



Translation Tasks for Learning Collocations: Effects of Machine Translation plus Post-editing and Sight Translation

翻譯活動對於搭配詞學習的助益:以機器翻譯後編輯與視譯為例

Chia-chien Chang¹  · Masaru Yamada²

Received: 29 December 2019 / Revised: 7 May 2020 / Accepted: 4 June 2020

Published online: 27 June 2020

© National Taiwan Normal University 2020

Abstract

This study tested the effectiveness of employing translation tasks to enhance students' use of collocations in spoken English. Two intact classes from a university in Taiwan participated in this study. Both classes first were taught a reading text with a focus on comprehension. Afterwards, while the control group had a free discussion on topics related to the reading, the experimental group completed two translation tasks. The first was a post-editing task in which students edited a Chinese translation of the reading generated by Google Translate. The second was a sight translation task in which the students orally translated part of the Chinese translation back into English. In the posttest, the experimental group used more target vocabulary correctly than the control group. Through a questionnaire, students in the experimental group reported that the two translation activities were helpful for learning to use collocations correctly in spoken English.

摘要

本研究旨在探討英語教學時使用翻譯活動是否有助於英語學習者口說時正確使用搭配詞。研究參與者為兩班剛入學的大一新生,學生先針對一篇高中程度的英文文章進行聽力與閱讀練習。之後,對照組針對文章主題進行自由討論活動,實驗組則進行兩項翻譯活動練習。第一項活動為機器翻譯後編輯練習,學生對照英文原文與Google Translate機器翻譯自動產出的中文譯文,進行挑錯與更正。第二項活動為視譯,學生使用文章的中譯版本,邊閱讀中文譯文,邊以口頭方式將內容重新翻回英文。後測結果顯示,實驗組比對照組在口說時更能正確運用課文中的搭配詞。問卷結果亦指出,學生認為兩種翻譯活動有助於他們學會並使用英文詞彙。

Keywords Machine translation post-editing · Sight translation · Language learning · Speaking · Collocation

✉ Chia-chien Chang
chiachienchang@ntu.edu.tw

Introduction

The role of translation in language learning and teaching has long been controversial as it is often associated with the outdated grammar translation approach. However, in the last decade, we have seen a revival of translation in language teaching, with the argument that translation could serve as an effective communicative activity given the right parameters [1]. It is increasingly recognized that translation can be a useful tool for language teaching, especially in single-language classes taught by bilingual teachers [2]. There is also an increasing awareness of the role of translation in language teaching among scholars in translation studies as well as second language acquisition [3–7].

More recently, the advent of Google Translate, a new system based on artificial intelligence (AI), has revolutionized translation. With an estimated 60% reduction in translation errors from its machine translation predecessor [8], the new Google Translate yielded quality improvements, leading some to pronounce, “[T]his is a threat to translation industry” [9, p. 8]. These developments in machine translation have accelerated academic studies on the combination of machine translation and language learning [10, 11].

The present paper is thus motivated to test the effectiveness of using two translation tasks for language learning: machine translation plus post-editing and sight translation. As the focus of vocabulary instruction shifted from individual words to context and collocations, we think that it is time to take advantage of translation activities that are in line with the cognitive processing of language learners. Unlike the bilingual word lists traditionally associated with the grammar translation method, the two activities proposed in this study aim to help students become familiar with vocabulary in context, with a special focus on decreasing collocation errors.

Collocations in Language Learning

Frequent use of collocations by language learners is, to an extent, an indicator of proficiency that can distinguish advanced learners from beginners; however, error frequency does not necessarily decrease with proficiency [12]. Compared with single-word vocabulary, learners seem to be less attentive to multiword expressions such as collocations and also have more difficulty transferring them from input to output [13]. Nesselhauf [14] indicates that the learner L1 has a degree of influence on learning target language collocations, regardless of the difficulty of acquisition and the learner’s proficiency. Collocations seem to be an inherent problem for L2 learners [14–16], especially when collocations do not have L1 equivalents.

As the issue of collocation learning has been recognized, there is an increasing number of empirical studies on effective instructional methods involving translation that deals with both L1 and L2. Among the possible alternatives is the use of translation activities that take advantage of cross-linguistic influence on lexical learning. However, the results are somewhat mixed. For example, Eyckmans, Boers, and Lindstromberg [17] tested the effectiveness of instructing students to check for non-congruency with their L1 and found that it was not effective for enhancing recall when compared with instructing students to look for alliteration in collocations. On the other hand, Laufer

and Girsai [18] used translation into L1 and L2 in class and found that, for helping students learn both collocations and single words, such contrastive form-focused tasks were more effective than were meaning-focused instruction and non-contrastive form-focused instruction.

In this study, we aim to take advantage of the features of translation tasks both as authentic communication activities and as cross-linguistic activities that can raise learners' awareness of the similarities and differences between the collocation patterns in the target language and those in their mother tongue. We designed two translation activities for the purpose of helping learners learn to use collocations correctly for speaking. The first task is machine translation plus post-editing, with students editing a Chinese translation of an English reading generated by Google Translate. The second is a sight translation task in which the students orally translate part of the Chinese translation back into English. We elaborate on the rationale for using these two activities in the next section.

