆斯特朗和他的五个同伙参加了从1990年代末开始的长达十年的兴奋剂阴谋。根据USADA的说法,阿姆斯特朗使用了增强性能的药物。这些指控导致他立即被禁止参加铁人三项比赛。那年8月,他拒绝进入USADA的仲裁程

baidu

2012年6月,美国反兴奋剂机构(USADA)指控阿姆斯特朗及其5名同伙参与了上世纪90年代末开始的十代龙兴奋剂阴谋。根据USADA的说法,阿姆斯特朗使用了兴奋剂。这些指控导

youdao

2012年6月,美国反兴奋剂机构(USADA)指控阿姆斯特朗和他的五名助手参与了始于20世纪90年代末的长达10年的兴奋剂阴谋。据美国反兴奋剂机构称,阿姆斯特朗使用了提高成绩的药物。这些指控导致他立即被禁止参加铁人三项比赛。当年8月他拒绝进入美国的仲裁过程。

Figure 5. S2 using multiple MT tools in the post-task

As shown in Figure 5, the first part is the English source text, followed by the translation of Google, Baidu, and Youdao. This word ‘decadelong’ was translated to Chinese ‘長達十年 (ten-year-long)’ by Google, ‘十代龍 (ten generation dragon)’ by Baidu, and ‘長達十年 (ten-year-long)’ by Youdao. This term ‘performance-enhancing drugs’ was translated to Chinese ‘增强性能的藥物 (function-enhancing drugs)’ by Google, ‘興奮劑 (dope)’ by Baidu, and ‘提高成績的藥物 (grades-improving drugs)’ by Youdao. For evaluating MT output, S2 adopted Google and Baidu as dictionaries together with Pleco to cross-check word meanings and translations. She also looked at bilingual example sentences on Baidu to see how words were used in the context of sentences as shown in Figure 6. With the aid of multiple MT tools and dictionaries, S2 finally translated ‘decadelong’ to ‘長達十年 (ten-year-long)’ and ‘performance-enhancing drugs’ to ‘提高運動成績的藥物 (sport performance-enhancing drugs)’.



Figure 6. S2 looking at example sentences for word use

Student 4 (S4) – L1 English & B2 Chinese level

Table 4. S4 using MT in the pre- and post-tasks

S4 Pre-task				
Text types	C to E	C to E	E to C	E to C
	1. Magazine			
MT adoption	1. Magazine article	2. Idiom story	3. Biography	4. Short novel
Number of MT tools used	1	1	1	1
Using other resources to cross-check MT accuracy	Yes	Yes	Yes	Yes
Process of translation and problem solving	Firefox English-Chinese pop-up dictionary, Pleco, search engines			
	For Text 1: HT with Google Translate assistance for words/phrases → Translating sentences using Google Translate for doing PEMT → Comparing translation made by HT&PEMT and MT for revision For Text 2, 3 & 4: Translating paragraphs using Google Translate for doing PEMT (using other resources to double-check MT output)			
S4 Post-task				
Text types	C to E	C to E	E to C	E to C
	1. Magazine			
MT adoption	1. Magazine article	2. Idiom story	3. Biography	4. Short novel
Number of MT tools used	2	2	2	2
Using other resources to cross-check MT accuracy	Yes	Yes	Yes	Yes
Process of translation and problem solving	Firefox English-Chinese pop-up dictionary, Pleco, search engines			
	For Text 1: HT with Google Translate assistance for words/phrases → Translating sentences using Google Translate for doing PEMT → Comparing translation made by HT&PEMT and MT for revision For Text 2, 3 & 4: Translating paragraphs using Google and Baidu for doing PEMT (using other resources to double-check MT output)			

As shown in Table 4, S4 used HT and PEMT to translate Text 1 and adopted PEMT to translate the other three texts in both pre- and post-tasks. Additional resources were used to double-check MT accuracy. The difference between pre- and post-tasks was that she used only Google Translate in the pre-task but adopted two (Google and Baidu) in the post-task. Table 5 shows S4 performing PEMT using only Google Translate and dictionaries for Text 3 English to Chinese translation in the pre-task.



Table 5. S4 using one MT tool in the pre-task

Text 3: English to Chinese translation		
<p>Part of the source text:</p> <p>... an investigation revealed that he was the key figure in a wide-ranging doping conspiracy while he compiled his Tour victories.</p> <p>Reference translation:</p> <p>...調查顯示他在取得環法自行車賽勝利的同時,也是大規模興奮劑陰謀案的關鍵人物。</p>	<p>Google translation:</p> <p>...調查顯示他是他在*收集環法自行車賽勝利時進行廣泛興奮劑陰謀的關鍵人物。(...an investigation showed that he was when he was *collecting the Tour victories the key figure in a wide-ranging doping conspiracy.)</p>	<p>PEMT version:</p> <p>...調查顯示他是他在取得環法自行車賽*成績時進行廣泛興奮劑案件的關鍵人物。(...an investigation showed that he was when he was receiving the Tour *scores the key figure in a wide-ranging doping conspiracy.)</p>

As shown in Table 5, S4 used Google Translate output as a draft to process PEMT. The source text English sentence ‘he...while he...’ was mistranslated to Chinese ‘他是他在...時 (he was when he was)’, as indicated by the underlined text. In addition, it was also unsuitable to translate the English word ‘compile’ to Chinese ‘收集 (collect)’ in context as indicated by the asterisk, because ‘收集勝利 (collect victories)’ was inappropriate in Chinese collocation. Analysed from the PEMT version, S4 did not detect the MT error in the ‘while’ sentence and correct it. Even though she post-edited ‘收集 (collect)’ to a more suitable translation ‘取得 (receive/gain)’ with the use of Pleco and the Firefox pop-up dictionary, S4 changed this appropriate MT output ‘勝利 (victories)’ to a less suitable word ‘成績 (scores/results)’. Accordingly, S4 could not successfully judge MT accuracy using only one MT tool and dictionaries.

