

## 人工翻譯與機器翻譯加後編輯之比較

李家璿 廖柏森

日新月異的科技演進已大幅增進機器翻譯 (machine translation) 的品質, 成為翻譯時的輔助工具。然而我們目前對於人工翻譯與自動機器翻譯系統加後編輯的差異仍缺乏足夠認識, 因此本研究旨在比較這兩種翻譯模式的異同與過程。研究方法以不同英文程度的兩組大學生共 140 人為研究對象, 請他們翻譯英文版的手機保養指南。兩組中各有一部份學生只拿到英文原文, 另一部份學生則多了機器翻譯的中文譯文以進行後編輯。考量機器翻譯使用的便利性及免付費, 故中文譯文取自於 Google Translate。研究中記錄研究對象翻譯過程所花費的時間, 並利用錯誤分析評估其譯文表現。統計結果顯示, 使用機器翻譯的譯文能夠顯著減少某些學生的翻譯錯誤, 亦能縮減學生英文程度落差對於譯文表現的影響。此外, 質性分析更進一步闡明學生如何使用機器翻譯譯文的過程, 以及不同組別之間的譯文差異。這些機器翻譯後端的分析可說明人工應用機器輔助翻譯的實際情況, 也指出學生翻譯上的問題, 希望能作為日後機器翻譯研究與教學的參考。

關鍵詞: 機器翻譯、後編輯、Google Translate

收件: 2011 年 4 月 27 日; 修改: 2011 年 7 月 1 日; 接受: 2011 年 7 月 4 日

# A Comparative Study of Human Translation and Machine Translation with Post-editing

Jason Lee      Posen Liao

With rapid advances in technology, many improvements have been made to make machine translation (MT) an effective tool for translators in the process of translation. However, an important issue that has not been extensively investigated is the use of automatic MT systems with post-editing in comparison to human translation. The purpose of the present study, therefore, is to explore the similarities, differences, and processes of these two modes of translation. In terms of research methods, two groups of college students with different English proficiency levels were asked to translate a cell phone user guide written in English. In each group, some students received only the English source text while others received the source text in addition to a machine-translated Chinese text for post-editing. Google Translate was the MT system used in this study because of its easy and free accessibility. A total of 140 subjects were timed on their translation tasks, and their performance was determined by the number of errors in the translated text. The statistical results indicated that the MT text was very helpful in reducing errors in some student translations; the use of MT also shortened the gap between students of divergent language proficiency levels. Further qualitative analysis elucidated how the MT text was utilized and the discrepancies in lexical choice and other aspects between the two groups of students. This back-end analysis of the machine translation process may offer insights into the practical use of human-assisted machine translation, as well as problems encountered by students when translating. It is hoped that these results may serve as a basis to facilitate future machine translation studies and teaching.

Keywords: machine translation, post-editing, Google Translate

Received: April 27, 2011; Revised: July 1, 2011; Accepted: July 4, 2011

## Introduction

Machine translation (MT), according to Hutchins (2001), refers to computerized systems responsible for the production of translations with or without human assistance. The mid 1930s heralded the beginning of this revolutionary idea, but it was not until the 1970s that various research projects paved way to working, commercially available systems (Arnold et al., 1994). Then, in the 1990s the Internet boom in the 1990s ushered MT into a new era, as the enormous amount of available multilingual documents enabled corpus-based MT techniques, consequently motivating ever more research effort into online MT systems. Today, numerous translation software applications are available on the market that can help those who are lacking language skills to translate a source text into a target text with a simple mouse click. The results of these systems should mean that MT has been established as a useful application of technology in translation, but the fact is that MT remains elusive to many, even veteran translators.

Unacquaintance aside, the development of MT systems has propelled and inspired numerous studies, ranging from the evaluation of MT performance (Chan & Ng, 2008; Liu et al., 2005), to improved add-ons from the programming perspective (Chang et al., 2008; Mauser et al., 2008). The continuous publication of studies on MT means it is a topic worth exploring. However, only a few researchers so far have looked into and compared the translation time and translation accuracy between human translation (HT) and machine translation with post-editing (MT PE). For example, Allen (2004) provided proof that using MT can help a translator produce approximately three times more in a day than the average translator without MT. Guerra (2003) demonstrated that MT with back-end human intervention, i.e. full post-editing, is faster than human translation from scratch by as much as 70%. And Shih (2007) explored students' translation process through a comparative study, evidencing that with the help of MT, students produced fewer translation errors, took less time to complete a translation task while using smaller translation units. These studies demonstrated that translators can perform a

translation task faster and better with the use of MT.

Although the aforementioned studies comparing MT PE to HT were empirical investigations rather than theoretical postulations, the sample size of the latter two experiments might not be considered adequate. There were only two translators involved in Guerra's (2003) investigation, one being herself who post-edited MT texts, the other who translated the texts from scratch without MT. On the other hand, the sample size in Shih's study was bigger with 20 students, but only descriptive statistics were used to describe the results without further calculations of statistical significance. Moreover, the language pairs in these previous studies were French, Spanish or Chinese to English, respectively, with Shih's study exploring Chinese to English direction. There has yet to be an empirical investigation on MT efficacy concerning English to Chinese translation with the reference of MT text, so the present study will endeavor to offer insights into this unexplored territory.