Machine Translation plus Post-editing and Sight Translation

The authors designed two translation tasks to enhance students' learning of collocations for speaking: machine translation plus post-editing and sight translation. The machine translation plus post-editing (hereafter referred to as post-editing) task was proposed based on Schmidt's [19] noticing hypothesis, which holds that extra attention paid to vocabulary words can help convert input into intake. We also hypothesized that the identification and correction of translation errors can raise students' awareness of the similarities and differences in form and meaning between the vocabulary words in their L2 and those in their L1 and hence counteract interference mistakes [20]. While post-editing is often incorporated into professional translator training to help prepare trainee translators for the increasing use of technology in the translation industry, post-editing for language learning has been used as a "bad model" for learners to identify and correct raw machine translation output errors [21]. Even though the quality of machine translation has been constantly improving, especially after the introduction of neural machine translation, there are still errors prominent enough for students to be able to identify in the post-editing process [21]. The interaction and autonomy promoted by the discussion of linguistic problems with teachers and other learners may also contribute to the acquisition of the vocabulary words [21]. In addition, the post-editing activity is of high involvement load, as defined by Laufer and Hulstijn's [22] task-induced involvement load hypothesis. While the PE task mainly involves evaluation of the machine translation output, the learners need to decide whether the translation is appropriate and search for better translation alternatives when necessary. An additional advantage of using this activity is being able to check comprehension, as students need to understand the source text correctly in order to successfully edit the errors in the machine translation output.

The sight translation task has been proposed based on Swain's [23] output hypothesis, which posits that language acquisition can be facilitated when learners are pushed to expand or restructure their linguistic resources. Sight translation is "the oral translation of a written text," in which a translator reads silently a written text in one language, processes the meaning quickly, and immediately translates the text out loud

into another language [24, p. 144]. It is often included as part of interpreter and sometimes translator training to develop “quick reactions, thorough comprehension, and flexible oral skills” [24, p. 146]. However, instead of translating a new text, the sight translation task we proposed in this study is for students to take a translated version of a text they have read and orally render it back into the original language. While other activities aiming to encourage learners to use newly learned expressions in their speaking have to devise creative and sometimes complicated ways to help learners recall expressions [25], this type of “back translation” activity provides a quick and clear means of pushing learners to expand their range of expressions as they apply new vocabulary to translate from L1 into L2. By using the sight translation task, the cognitive load for generating content on their own when speaking is also reduced so that learners can focus their effort more on memorizing, recalling, and automatizing the use of vocabulary words. As with the post-editing activity, the sight translation task is of high involvement load [22], as completing the task involves finding the vocabulary words needed to complete the task, searching for vocabulary words that express the L1 concept, and evaluating whether the vocabulary words found are appropriate.

The Study

The study was designed to test the effectiveness of using translation activities for language learning and answer the following research question: Can post-editing and sight translation help students learn and use collocations? The effectiveness of the instructional method was measured by the number of target vocabulary words used correctly by the learners in a real-time speaking task.

Participants

Two intact English-as-a-foreign-language (EFL) classes at a top university in northern Taiwan participated in this study. The experimental group consisted of 27 students and the control group consisted of 24 students. At the time of the study, all students were in their first month of university. As these students had been admitted to the university after passing a competitive entrance exam which included a written test of English, it was assumed that they could handle the secondary-level English learning materials to be used in this study and should already have had the target items as part of their passive vocabulary. To account for possible differences in English speaking proficiency between the two groups, the General English Proficiency Test (GEPT) intermediate-level speaking test was administered a week before the study to be used later as a covariant.

Teaching Materials

The teaching materials were based on an article about procrastination among college students taken from a widely used senior high school English textbook in Taiwan [26]. Eleven vocabulary words along with their collocations were picked as target items for learning: *extremely important*, *master a skill*, *feel overwhelmed*, *drop out of school/quit school*, *have poor time management skills*, *break a bad habit*, *reward yourself*, in

order of **priority**, **budget your time**, and **create a schedule**. The first eight items selected each have a corpus frequency of over 100 occurrences in the Corpus of Contemporary American English (COCA). The collocation item *in order of priority* was chosen because it is relatively long and may be more challenging for learners. The two additional collocation items, *create a schedule* and *budget your time*, were included because the machine translations of the items were unsatisfactory and therefore possibly more noticeable in the machine translation post-editing activity. The target items were highlighted in the reading text to draw students' attention to them. However, instead of highlighting the whole collocation, only the keywords were highlighted (e.g., **extremely important**) to resemble the typical layout of textbooks.

To prepare the machine translation text for post-editing, a Chinese translation of the English reading was generated by Google Translate one month before the study was conducted. Two translation teachers reviewed the machine translation output and concluded that the translations were generally appropriate, with only two obviously inappropriate translations (*quit school* 戒掉學校 and *budget your time* 預算你的時間). However, considering that not all the translations were completely idiomatic, the student editors had opportunities to improve some translations.

On the handout for the post-editing task, the translation was included beneath each of the six paragraphs of the reading, each paragraph was numbered, the eleven vocabulary items were highlighted, the target words were bolded, and the meaning unit of each target word underlined. Their Chinese equivalents were also underlined to facilitate comparison between English and Chinese. Under each paragraph was a box with two options for the students to check: (1) "I feel that the AI translation is accurate and natural" or (2) "I feel there is something wrong. We should say it in Chinese this way: (provide the appropriate translation)."