In the post-task, S4 adopted two MT tools and the same dictionaries to perform PEMT. As shown in Figure 7, this student used Google Translate output as a draft (marked as GT) and Baidu translation (marked as BT) together with the Firefox pop-up dictionary to cross-check MT Chinese output.

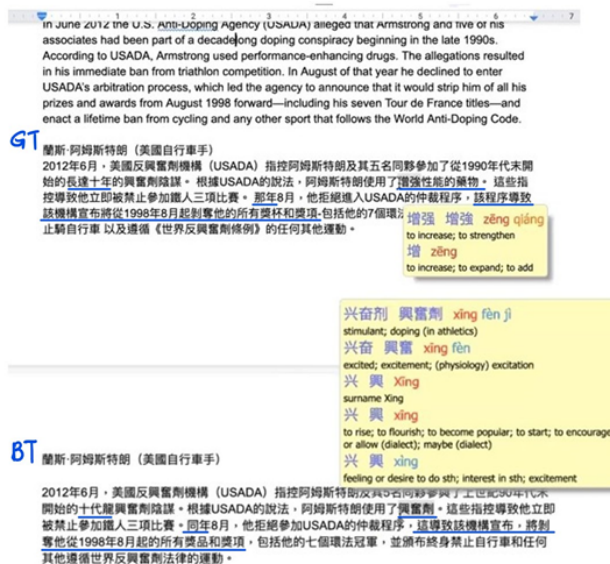


Figure 7. S4 using two MT tools and dictionaries in the post-task



Like S2, S4 analytically compared and evaluated the different outputs of Google and Baidu based on the source text, as indicated by the blue underlined text. For example, Google mistranslated the source text English sentence 'he declined to enter USADA's arbitration process, which led the agency to announce that it would strip him of all his prizes and awards from August 1998 forward' to Chinese '他拒絕進入USADA的仲裁程序' 該程序導致該機構宣布將從1998年8月起剝奪... (He refused to enter USADA's arbitration process. This process led the agency to announce that starting from August 1998, they would strip him of all his prizes...). S4 identified this error and post-edited it when cross-referencing with BT in which the sentence was translated accurately.

Integration of screen observation and focus group discussion results

The screen observation results show that when these students employ more than one MT tool together with other commonly used resources, they become more capable of evaluating MT accuracy. This in turn may enable them to produce more satisfactory machine-assisted translation. Although an analysis of students' translation accuracy/quality goes beyond the aim of this study, the assessment of the four students' translations from Chinese to English (Text1) and from English to Chinese (Text 3) in the pre- and post-task is shown in Appendix D for reference. Table 6 displays the combined results of screen observations and focus group discussions to show the students' strategies to evaluate MT accuracy.

As shown in Table 6, three out of four students used more than one MT tool and all of them adopted dictionaries and/or search engines to cross-check MT output. The students noticed that different MT tools might produce varied translations, therefore, S1, S2 and S4 had adopted a pattern of using more than one MT tool to double-check MT output of different language units (e.g., words, phrases, and sentences). Additionally, dictionaries were not only used to check word meaning but also to understand word usage by looking at example sentences. Another online resource, search engines, particularly Google Search, were employed to 1) look for word definitions and translations; 2) access Google Image to verify MT output (e.g., green beans); 3) access Google News to check word collocations for news texts; and 4) check collocations produced by MT.

Table 6. Combined results of screen observations and focus group discussions

Data sources Resources	Screen observations	Focus group discussions
More than one MT tool	<ul style="list-style-type: none"> ▶ Choosing the best MT version in the view of the student as a draft to do PEMT while cross-referencing with other MT output [S1, S4] ▶ Using different versions of MT output for cross-referencing [S2] ▶ Cross-checking MT output of different language units [S1, S2, S4] 	<p>“I’ve learned two methods in this course. One is making comparisons of machine translations using multiple MT tools. If translations from MT tools are the same, the translation accuracy rate will be higher; but if MT translations are different, I will then pay more attention to the difference” [S1].</p> <p>“I definitely learned about how different machine translation software differ from each other (before I thought that it didn’t matter which one you used)” [S4].</p>
Dictionaries – smartphone apps, online dictionaries, and MT build-in dictionaries	<ul style="list-style-type: none"> ▶ Using dictionaries to double-check word meaning [S1, S2, S3, S4] ▶ looking at Chinese example sentences to have better ideas of Chinese word usage [S2, S4] 	<p>“I typically use Baidu to search for example sentences to get a better idea if I know if I am using a word properly...The English in these examples sentences is not very good... Sometimes I won’t read the English at all, just the Chinese example sentence” [S3].</p>
Search engines	<ul style="list-style-type: none"> ▶ Looking for definitions and/or translations of words/phrases [S1, S2, S3, S4] ▶ Accessing Wikipedia in Chinese to look for Chinese proper nouns when translating Text 3 from English to Chinese [S1, S2, S3, S4] ▶ Putting Chinese phrases provided by MT on the Google search engine to check usage and suitability [S1, S2, S3, S4] 	<p>“...using Google search engine, for example, if I want to know the Chinese translation of ‘green beans’, I will use MT and then put the translation on Google Image to see if it is correct” [S3].</p> <p>“I use Google News on Google search engine to check word collocations” [S2].</p> <p>“...placing collocations provided by MT on the Google search engine to see if Chinese speakers would use this collocation or to see if that is a common collocation” [S1].</p>