The sight translation handout had a similar format to the machine translation handout. The reading was divided by paragraph, and the Chinese translation was taken from the teacher's manual provided by the publisher of the textbook. Since students were to back-translate the text from its Chinese translation into English, each Chinese paragraph was followed by the equivalent English paragraph. The target words were bolded and the meaning unit of the target words underlined. Their Chinese equivalents were also underlined to facilitate comparison between English and Chinese. Since the purpose of using a sight translation task in this study is only to push learners to use the target vocabulary in context, learners were instructed to use the target vocabulary to convey the meaning of the underlined phrases but could convey the message in their own words for the rest of the underlined sections.

Pretest and Posttest Design

A speaking test was designed to assess the students' ability to use the target items. A message in which the eleven target items can be used was created and a context was provided for students to use the target items: *You have a friend who is a serious procrastinator. He often puts off handing in assignments or preparing for tests. Please listen to the following one-minute Chinese message and then convey the ideas in the message in English to your friend.* The test was administered to the two groups before the lesson as a pretest and again at the end of the lesson as a posttest. In the posttest, students were told to use as many target items from the reading as possible.

To prevent individual differences in note-taking skills from influencing performance, students were not allowed to take notes while listening to the Chinese message. To help them remember the content of the one-minute talk and the target items they needed to use in their response, they were given a sheet of paper with ten pictures in the left column representing the ideas presented in the message and ten Chinese prompts for the target items in the right column. After listening to the Chinese message, they were given one minute to prepare before they started recording their answer.

The translation prompts used in the speaking test were different from those used for the machine translation post-editing and sight translation. In fact, we deliberately used three different translations of these target vocabulary words in the machine translation post-editing, sight translation, and pretest/posttest, respectively. Since we used translation as prompts in our posttest, we tried to decrease an advantage the experimental group may have had over the control group by using translations that were different from the ones they were exposed to during instruction (e.g., for *extremely important*, pretest/posttest: 至關重要, machine translation: 非常重要, sight translation: 極度重要). The exposure to a variety of translations for the same vocabulary words was also intended to counteract learners' tendency to link vocabulary words exclusively to one L1 equivalent.

Procedure

The data collection was completed in a regular class, which had been presented as a speaking workshop aimed to help the students transition away from the high school strategy of reading to learn new words and grammar toward the college strategy of using what they already know for speaking and listening. The class for the control group and that for the experimental group were both conducted by one of the authors within the same week. The speaking pretest was administered right before the workshop, which was divided into two parts, each lasting 25 to 30 minutes. The first part was identical for both groups. The students were first given a handout with the reading on procrastination and were told that they would learn how to practice reading, listening, and speaking using the article. The students read the article for two minutes. Then, they were shown a video on procrastination. They were told beforehand to be ready to answer a few questions based on the content of the video, including questions on the definition of procrastination, the problems the speaker experienced, and her solutions to the problems. After students watched the video, the instructor briefly discussed the answers to the questions. The instructor then presented the article paragraph by paragraph with an audio file while the students worked in pairs to summarize each paragraph and talk about their opinions after listening. After this activity, the students were given a ten-minute break before the second part of the workshop started.

During the second part of the workshop, entitled "Expand your vocabulary," students were explicitly told to use the words from the reading for speaking. The two groups received a different treatment: The control group had a free discussion on topics related to the reading while the experimental group did the two translation tasks.

During the free discussion, the control group continued to receive meaning-based instruction and completed two tasks in pairs. The first task was to summarize the reading and discuss personal experiences with procrastination. The second task was to

come up with solutions to the problems and share with classmates. Both prompts were designed to enable students to use the target items in their answers, and the students were explicitly encouraged to use as much of the target vocabulary highlighted in the reading text as possible. After each pair discussion, a few students were invited to share their thoughts with the whole class, using as many highlighted vocabulary words as possible in their presentation. These interactions took place using English only.

For a similar amount of time, the experimental group completed two translation tasks. The first was a post-editing task in which students edited the Chinese translation of the reading generated by Google Translate. The activity was presented as a human vs. AI competition where the students needed to outwit the AI. Students worked in pairs and focused on the Chinese translations of the target vocabulary highlighted in the reading paragraph by paragraph. They were asked to check if the machine translation outputs were appropriate and, if not, to provide alternative translations from the perspective of human translators. In the second task, sight translation, the students orally rendered part of the Chinese translation back into English. Students were provided a different Chinese translation, taken from the teacher's manual provided by the publisher of the high school textbook, and were asked to work in pairs to sight translate the reading paragraph by paragraph from Chinese back into English. During the sight translation of each paragraph, students needed to provide only a loose translation of the Chinese text that conveyed the main ideas. However, when it came to the underlined sentences, they were required to include in their translation the highlighted vocabulary items from the English reading. If a student used a different word or made a mistake related to the target vocabulary, that student's partner would immediately correct it.

After these different treatments, another video on tips to overcome procrastination was shown to both classes as a wrap-up to the lesson. At the end of the instruction, all handouts were collected, and the students took a posttest, which was identical to the pretest except that this time the students were instructed to use as much of the highlighted vocabulary from the reading as possible. After the posttest, a questionnaire was administered to both groups. The students were asked to answer whether they had used the vocabulary from the reading and, if not, the reasons they had not used them. Additionally, the experimental group was asked whether the two translation tasks helped them use the target vocabulary in the posttest speaking task.

The complete procedure can be summarized as follows: Experimental group: pretest>reading input>video input and oral output>reading with audio input and oral output>break>**post-editing task**>**sight translation task**>video input and oral output>posttest>questionnaire. Control group: pretest>reading input>video input and oral output>reading with audio input and oral output>break>**free discussion task 1**>**free discussion task 2**>video input and oral output>posttest>questionnaire.

Results and Discussion

Pretest and Posttest Results

All the recordings from the control and the experimental groups were fully transcribed for grading. There were in total eleven target items chosen for the study, but two of

them (*drop out of school* and *quit school*) are synonymous and were treated as one item in the pretest/posttest. As each correct answer was worth one point, the maximum score was ten points. The grading criteria for both the pretest and posttest were as follows:

1. Since the purpose of the study was to prompt the students to activate the vocabulary learned from the reading, only when the target word/phrase was used would the students be given a point, even when the student used a suitable word in place of the target word. For example, the answer “create a schedule” received one point but “make a schedule” did not. Similarly, the answer “extremely important” received one point but “very important” did not. Meanwhile, either “quit school” or “drop out of school” could be used interchangeably in the test and receive one point. In addition, using “reward” as a noun or using “prioritize” instead of “priority” also earned one point since both were used in the two videos students watched.
2. To be marked as correct, target items needed to have been used correctly. For example, the answer “you may feel overwhelmed” or “it may be overwhelming” received one point, but “you may feel overwhelming” did not.
3. The misuse of article the/a or plural/singular form was disregarded as it is beyond the purpose of the study to help students correct these tenacious mistakes, such as “*a(n) extremely important thing” and “give yourself some *reward(s).” However, when the error was part of a set phrase, it was counted as an error, such as “drop out of *the school,” or “quit *the school.”

Table 1 shows the results of the pretest and posttest for both groups. In the pretest, the experimental group received an average score of 0.78 while the control group received an average score of 0.46. To account for possible differences in English-speaking proficiency between the two intact classes, a one-way analysis of covariance (ANCOVA) was conducted, using the GEPT intermediate speaking test scores as a covariant. Adjusted for initial speaking proficiency, the ANCOVA results ($F(1,48) = 1.125, p = 0.294$) indicate that the two groups did not differ significantly in their pretest performance.

In the posttest, the experimental group received an average score of 6.26 while the control group received an average score of 3.04. A one-way analysis of covariance (ANCOVA) was conducted using the GEPT intermediate speaking test scores as a covariant. The results ($F(1,48) = 35.551, p = 0.000$) show that there was a significant difference in learners’ performance in the posttest between the two groups, adjusting for initial speaking proficiency. The partial eta squared value (0.425) indicates that the effect size is moderate.

Table 1 Pretest and posttest results

Group	Pretest, max. = 10		Posttest, max. = 10	
	Mean	SD	Mean	SD
Experimental, $n = 27$	0.78	0.80	6.26	1.97
Control, $n = 24$	0.46	0.78	3.04	1.49

Table 2 Posttest results for individual items

	Target items	Experimental group mean (SD)	Control group mean (SD)
1	Extremely important	0.85 (0.36)	0.08 (0.28)
2	Master a skill	0.48 (0.51)	0.25 (0.44)
3	Have poor time management skills	0.48 (0.51)	0.25 (0.44)
4	Feel overwhelmed	0.78 (0.42)	0.42 (0.50)
5	Drop out of school/quit school	0.67 (0.48)	0.54 (0.51)
6	Break a bad habit	0.78 (0.42)	0.38 (0.49)
7	In order of priority	0.56 (0.51)	0.29 (0.46)
8	Budget your time	0.41 (0.50)	0.08 (0.28)
9	Create a schedule	0.44 (0.51)	0.13 (0.34)
10	Reward yourself	0.81 (0.40)	0.63 (0.49)

To gain a deeper understanding of the results, we conducted an error analysis of each vocabulary item from the posttest data. Table 2 shows the differences in performance between the experimental group and the control group on each vocabulary item.

In addition, to see if correcting the erroneous translation in the machine translation output indeed left a strong impression on the students, we calculated the number of corrections made by students during the post-editing task on each item. Table 3 shows the number of machine translation post-editing instances from the 15 student pairs¹ in the experimental group. The machine translation output of T5(2) (quit school: 戒掉學校) and T8 (budget your time: 預算你的時間) was clearly wrong and was corrected by all student pairs. The machine translation output for T6 (break a bad habit: 打破這個壞習慣) is a collocation that sounds unnatural in Chinese and hence was corrected by most pairs.

Based on the above analyses, a more detailed discussion of student performance on each target item follows.