In the focus group discussion, S1 and S3 argued that the key determinant for effectively judging MT accuracy depended on translators’ language proficiency in both languages. Nonetheless, three of the students (S1, S2 & S4) considered it advantageous to adopt varied MT tools to perform PEMT because this could enable them to ensure greater accuracy in their translation. As S4 stated, “I have a better understanding of how to process PEMT and how to make use of PEMT to improve the accuracy of my translation”. In addition, the



multiple-resource-based PEMT strategies may help students enhance translation and problem-solving abilities as indicated in the observation data. S2 suggested, “After this course, I knew more translation methods so I could translate faster and more accurately.”

More importantly, the students mentioned that they all gained a clearer understanding of MT and how to process PEMT appropriately after the intervention. S3 said, “I now better understand how to use MT and in what context I would like to use.” This echoes the opinion of S1, “I’d never thought of using multiple MT tools when translating, nor did I understand the limitations of MT. Now I have a better understanding of the chance of translation errors made by MT.”

Accordingly, this translation course seems to help the students understand how to use MT and its benefits and limitations. PEMT is considered to be essential in seeking to use MT appropriately and achieve greater accuracy in translation. Yet, judging MT accuracy is the first step prior to post-editing. The findings of this study reveal that students’ strategies to effectively judge the accuracy of MT output include a recursive process: 1) using more than one MT tool to cross-check MT output of different language units; 2) adopting dictionaries to check word meaning and understand word usage by looking at example sentences; and 3) employing search engines to check word definitions, translations, and collocations (not necessarily in this sequence).

Discussion

As the students notice that MT output is not as precise so far, they usually adopt additional resources to proofread MT output as shown by screen observation results. In accordance with the present results, previous research has demonstrated that students will consult preferred resources, for example dictionaries, when performing post-editing (e.g., La Torre, 1999; Nino, 2008, 2009). Nevertheless, this study showed that the students seemed less able to evaluate MT accuracy before explicit teaching of this was provided in the course. This finding may indicate that even though students always practise proofreading or post-editing when using MT, they do not necessarily know how to judge the extent of MT accuracy or how to undertake PEMT. Screen observation results showed that S1 consulted Google Translate to check and revise his translation produced by HT. This finding is consistent with those of Lee (2020) and Tsai (2020) who argue that MT can be used as a revision tool to improve written production. However, the results also indicated that the inadequate translation of S1 in the revised version might result from the fact that only one MT tool was consulted. The reason why S1 checked his English translation with MT English output was because he could not fully comprehend the L2 Chinese source text. This prevented the student from correctly judging the accuracy of MT output from Chinese to English. Similarly, owing to the linguistic gap between English and Chinese, S2 and S4 were unable to successfully judge the accuracy of MT output from English to Chinese with the use of only one MT tool and dictionaries in the pre-task.

In the post-task, the three students (S1, S2 and S4) all changed their approach to use more than one MT tool together with their frequently used resources to process PEMT as they believed it beneficial to help them produce more accurate machine-assisted translation. These findings may be explained by the fact that dictionaries are more commonly used to obtain meanings and forms of single words (Lai & Chen, 2015), and search engines are usually adopted to look for word usage and collocation (Han & Shin, 2017). In contrast, MT can translate not only small language units (e.g., words) but also larger language units (e.g., sentences) in context, which may provide L2 learners with situationally appropriate L2 usage (Chon *et al.*, 2020; Lee, 2020). Given that MT can serve as a dictionary and more than a dictionary (i.e., MT can translate larger language chunks), using more than one MT tool to cross-check MT output may increase accuracy rate, as indicated by S1 in Table 6.

Based on the three-phase model of the translation process (Nord, 2005), translation comprises an understanding of the source text, a transfer of meaning corresponding to the expectations of the target language culture, and a reconstruction of the target text. Thus, a culturally and linguistically appropriate transfer of meaning from the source text/language to the target text/language requires a good command of L2 in both input (e.g., reading) and output (e.g., writing). In terms of translating from L2 Chinese to English, Chinese source text comprehension serves as an important element to determine English translation quality. Having different versions of MT output in L1 or in a fluent language for cross-referencing and cross-checking may help students judge MT accuracy and gain a better understanding of the source text in L2. In terms of translating from English to L2 Chinese, the capabilities of appropriate use of Chinese language play the main role in Chinese translation. Using varied versions of MT output in L2 for cross-referencing may empower students to judge MT accuracy and post-edit MT errors to render a better translation in L2 (see Appendix D). In addition to the adoption of different MT tools, other commonly used resources such as dictionaries and search engines indeed account for essential tools to check and confirm MT output, as indicated by the results of this study.

This research found that students employed bilingual dictionaries to double-check the meanings of words provided by MT and to look up word usage by looking at example sentences in the dictionaries. This means students not only use dictionaries to look up word meanings and form as shown in the research by Lai & Chen (2015), but also to gain a better understanding of how a word or phrase is used by checking example sentences. Another interesting finding was that students utilised search engines to look for word definitions, translations, and collocations for confirmation. Search engines as corpora and concordancers collect a considerable corpus of linguistic data from the web (Han & Shin, 2017; Yoon, 2016). Learners can employ search engines as a tool to look for examples of how a word, phrase or grammar is used to understand the correct use of language. The results of this study corroborate the findings of previous work on L2 learners using multiple resources as a complementary strategy to help with L2 writing and ensure the accuracy of L2 production

(Conroy, 2010; Lai & Chen, 2015; Yoon, 2016). As discussed thus far, consulting different MT resources pushes students to make a purposeful comparison of MT output while having more translation alternatives from MT at hand as points of reference. To further process critical evaluation of MT accuracy, students subsequently utilise dictionaries in combination with search engines as needed in the process of PEMT.