1. **Extremely important:** In the pretest, students from both the experimental and the control groups used “very” or “quite” to intensify the level of importance. In the posttest, 23 students (85%) from the experimental group correctly used “extremely” to intensify “important” while only three continued to use “very” or “quite” and one student used “very extremely.” However, in the control group, only two students (8%) used “extremely” while 16 students (67%) continued to use “very/ really/pretty” or used no intensifier.
2. **Master a skill:** In the pretest, students from both groups tended to use the generic verb “have” as in “have the ability/skill.” Many were affected by the Chinese prompt 掌控 (hand+control) and incorrectly used verbs such as “control,” “hold,” or “handle” the skill. In the posttest, 13 students (48%) in the experimental group used the collocation “master the skill” correctly compared with only six students (25%) in the control group.

¹ Among the 27 participants in the experimental class, three participants were paired with students who were ineligible for the study because they were senior students, so there were in total results from 15 pairs.

Table 3 Number of post-editing instances

Target item	T1	T2	T3	T4	T5 (1)/(2)	T6	T7	T8	T9	T10
Number of editing	6	2	3	9	1/15	13	6	15	8	0

3. Have poor time management skills: In the pretest, almost all students were affected by the Chinese prompt “不佳” (not good) and used “bad” time management skills. No students from either used the word “poor.” In the posttest, 13 students (48%) from the experimental group and only six students (25%) from the control group used “poor” correctly in context.
4. Feel overwhelmed: In the pretest, more than half of the students from both the control and experimental groups were affected by the Chinese prompt “無法承受” (can’t bear) and used words such as “can’t afford the burden,” “feel unaffordable,” or “cannot be able to bear.” In the posttest, 21 students (78%) from the experimental group were able to use “feel overwhelmed” correctly while ten students (42%) from the control group were able to do so. The students seemed to have embraced this phrase. When the control group was asked to use the target items in their free discussion, the researcher observed that “feel overwhelmed” was one of the most commonly used collocations, along with “break the bad habit.”
5. Drop out of school/quit school: Students who chose to use either of these two expressions were awarded one point. Eighteen students (67%) (13 “quit,” 5 “drop out of”) from the experimental group and 13 students (54%) (7 “quit,” 5 “drop out of,” 1 both) from the control group used either item correctly. It is interesting to note that many more students from the experimental group used “quit school,” possibly indicating a stronger impression left from correcting the machine translation errors.

It should also be noted that “to drop out of school” was already a familiar expression to many of the students, being used by many students in the pretest. It is interesting, however, that quite a few students (ten from the control group, three from the experimental) used “drop out of *the school” or “quit *the school,” inserting the article “the” into the expression in the posttest. Many also did so in the pretest, giving support to Nesselhauf’s [14] claim that articles should be emphasized in the learning of collocational expressions.

6. Break a bad habit: The target item “break” appears in the reading as “you may want to do something to **break** this bad habit.” It seems that the collocation of “break” and “bad habit” became salient once they were brought to students’ attention as they are incongruent with the Chinese pattern of collocation. No students used the verb “break” in the pretest. In the posttest, 21 students (78%) from the experimental group used “break” correctly while nine students (38%) in the control group did. Six students (25%) from the control group used “change” compared with only one (4%) from the experimental group. It is interesting to note, however, that four students from the control group used “break out” or “break down” and one student from the experimental group used “break up” in the posttest, showing that, even though they noticed the collocation, they mistakenly connected the verb “break” with its other collocations. It is possible that

some students in the experimental group may also have made the same mistake during practice but were corrected by their partner during the sight translation activity, as is discussed in the next section on questionnaire results.

7. In order of priority: As this phrase is the longest among the ten target items, it is understandable that students had a harder time memorizing the whole collocation. In the posttest, 15 students (56%) from the experimental group used “in order of priority” correctly while as many as ten students (37%) made a variety of minor errors, including “in priority” and “in order for priority.” Seven students (29%) from the control group used the word “priority” correctly, and among them, only one used “in order of priority” while others used alternative forms such as “set the priority,” “prioritize,” and “choose what is your top priority.” It seems that students in the control group tended to focus on the single word “priority.” Students in the experimental group, on the other hand, were able to pay more attention to the form of the phrase “in order of priority,” even though many still had problems producing the complete multiword unit accurately.
8. Budget your time: This target item received the lowest score among the ten vocabulary items in the posttest both for the experimental group and the control group. Eleven students (41%) from the experimental group used “budget” correctly while only two students (8%) from the control group did. Instead, many students used “manage/allocate your time” as acceptable alternatives while a few erroneously used “assign/separate/divide your time.”

These low scores are even more interesting considering that the word “budget” was mistranslated by the machine translation and all student pairs in the experimental group corrected it during the post-editing activity. As a result, one might expect the word to have left a stronger impression on the learners. It is possible that it is hard to associate the Chinese prompt “分配” with the verb “budget.” Some students also mentioned in the questionnaire that they found it weird to use the word “budget” with time as the expression in Chinese is almost exclusively associated with money. For vocabulary items like this, following the machine translation task with a contrastive analysis between the two languages may help students activate the vocabulary.

9. Create a schedule: In the pretest, only one student in the experimental group used the collocation “create a schedule,” and no student from the control group did. In the posttest, 12 students (44%) in the experimental group used “create” correctly while only three students (13%) in the control group did so. It is interesting to note that most people chose to use “make a schedule” in both the pretest and posttest. Even in the experimental group, there were still many people using “make” instead of “create” in the posttest. This could be because students were more used to using “make” and were unable to change the word choice under the pressure of a speaking test.
10. Reward yourself: In the pretest, only six students from the experimental group and three from the control group used “reward” correctly in context. Many students confused the word with “award,” as in “you can award yourself.” In addition, because of the translation prompt where “reward” was used as a noun in Chinese as in “給自己一點獎勵” (give yourself some rewards), almost all students who used “reward” correctly used it as a noun. In the posttest, 22 students (81%) in the experimental group used “reward” correctly, and among them, 16 students

(73%) used it as a verb. In the control group, 15 students (63%) used “reward” correctly, but among them, only seven students (47%) used it as a verb. Although either usage was considered correct based on our grading criteria, it seems that students who completed the translation activities more effectively resisted the effect of the Chinese translation prompt on their choice of part of speech.

Learners’ Perceptions of the Effectiveness of Using Translation Tasks for Learning Collocations

To understand students’ perceptions of the two translation activities, students in the experimental group were asked in the post-study questionnaire to rate on a 5-point Likert scale the extent to which they agree or disagree with the statements that the two translation activities can help them remember and use the target vocabulary, with 5 being strongly agree and 1 being strongly disagree. As shown in Table 4, students reportedly found both activities useful for their learning (3.63 and 4.04). In the following sections, we will discuss students’ perceptions of these two activities respectively.

Post-Editing

Regarding the machine translation post-editing activity, of the 27 students, four gave a rating of 5 and 12 gave a rating of 4 (59%) while eight gave a rating of 3 and three gave a rating of 2 (41%). Most students commented that the process of identifying inappropriate translation and coming up with a better translation helped them dig deeper into the meaning of the English words and hence retain the words, as demonstrated by the following comments:

The process of improving machine translation allowed me to gain deeper impressions of the words I am not familiar with. (E17, posttest score 8)

When I saw a clumsy Chinese machine translation, I would go back to the English, then think about why the machine translation would have rendered the translation as such. (E21, posttest score 5)

Ten students mentioned that correcting the mistakes in the machine translation created a deeper impression of those words. As suggested in the comment by E21, these students would start with the Chinese and, when they saw something awkward, look at the English and try to imagine what caused the machine to generate such an inappropriate

Table 4 Rating of the usefulness of the two translation activities

Group	Post-editing		Sight translation	
	Mean	SD	Mean	SD
Experimental, <i>n</i> = 27	3.63	0.88	4.04	0.76

translation and how to translate the words appropriately. Nevertheless, if the original translation was appropriate, it did not leave a particular impression on the students, as indicated by the following two students who both gave a rating of 3 to the activity:

If there were mistakes in the machine translation, I would be able to remember it better. (E24, posttest score 6)

Odd machine translation outputs would leave an impression on me, but the correct ones, not so much. (E25, posttest score 9)

However, other students seemed to experience a different cognitive process as they raised the issue that, during the post-editing activity, they were more focused on correcting the Chinese translation so that the Chinese passage was more natural and fluent. As a result, their attention was on the Chinese, not the English. Of these 11 students, three cited the same reason for not finding the activity particularly helpful. As shown in the comments below, these students apparently paid more attention to the Chinese translation than to the English source text:

I did not really pay much attention to the original English text. I mainly focused on whether the Chinese is correct or not. (E9, posttest score 7)

Some machine translation outputs were odd, so I put a lot of effort into coming up with appropriate Chinese and hence it was difficult to remember the English. (E14, posttest score 5)

In other words, compared with providing students with a correct version of translation, these students found the poor translation detrimental to their learning process. Another student who gave a rating of 3 to the activity even stated that he found the machine translation post-editing task unhelpful because “machine translation output is often inaccurate, or downright incorrect” (E11, posttest score 6).

In short, it seems that identifying mistranslation is conducive to remembering the target vocabulary while correcting the translation seems to cause some students to shift their attention to the Chinese and hence may even be detrimental to retaining the vocabulary.

Sight Translation

Regarding the sight translation activity, fifteen gave a score of 4 and seven a score of 5, indicating that 81.5% of the students found sight translation helpful for remembering the target items. Eleven students cited the reason being that the process of actively recalling, searching, and producing the vocabulary words helped them remember the target items better, giving support to the involvement load hypothesis. For example:

Translating and actually using new words in sentences helped me better incorporate them into my vocabulary knowledge. (E9, posttest score 7)

I would be able to remember the new words because I actually had to use them. (E10, posttest score 9)

Another common reason offered was that the experience of being corrected by a partner while practicing helped reinforce their learning. As students had to pair up in the sight translation activity, when one person used a different word instead of the target item or used the target item but made a mistake regarding the collocation or other aspects of the target item, that person's partner would correct it. Six students (E3, 7, 12, 14, 18, 19) mentioned that making mistakes and being corrected helped them remember the correct usage better. Among them, three described specific benefits of this corrective feedback:

During sight translation practice, I have to actively translate the Chinese into English. After being corrected by my classmate again and again, I was able to remember complete phrases or sentences. I could use them better in conversations. (E7, posttest score 10)

During sight translation practice, my usages may differ from those in the textbook. In these circumstances, my classmates would correct me, allowing me to have a stronger impression of these words and phrases. (E12, posttest score 8)

Because there was someone to remind me when I used different words (those I am more familiar with.) (E19, posttest score 8).