Making effective judgements of MT accuracy and performing PEMT are important language learning and translation capabilities while MT output remains imperfect. This research serves as a preliminary small-scale study and sheds light on approaches that enhance the capabilities of advanced learners of Chinese language to evaluate MT output through various strategies. One might hypothesise that the findings are beneficial for both English-speakers learning Chinese and Chinese-speakers learning English.

Conclusion

The aim of this study was to explore how L2 Chinese learners judge the accuracy of MT output when using MT or PEMT. This research was undertaken through a teaching intervention in a translation course which involved using different MT tools for cross-checking when processing PEMT. The results of this study showed that students seemed less able to judge MT accuracy when using only one MT tool and their preferred resources such as dictionaries. They, however, became more capable of judging MT accuracy when using more than one MT tool together with the same resources. The major finding of this study, therefore, is that students employ multiple resources as complementary strategies to better evaluate MT accuracy, which includes: 1) using more than one MT tool to cross-check MT output; 2) adopting dictionaries to check word meaning and word usage by looking at example sentences; and 3) employing search engines to check word definitions, translations and collocations. Despite the limitations of the scale and delivery of this study owing to the disruption of COVID-19 in 2020, there are sufficiently useful findings to warrant further larger-scale research with implications for language teaching and learning pedagogy. The findings of this study suggest that language teachers can help L2 learners effectively judge MT accuracy by providing them with two or more useful MT tools for cross-checking and cross-referencing. It is important to teach students how to use the multiple-resource-based strategies to evaluate MT output critically or to process PEMT appropriately for not only generating more satisfactory L2 production but also improving computer assisted language learning outcomes.

Notes

1. L2 is used as an umbrella term throughout this study to refer to learning any language other than the first, whether it takes place in the second or foreign language context.
2. Although the texts were selected from authentic texts, the difficulty level

of each text was also considered to avoid choosing materials that were far beyond students' L2 Chinese proficiency.

3. For consistency purposes, traditional Chinese is used in the text for discussion, regardless of whether simplified Chinese or traditional Chinese is shown in the figures.

Acknowledgements

I would like to thank the anonymous reviewers for their feedback, Kathleen Heugh and Enza Tudini for the comments on the draft of this paper, and the students who participated in this study. I also want to acknowledge the support of an Australian Government Research Training Program Scholarship.

References

- Angelelli, C. V. (2009). Using a rubric to assess translation ability. *Testing and assessment in translation and interpreting studies*, 13–49.
- Bodomo, A. (2006). Human language technology in multilingual perspectives: Institutions and applications. *International Journal of Technology and Human Interaction*, 2(1), i–vi.
- Bowker, L., & Buitrago-Ciro, J. (2019). *Machine translation and global research: Towards improved machine translation literacy in the scholarly community*. Bingley: Emerald Publishing Limited.
- Braun V., & Clarke V., (2012) Thematic Analysis. In P. Liamputtong (Eds.), *Handbook of research methods in health social sciences* (pp. 843–860). Singapore: Springer.
https://doi.org/10.1007/978-981-10-5251-4_103
- Briggs, N. (2018). Neural machine translation tools in the language learning classroom: Students' use, perceptions, and analyses. *The JALT CALL Journal*, 14(1), 2–24.
- Burns, A. (2009). *Doing action research in English language teaching: A guide for practitioners*. New York: Routledge.
- 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.
- Case, M. (2015). Machine translation and the disruption of foreign language learning activities. *eLearning Papers*(45), 4–16.
- Chon, Y.V., Shin, D., & Kim, G. E. (2020). Comparing L2 Learners' writing against parallel machine-translated texts: Raters' assessment, linguistic complexity and errors. *System*, 102408.
- 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).
- Conroy, M.A. (2010). Internet tools for language learning: University students taking control of their writing. *Australasian Journal of Educational Technology*, 26(6), 861–882.

- 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.
- Forcada, M.L. (2017). Making sense of neural machine translation. *Translation spaces*, 6(2), 291–309.
- 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, 100957.
- Han, S., & Shin, J.-A. (2017). Teaching Google search techniques in an L2 academic writing context. *Language Learning & Technology*, 21(3), 172–194.
- Hassan, H., Aue, A., Chen, C., Chowdhary, V., Clark, J., Federmann, C., Huang, X., Junczys-Downmunt, M., Lewis, W., & Li, M. (2018). Achieving human parity on automatic chinese to english news translation. *arXiv preprint arXiv:1803.05567*.
- He, Z. (2015). Baidu translate: Research and products. In *Proceedings of the Fourth Workshop on Hybrid Approaches to Translation (HyTra)* (pp. 61–62). Beijing: Association for Computational Linguistics.
- Jiménez-Crespo, M.A. (2017). The role of translation technologies in Spanish language learning. *Journal of Spanish Language Teaching*, 4(2), 181–193.
- Jolley, J.R., & Maimone, L. (2015). Free online machine translation: Use and perceptions by Spanish students and instructors. *Learn Language, Explore Cultures, Transform Lives. Eau Claire, WI: Robert M. Terry*, 181–200.
- La Torre, M.D. (1999). A web-based resource to improve translation skills. *ReCALL*, 11(3), 41–49.
- Lai, S.L., & Chen, H.J.H. (2015). Dictionaries vs concordancers: Actual practice of the two different tools in EFL writing. *Computer Assisted Language Learning*, 28(4), 341–363.
- 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>
- Moran-Ellis, J., Alexander, V. D., Cronin, A., Dickinson, M., Fielding, J., Sleney, J., & Thomas, H. (2006). Triangulation and integration: processes, claims and implications. *Qualitative Research*, 6(1), 45–59.
- 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.
- Niño, A. (2008). Evaluating the use of machine translation post-editing in the foreign language class. *Computer Assisted Language Learning*, 21(1), 29–49.
- Niño, A. (2009). Machine translation in foreign language learning: language learners' and tutors' perceptions of its advantages and disadvantages. *ReCALL*, 21(2), 241–258.
- Niño, A. (2020). Exploring the use of online machine translation for independent language learning. *Research in Learning Technology*, 28.
- Nord, C. (2005). *Text analysis in translation: Theory, methodology, and didactic application of a model for translation-oriented text analysis*. Rodopi.



- Sun, M., Jiang, B., Xiong, H., He, Z., Wu, H., & Wang, H. (2019). Baidu neural machine translation systems for WMT19. Proceedings of the Fourth Conference on Machine Translation (Volume 2: Shared Task Papers, Day 1),
- 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>
- Varile, G. B., Cole, R., Cole, R. A., Zampolli, A., Mariani, J., Uszkoreit, H., & Zaenen, A. (1997). *Survey of the state of the art in human language technology* (Vol. 13). Cambridge: Cambridge University Press.
- Wang, C. (2017). Machine translation and Chinese college Students' English learning. 2nd International Conference on Contemporary Education, Social Sciences and Humanities (ICCESSH 2017). <https://dx.doi.org/10.2991/iccessh-17.2017.79>
- Wu, Y., Schuster, M., Chen, Z., Le, Q.V., Norouzi, M., Macherey, W., Krikun, M., Cao, Y., Gao, Q., & Macherey, K. (2016). Google's neural machine translation system: Bridging the gap between human and machine translation. *arXiv preprint arXiv:1609.08144*. <https://doi.org/10.48550/arXiv.1609.08144>
- Yoon, C. (2016). Concordancers and dictionaries as problem-solving tools for ESL academic writing. *Language Learning & Technology*, 20(1), 209–229.
- Zhu, J., & Zhang, M. (2021). Analysis of errors in machine translation from Roger T. Bell's Translation Process Model. *International Journal of Social Science and Education Research*, 4(2), 391–399.

Appendix A

Translation course schedule and teaching focus

Teaching content 8 weeks 2hrs × 15 sessions	Teaching and learning activity & materials	Teaching phase & focus
Week 1-1	Pre-task (translating 4 texts)	<i>Phase 1</i> Introducing basic knowledge of translation to students, including definitions of translation, standards of translation, and classification of translation.
Week 1-2	Course outline distribution // Introduction to translation (incl. HT, MT and PEMT)	
Week 2-1	HT exercise from Chinese to English and vice versa	<i>Phase 2</i> Processing HT/PEMT and trying out the recommended MT tools
Week 2-2	PEMT exercise from Chinese to English and vice versa	
Week 3-1	Chinese to English translation activity via HT and PEMT (biography)	<i>Phase 3</i> Comparing HT and PEMT, including advantages and disadvantages of both methods; translating materials are selected by the teacher
Week 3-2	Chinese to English translation activity via HT and PEMT (folk story)	
Week 4-1	English to Chinese translation activity via HT & PEMT (recipe)	
Week 4-2	English to Chinese translation activity via HT & PEMT (short novel)	
Week 5-1	Chinese to English translation activity (prose)	<i>Phase 4</i> Undertaking more translation practices via HT, PEMT or a mix of both; translating materials are chosen by students
Week 5-2	Chinese to English translation activity (prose)	
Week 6-1	English to Chinese translation activity (news article)	
Week 6-2	English to Chinese translation activity (short novel)	
Week 7-1	Post-task	Phase 5 // Reviewing translation methods and discussing students' translations made in the pre- and post-tasks
Week 7-2	Review and reinforcement	
Week 8	Focus group discussions	



The
JALT CALL
Journal
vol. 18 no.1

Chang: Chinese language learners' evaluating machine translation accuracy

Appendix B

Observation sheets to record how students judge MT accuracy in the pre-/ post-tasks



The
JALT CALL
Journal
vol. 18 no.1

Text 1: Magazine article	
Chinese to English translation	Notes
Number of MT tools used	
Using other resources to cross-check MT accuracy	
Process of translation and problem solving	
Translation difficulties	Problem solving strategies
Text 2: Idiom story	
Chinese to English translation	Notes
Number of MT tools used	
Using other resources to cross-check MT accuracy	
Process of translation and problem solving	
Translation difficulties	Problem solving strategies
Text 3: Biography	
English to Chinese translation	Notes
Number of MT tools used	
Using other resources to cross-check MT accuracy	
Process of translation and problem solving	
Translation difficulties	Problem solving strategies
Text 4: Short novel	
English to Chinese translation	Notes
Number of MT tools used	
Using other resources to cross-check MT accuracy	
Process of translation and problem solving	
Translation difficulties	Problem solving strategies