Some students (E21, 23, 27) mentioned that, when doing the sight translation activity, they noticed the gap in their knowledge when they were not able to sight translate a particular vocabulary item and hence were able to fill the gap with practice, as indicated in the following comments:

Since I was not only passively taking in the words—I had to say the words out loud—I could discover what I was still unfamiliar with. (E21, posttest score 5)

Sight translation practice helped me identify what I was still unsure of, and make stronger connections between Chinese and English, which aids in memory. (E23, posttest score 9)

Two students (E3, E17) also mentioned that the sight translation practice helped them establish the link between the Chinese phrases and the English vocabulary words learned from the reading, promoting association between the Chinese prompts and the English words during posttest.

Because while reading the Chinese source text [during the sight translation exercise], I could instantly think of the corresponding English, when doing the posttest, I was also able to instantly connect the Chinese

prompts with the Chinese used in the sight translation. (E17, posttest score 8)

Although we had tried to make the translations used in the sight translation task and in the posttest different, this inter/intra-lingual connection may have given an advantage to students in the experimental group. Nevertheless, the establishment of the connection can benefit learners when they need to translate from Chinese into English in the future.

On the other hand, there were still some students who found the activity not so helpful. One student gave a score of 2 and four students gave a score of 3 regarding the usefulness of sight translation. One student (E1, posttest score 5) who gave the score of 3 mentioned that the task was too difficult. The student who gave a score of 2 stated that he needed to come up with the translation of the whole sentence, so he was unable to focus on the target item only.

I had to think about other things. It was not easy to focus on target item. (E02, posttest score 4)

In general, questionnaire results indicate that students reacted positively to both the input-oriented post-editing task and the output-oriented sight translation task. One student (E07), who received a score of 0 in the pretest but the maximum score of 10 in the posttest, gave additional comments at the end of the questionnaire, which summarize the potential benefits of both activities nicely:

I like the effects brought by sight translation. It was so effective! Post-editing was also a brand new experience. In our old way of learning, we are used to merely understanding the meaning of words in isolation, but [with post-editing] we see how these words work within sentences. We need to consider the entire context. We need to correctly identify the meaning of these words and their parts of speech in context. I felt I was truly learning these words and it felt great. (E07, posttest score 10)

Conclusion and Pedagogical Implications

This study tested the effectiveness of using translation tasks to enhance students' learning of collocations for speaking. The machine translation post-editing was introduced as the input-oriented activity, which was intended to trigger noticing differences between L1 and L2 in both their forms and meaning, while the sight translation activity was employed for output purposes. The results show that a combination of these two activities can help learners memorize and actively use the collocations they have learned in speaking.

Two implications can be drawn from the study. First, post-editing is an easy and engaging activity. Teachers can generate the translation with the click of a button and demonstrate how AI like Google Translate can come up with a generally high-quality

translation in a second. Additionally, it can be used to check students' comprehension of the reading as they need to understand the original text in order to evaluate the machine translation. The idea of competing with the AI also adds variety to the activities used in the English classroom. However, to prevent students from paying more attention to the translation than to the English text, the cognitive process of machine translation post-editing needs to be carefully controlled. Students should be instructed to focus their attention more on the English text and, when identifying a mistranslation, try to figure out why the AI made such a mistake. In this way, the involvement load of searching and evaluating can better work to enhance students' memory of the English target words. In ordinary classroom instruction, the teacher can also discuss possible translations for the target vocabulary and their collocations so as to reinforce the effects.

Second, sight translation may be an effective way of activating vocabulary. By forcing students to produce what they have learned, they can notice the gap between their comprehension and production abilities, giving them the motivation to activate their passive vocabulary. It is also beneficial for students to work in pairs to receive feedback from each other. Because the translation activity is a back translation into the original English, with the "correct answer" on the handout, the feedback provided by student partners can be as good as native speaker feedback [27], and the process of being corrected by a partner can reinforce learning. However, studies have suggested that sight translation practiced by professional interpreters is a cognitively demanding task as one has to divide attention between visual input and oral production, and the risk of linguistic interference is also higher as the words and sentence structures of the source language remain visible [28]. To decrease the cognitive load of the sight translation process, the task of sight translation should be made easier for language learners. The amount of content to be sight translated, full sentences or just key words, can also be adjusted according to students' levels.

The study is limited in many ways, including the small number of target words selected, the limited number of participants, and the lack of a delayed posttest. However, it has demonstrated that translation tasks traditionally reserved for training professional translators and interpreters may hold great potential as effective learning activities if adopted appropriately in the language classroom and therefore deserve more attention from foreign language teachers and researchers alike.

Compliance with Ethical Standards

Conflict of Interest The authors declare that they have no conflict of interest.