Instructions and source texts for pre- and post-tasks



<p>Task instructions: You have 90 minutes to translate <i>FOUR</i> texts from Chinese to English and English to Chinese. Internet access to make use of any online resources is permitted while translating. The process of your translation will be recorded for the purpose of research data collection; however, it does not include your facial image.</p>	
Source texts for pre-tasks	Source texts for post-tasks
<p>1. AI 人工智慧 科技發展之際,教育的方式得大幅改變才行。隨著AI崛起,教育變革成為必要之事。研究表示,目前的教育太過將人視作陶土,塑形燒製之後就無法改變外形,而非可以重塑的黏土。未來,愈來愈多工作得面臨自動化的挑戰,而最有價值的技能亦將不斷改變。</p> <p>Excerpted from <i>The Economist</i>: https://www.cw.com.tw/article/article.action?id=5082553</p>	<p>1. AI 人工智慧 因應科技發展,教育方式亦隨之改變,一如工業革命造成的影響,工業取代農業。這催生了工廠模式的普及與國家教育,亦即由學校為工廠提供具備合適資格的勞工。AI崛起亦有可能帶來同樣的效應,讓教育變革成為必要之事,並以適性教育來促成教育變革。</p> <p>Excerpted from <i>The Economist</i>: https://www.cw.com.tw/article/article.action?id=5082553</p>
<p>2. 同舟共濟 很久以前,有兩個國家的人民將對方視為仇敵。有一次,兩國的人剛好同坐一艘船渡河。在船上,正當他們差一點發生衝突時,突然遇到暴風雨,眼見船就要翻了,為了保住性命,他們顧不得彼此的仇恨,反而互相救助,合力使船身穩定,最後終於安全到達岸邊。</p> <p>Excerpted from the link: http://www.youjiao.com/e/20180609/5b1bd45a5816d.shtml</p>	<p>2. 同舟共濟 有人問孫武:“戰場上的士兵是否會相互照應?”他回說:“必然的。比如兩個仇人,同乘上一條船渡海,遇到了狂風惡浪,眼看就有葬身海底的危險時,他們就會忘記舊仇,同心協力與風浪搏鬥以避免翻船。連仇人在危險之時尚能同舟共濟,何況沒有冤仇的將士呢?”</p> <p>Excerpted from the link: https://baike.baidu.com/item/%E5%90%8C%E8%88%9F%E5%85%B1%E6%B5%8E/13767</p>
<p>3. Lance Armstrong (AMERICAN CYCLIST) Lance Armstrong, (born September 18, 1971, Plano, Texas, U.S.), American cyclist, who was the only rider to win seven Tour de France titles (1999–2005) but who was later stripped of all his titles after an investigation revealed that he was the key figure in a wide-ranging doping conspiracy while he compiled his Tour victories. Armstrong entered sports at a young age, excelling in both swimming and cycling, and by the time he was a teenager he was competing in triathlons and swimming competitions. Before his high-school graduation the junior national team of the U.S. Cycling Federation had recruited him.</p> <p>Excerpted from Encyclopaedia https://www.britannica.com/biography/Lance-Armstrong</p>	<p>3. Lance Armstrong (AMERICAN CYCLIST) In June 2012 the U.S. Anti-Doping Agency (USADA) alleged that Armstrong and five of his associates had been part of a decadelong doping conspiracy beginning in the late 1990s. According to USADA, Armstrong used performance-enhancing drugs. The allegations resulted in his immediate ban from triathlon competition. In August of that year he declined to enter USADA's arbitration process, which led the agency to announce that it would strip him of all his prizes and awards from August 1998 forward – including his seven Tour de France titles – and enact a lifetime ban from cycling and any other sport that follows the World Anti-Doping Code.</p> <p>Excerpted from Encyclopaedia https://www.britannica.com/biography/Lance-Armstrong</p>
<p>4. An unwritten novel Life's what you see in people's eyes; life's what they learn, and, having learnt it, never, though they seek to hide it, cease to be aware of – what? That life's like that, it seems. Five faces opposite – five mature faces – and the knowledge in each face. Strange, though, how people want to conceal it! Marks of reticence are on all those faces: lips shut, eyes shaded, each one of the five doing something to hide or stultify his knowledge. One smokes; another reads; a third checks entries in a pocketbook; a fourth stares at the map of the line framed opposite; and the fifth – the terrible thing about the fifth is that she does nothing at all.</p> <p>Excerpted from Virginia Woolf's an unwritten novel</p>	<p>4. An unwritten novel The best thing to do against life was to fold the paper so that it made a perfect square, crisp, thick, impervious even to life. This done, I glanced up quickly, armed with a shield of my own. She pierced through my shield; she gazed into my eyes as if searching any sediment of courage at the depths of them and damping it to clay. Her twitch alone denied all hope, discounted all illusion. So we rattled through Surrey and across the border into Sussex. But with my eyes upon life I did not see that the other travellers had left, one by one, till, save for the man who read, we were alone together.</p> <p>Excerpted from Virginia Woolf's an unwritten novel</p>

Assessment of students' translations in Text 1 and Text 3



<p>S1 - Text 1 Chinese to English translation in the pre-task (using one MT tool)</p>	<p>S1 - Text 1 Chinese to English translation in the post-task (using four MT tools)</p>
<p>AI With developments in science and technology, education methods must also undergo significant change. Trends in AI have made education reform a necessity. Research shows that today's education treats people too much like 'hardened clay', with no way to change <u>its their</u> initial form after it has been <u>hardenedmoulded and fired, unlike 'wet clay' that can be reshaped.unable to be re-molded into a new shape.</u> Yet <u>in</u> future, <u>as</u> more and more work <u>has</u> to face the challenge of automation, <u>and</u> the most valuable skills will continually change.</p>	<p>Artificial Intelligence Education methods change in response to technological developments, <u>just</u> such as when industry replaced agriculture in the wake of the Industrial Revolution. This led to the popularization of the factory model and state education, in which the role of schools was <u>to</u> provide suitably qualified workers for factories. The rise of AI is likely to have a similar effect, making the transformation of education an unavoidable necessity, with 'adaptive learning' forming the basis of (<u>driving</u>) education reform.</p>
<p>S1 - Text 3 English to Chinese translation in the pre-task (using one MT tool)</p>	<p>S1 - Text 1 English to Chinese translation in the post-task (using four MT tools)</p>
<p>蘭斯·阿姆斯特朗 (美國自行車手) 蘭斯·阿姆斯特朗 (生於1971年9月18日, 美國得克薩斯州普萊諾), 美國自行車手, 是唯一獲得七次環法自由車賽冠軍 (1999-2005) 的車手。而後來由於 <u>調查檢測</u> 發現他在獲得環法獎牌的同時, <u>也是廣泛大規模禁藥行為</u> 陰謀的關鍵人物, 因此被剝奪所有冠軍頭銜。阿姆斯特朗 <u>小歲時候就</u> 進入運動界, 在游泳與騎 <u>自行車</u> 都表現出色, 以致青年時已經參加鐵人三項和游泳比賽。在他高中畢業前, 美國自由車聯合會的青年國家隊已招募他了。</p>	<p>藍斯·阿姆斯壯 (美國自行車手) 2012年6月, 美國反興奮劑機構 (USADA) 指控阿姆斯壯及其5名同夥參與了從 <u>上世紀</u> 90年代末開始長達 <u>的</u> 10年的興奮劑陰謀。據USADA稱, 阿姆斯壯使用了性能增強藥物。這些指控導致他立即被禁止參加鐵人三項比賽。同年8月, 他拒絕參加USADA的仲裁程序, 這導致該機構宣佈, 將剝奪他從1998年8月起所有獲得的獎品和獎項, 包括他的七個環法冠軍, 並 <u>頒佈</u> 終身禁止他參加自行車賽和任何遵循世界反興奮劑法律的其他運動。</p>
<p>S2 - Text 1 Chinese to English translation in the pre-task (using one MT tool)</p>	<p>S2 - Text 1 Chinese to English translation in the post-task (using four MT tools)</p>
<p>Artificial intelligence At <u>a</u> <u>the</u> time of development of science and technology, education methods have to undergo large changes as well. Along with the rise of AI, the transformation of education has become a necessity. <u>According to R</u> research <u>shows that</u>; in the current education system people are too much treated as 'hardened clay'. After <u>the base of clay+being</u> shaped and fired, there is no way to change its <u>shape form or remodel the clay-</u>, <u>unlike 'wet clay' that can be reshaped.</u> In the future, more and more jobs will have to deal with the challenge of automation, and the most valuable skills will also continue to change.</p>	<p>Artificial intelligence In response to the development of science and technology, education methods have also changed. Just as the impact of the industrial revolution, industry has replaced agriculture. This has led to <u>the</u> popularization of the factory model and national education, i.e. schools provided <u>ed</u> workers with suitable qualifications for factories. The rise of AI may also bring about the same effect, making educational reform a necessity, and using <u>adaptiveappropriate</u> education to facilitate educational reform.</p>
<p>S2 - Text 3 English to Chinese translation in the pre-task (using one MT tool)</p>	<p>S2 - Text 3 English to Chinese translation in the post-task (using four MT tools)</p>
<p>藍斯·阿姆斯壯 (美國的自行車賽車手) 藍斯·阿姆斯壯 (出生於1971年9月18日, <u>位於</u> 美國德克薩斯州普萊諾), 美國自行車賽車手, 他是唯一 <u>贏得</u> 七次 <u>贏得</u> 環法自行車賽冠軍的自行車賽車手 (1999年到2005年)。不過他 <u>之所以</u> 後來失去所有獲得的冠軍 <u>頭銜</u>, 是因為一項調查 <u>暴露</u> 顯示, 他在獲得這些冠軍時, <u>也是</u> 在一個 <u>廣泛大規模</u> 的興奮劑陰謀中 <u>扮演</u> 的主要的人物角色。阿姆斯壯 <u>小時候就</u> 開始 <u>體育</u> 運動, 他擅長游泳與騎 <u>自行車</u>。當他是 <u>成年人</u> 青少年時, 他就參加三項全能和游泳比賽。高中畢業之前, 他 <u>已被</u> 美國自行車聯合會的國家青年 <u>自行車隊</u> 招致 <u>招募</u>。</p>	<p>蘭斯·阿姆斯特朗 (美國 <u>前職業公路</u> 自行車賽車手) 2012年6月, 美國反興奮劑機構 (USADA) 指控阿姆斯特朗及他五名同夥參與了始於90年代末的 <u>長達</u> 十年的興奮劑陰謀。根據美國反興奮劑機構的說法, 阿姆斯特朗使用了提高運動成績的藥物。這些指控導致他立即被禁止參加三項 <u>全能</u> 比賽。同年8月, 他拒絕進入USADA的仲裁程序, 這導致該機構宣布, 將剝奪他從1998年8月起所有的獎品與獎項, 包括他的七個環法自行車賽冠軍。此外, 也終身禁止他參加自行車以及任何遵循《世界反興奮劑條例》的運動比賽。</p>