References

1. Kerr, P. (2014). *Translation and own-language activities*. Cambridge: Cambridge University Press.
2. Cook, G. (2010). *Translation in language teaching: an argument for reassessment*. Oxford: Oxford University Press.
3. Pym, A. (2017). Where translation studies lost the plot: relations with language teaching. *Translation and Translanguaging in Multilingual Contexts*, 4(3), 203–222.
4. Carreres, Á., & Noriega-Sánchez, M. (2011). Translation in language teaching: insights from professional translator training. *The Language Learning Journal*, 39(3), 281–297.

5. Carreres, Á. (2014). Translation as a means and as an end: reassessing the divide. *The interpreter and Translator Trainer*, 8(1), 123–135.
6. Takimoto, M., & Hashimoto, H. (2010). An “eye-opening” learning experience: language learning through interpreting and translation. *Electronic Journal of Foreign Language Teaching*, 7(1), 86–95.
7. Wu, Y., & Liao, P. (2019). Interpreting in the language classroom: effects of Chinese-to-English interpreting strategy training on EFL undergraduates’ oral proficiency. *Compilation & Translation Review*, 12(2).
8. Wu, Y., Schuster, M., Chen, Z., Le, Q. V., Norouzi, M., Macherey, W., et al. (2016). Google’s neural machine translation system: bridging the gap between human and machine translation. *arXiv preprint arXiv:1609.08144*, 1–23.
9. Japan Translation Federation. (2017). Google NMT shock. *JTF Journal*, 288, 8–15.
10. Gally, T. (2018). Machine translation and English education in Japan. *Komaba Journal of English Education*, 9, 43–55.
11. 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.
12. Laufer, B., & Waldman, T. (2011). Verb-noun collocations in second language writing: a corpus analysis of learners’ English. *Language Learning*, 61(2), 647–672. <https://doi.org/10.1111/j.1467-9922.2010.00621.x>.
13. Hoang, H., & Boers, F. (2016). Re-telling a story in a second language: how well do adult learners mine an input text for multiword expressions? *Studies in Second Language Learning and Teaching*, 6(3), 513–535.
14. Nesselhauf, N. (2003). The use of collocations by advanced learners of English and some implications for teaching. *Applied Linguistics*, 24(2), 223–242.
15. Nesselhauf, N. (2005). *Collocations in a learner corpus*. Amsterdam: John Benjamins Publishing Company.
16. Xiao, R., & McEnery, T. (2006). Collocation, semantic prosody, and near synonymy: a cross-linguistic perspective. *Applied Linguistics*, 27(1), 103–129.
17. Eyckmans, J., Boers, F., & Lindstromberg, S. (2016). The impact of imposing processing strategies on L2 learners’ deliberate study of lexical phrases. *System*, 56, 127–139. <https://doi.org/10.1016/j.system.2015.12.001>.
18. Laufer, B., & Girsai, N. (2008). Form-focused instruction in second language vocabulary learning: a case for contrastive analysis and translation. *Applied Linguistics*, 29(4), 694–716. <https://doi.org/10.1093/applin/amm018a>.
19. Schmidt, R. W. (1990). The role of consciousness in second language learning. *Applied Linguistics*, 11(2), 129–158. <https://doi.org/10.1093/applin/11.2.129>.
20. Zojer, H. (2009). The methodological potential of translation in second language acquisition: re-evaluating translation as a teaching tool. In A. Witte, T. Harden, & A. R. D. O. Harden (Eds.), *Translation in second language learning and teaching* (pp. 31–51). Bern: Peter Lang.
21. Yamada, M. (2019). Language learners and non-professional translators as users. In M. O’Hagan (Ed.), *The Routledge handbook of translation and technology* (pp. 183–199). New York: Routledge.
22. Laufer, B., & Hulstijn, J. (2001). Incidental vocabulary acquisition in a second language: the construct of task-induced involvement. *Applied Linguistics*, 22(1), 1–26. <https://doi.org/10.1093/applin/22.1.1>.
23. Swain, M. (2005). The output hypothesis: theory and research. In *Handbook of research in second language teaching and learning* (pp. 495–508). New York: Routledge.
24. Chen, W. (2015). Sight translation. In H. Mikkelsen & R. Jourdenais (Eds.), *The Routledge handbook of interpreting* (pp. 144–153). New York: Routledge.
25. Wood, D. (2009). Effects of focused instruction of formulaic sequences on fluent expression in second language narratives: a case study. *Canadian Journal of Applied Linguistics/Revue canadienne de linguistique appliquée*, 12(1), 39–57.
26. Che, P.-C. (Ed.). (2015). *San Min English reader for senior high schools: teacher’s manual* (Vol. 3). Taipei: San Min Book Co., Ltd.
27. Wu, Y. (2019). Chinese-to-English back-interpreting practice in the undergraduate EFL classroom: task design and learners’ perspectives. *Taiwan Journal of TESOL*, 16(1), 67–98.
28. Agrifoglio, M. (2004). Sight translation and interpreting: a comparative analysis of constraints and failures. *Interpreting*, 6(1), 43–67.

Affiliations

Chia-chien Chang¹ · Masaru Yamada²

Masaru Yamada
yamada@apple-eye.com

¹ National Taiwan University, Taipei, Taiwan

² Kansai University, Osaka, Japan