<p>S3 - Text 1 Chinese to English translation in the pre-task (using one MT tool)</p>	<p>S3 - Text 1 Chinese to English translation in the post-task (using four MT tools)</p>
<p>AI With the development of technology, education has <u>to gone</u> through massive changes. With the rise of AI, a transformation in education is essential. Research has indicated; <u>that the current</u> education system <u>before</u> has treated people too much like 'potter's clay'. After molding and firing, there's no way to change their form, <u>unlike 'wet clay' that can be reshaped or alter them further.</u> In the future, more and more work <u>has to will</u> face the challenge of automatization and the most valuable skill will <u>continue to change be the ability to keep changing.</u></p>	<p>Artificial intelligence In order to adapt to the development of technology, education <u>has must</u> also changed along with it. Just as <u>with the effect of</u> the industrial revolution, industry replaced agriculture. This induced the ubiquitousness of factories as well as national education, namely, <u>that schools would providinge</u> factories with workers possessing <u>the</u> proper qualifications. With the rise of AI, there could be a similar effect, leading education <u>reform</u> to become very necessary, with <u>adaptivepractical</u> education leading to this shift.</p>
<p>S3 - Text 3 English to Chinese translation in the pre-task (using one MT tool)</p>	<p>S3 - Text 3 English to Chinese translation in the post-task (using four MT tools)</p>
<p>藍斯·阿姆斯壯 (美國自行車賽車手) 藍斯·阿姆斯壯 (生於1971年9月18日號出生,美國德克薩斯普萊諾,德克薩斯,美國)美國自行車賽車手,是唯一獲得七個位環法自行車賽冠軍頭銜的車手(1999-2005),但在一個調查顯示視察露出他在參與環法的同時,亦是<u>大規模</u>興奮劑陰謀的關鍵人物,因此他的一切頭銜皆被剝奪。阿姆斯壯年輕的時候就<u>就</u>開始參與運動,游泳,自行車<u>方面</u>都很成功。到他青少年<u>就</u>參與三鐵和游泳比賽。在他<u>高中畢業高中</u>之前,<u>已</u>他被美國自行車聯合的國家少年隊招募。</p>	<p>蘭斯·阿姆斯特朗(美國自行車選手) 2012年6月,美國反禁藥組織(USADA)指控阿姆斯特朗從1990年代末開始跟他的五個同夥<u>參與參加了</u>的長達十年的興奮劑陰謀。根據USADA所說,阿姆斯特朗使用了增強性能的藥物。這些指控導致他立即被禁止參加鐵人三項比賽。那年8月,他拒絕進入USADA的仲裁程序,導致該組織宣布將<u>從1998年8月起</u>剝奪他<u>從1998年8月起</u>的所有獎杯和獎項-包括他的7個環法自行車賽冠軍-並終身禁止<u>參加</u>騎自行車以及遵循《世界反興奮劑條例》的任何其他運動。</p>
<p>S4 - Text 1 Chinese to English translation in the pre-task (using one MT tool)</p>	<p>S4 - Text 1 Chinese to English translation in the post-task (using four MT tools)</p>
<p>Artificial Intelligence As technology develops, education methods must undergo substantial change. With the rise of AI, education reform has become necessary. Research shows that current education regards people too much like '<u>hardened clay</u>'; once molded and fired, there is no way to change its shape, <u>unlike 'wet clay' that can and it cannot</u> be reconstructed. In the future, more and more jobs must face the challenge <u>of</u> automation, and the most valuable skills will also continuously change.</p>	<p>Artificial Intelligence In response to technological development, educational methods are also changing, just like the effects brought about by the Industrial Revolution when industry replaced agriculture. This has induced the universality of the factory model and national education, namely, schools providing factories with workers with suitable qualifications. The rise of AI may also bring about the same effect, making educational reform a necessity, and using <u>adaptive-well-suited</u> education to promote educational reform.</p>
<p>S4 - Text 3 English to Chinese translation in the pre-task (using one MT tool)</p>	<p>S4 - Text 3 English to Chinese translation in the post-task (using four MT tools)</p>
<p>藍斯·阿姆斯壯 (美國腳踏自行車賽車手) 藍斯·阿姆斯壯(1971年9月18日出生於美國德克薩斯州普萊諾)是美國<u>腳踏自行車</u>賽車手。他是唯一獲得7個環自行車賽冠軍(1999-2005)的車手,但後來<u>他</u>被剝奪了所有冠軍頭銜,原因是調查顯示<u>他</u>是在取得環法自行車賽<u>成績勝利</u>時,<u>也是</u>進行<u>廣泛大規模</u>興奮劑陰謀案件的關鍵人物。阿姆斯特朗很小的時候就參加體育運動,在游泳和騎<u>腳踏自行車</u>方面都表現出色,到了十幾歲的時候,<u>他</u>就參加了三項全能和游泳比賽。在他高中畢業之前,美國<u>腳踏自行車</u>聯合會的青年國家隊就招募了他。</p>	<p>蘭斯·阿姆斯特朗(美國自行車手) 2012年6月,美國反興奮劑機構(USADA)指控阿姆斯特朗及其五名同夥<u>參與參加了</u>從1990年代末開始長達十年的興奮劑陰謀。根據USADA的說法,阿姆斯特朗使用了增強性能的藥物。這些指控導致他立即被禁止參加三項全能賽。同年8月,他拒絕進入USADA的仲裁程序,這導致該機構宣布將剝奪他從1998年8月起的所有獎品和獎項,包括他的七個環法自行車賽冠軍,並終身禁止參加自行車以及任何其他遵循《世界運動禁藥管制規範》的運動。</p>

Note: The 'track changes' have been left in to show the corrections to the translation of students.