Assessing college students' translation accuracy and timing their translation are the main purpose of this study, in addition to cross-comparing the results among students with MT and without MT at different language proficiency levels. The quantitative results are subjected to various statistical analyses to see if there is any disparity between different student groups. Other related issues, as will be discussed qualitatively in detail, include: to understand more about the pros and cons of using MT text in carrying out a translation task as well as to probe into some of the translation problems encountered by students participating in this study.

Gleaning into the practical use of a readily available MT system can help us understand its value in aiding the translation process, providing constructive and helpful suggestions on using MT in the classroom or for real translation assignments. The differences in the errors made by students may also elucidate possible ways to use MT more effectively, which may have benefits that go beyond improving the translation speed and quality and serve as a potential approach to leveling the playing field among students of different language proficiency levels in the same class. The present empirical investigation is one of the few studies to look at the use of MT systems for English-Chinese translation, with the hopes of drawing more attention to this fascinating but under-appreciated technology.

## Review of Related Literature

### *Machine Translation*

Machine translation exploits advanced computational linguistic analytic processing to automatically translate, sometimes with human intervention, source documents into target texts. In general, human and machine translation can be categorized into four types (Hutchins & Somers, 1992):

1. Human Translation (HT)
2. Machine-Assisted Human Translation (MAHT)
3. Human-Assisted Machine Translation (HAMT)
4. Fully Automatic Machine Translation (FAMT)

In this paper, HT and HAMT will be compared, while post-editing (PE) is involved in the back-end of the translation process. MAHT and FAMT will only be addressed briefly.

MT has seen evolutions in its translation approaches, starting from the direct approach, to the current rule-based, example-based, statistical systems, or hybrids that have been gaining popularity. In the beginning of the 1950s, the direct approach employed a lexicon-oriented method, where the SL (Source Language) input first underwent simple morphological analyses so that word endings were properly identified and the inflected forms reduced to their base forms. This was followed by dictionary look-up to supply TL (Target Language) word equivalents. Some local reordering rules may be applied to rearrange the target word order into a more acceptable pattern (Wong, 2008).

Then around the 1970s and 1980s, transfer-based or rule-based approach attempted to carry out deeper linguistic analysis using multiple levels of linguistic knowledge, describing SL linguistic structures and how they mapped onto the TL (Hutchins & Somers, 1992). This approach consists of three steps—(1) analysis: to transform the surface form of a SL sentence into an abstract representation relating to the linguistic characteristics of the SL; (2) transfer: to map this representation for the SL to that for the TL; and (3) generation: to transform the abstract representation for a TL sentence to its surface form (Wong, 2008). Different levels of linguistic analysis are applied

to produce the abstract representation. A syntactic analysis is common, e.g. examining the linguistic representations of the correspondent grammatical features between the SL and TL for each language pair. When the analysis and generation are executed at the semantic level, a representation model for the meaning in a language is needed, and the ideal model for this is interlingua, an abstract language-independent representation. Hence, this MT method is also called the interlingua approach. SYSTRAN is the most popular MT systems employing such an approach. Yet, this approach cannot settle most of the ambiguities that may arise owing to the lack of linguistic information about the source text. For this reason, this approach necessitates language experts and linguists to meticulously craft extensive lexicons and rules related to grammar, syntax, and semantics to generate text in a target language (Stix, 2006).

The most state-of-the-art MT system is corpus-based MT that makes use of existing translated bi-texts. There are two variants to this approach, namely, example-based MT (EBMT) and statistical MT (SMT). In EBMT, a set of existing translations for the appropriate fragments of the SL sentence are retrieved, followed by the amalgamation of the retrieved translations into a grammatical TL sentence (Hutchins, 2005). On the other end, SMT treats translation as a mathematical problem. When fed a large amount of translation data, a statistical model is used to crunch the relevant linguistic dependencies between the source and target texts, and yield specific probability values from the translation data as the parameters of the model. The actual translation process of SMT involves the selection of a target word combination of the highest probability as the translation for a source sentence, based on the probability distribution in the trained statistical model (Stix, 2006). One of the advantages of SMT is that it is able to see the translation as word segments or phrases, since it also uses a monolingual corpus; consequently, the machine produced translation takes into account of how words are collocated to generate more natural translations. Google Translate, the MT system used in the present study, also uses its own statistical MT system.

## *How People Perceive MT*

MT remains largely a controversial issue, and every so often, a source of ridicule or joke. For example, Hutchins (2003) pointed out that machine translation is still better known for its failures than for its successes. The report produced by ALPAC (Automatic Language Processing Advisory Committee) in 1966 claimed that “there is no immediate or predictable prospect of useful machine translation” (Hutchins & Somers, 1992, p. 32). Later on in 1975, Chomsky wrote, “as for machine translation and related enterprises, they seemed to me pointless as well as probably quite hopeless” (Chomsky, 1975, p. 40).

Despite negative complaints, MT still receives some well-deserved praises. According to Newton (1992), even though MT cannot be indiscriminately used, it can still be quite helpful because machines are consistent in rendering terminology, do not skip paragraphs or sentences and do not make wrong assumptions that experienced human translators sometimes might make. Gross (1992) also noted that machines are faster, more cost effective, and in some cases more accurate than their human counterparts. He argues that, computers are competitive if given a limited and particular subject domain, similar to human translators specializing in a particular translating domain to be highly efficient.

## *Post-editing of MT Output*

Current MT systems have technical, linguistic and extra-linguistic limitations, thus their outputs require editing to facilitate the emendations of the artificial or imperfect MT outputs into publishable texts (Shih, 2006). Post-editing (PE) is the term used for editing, modifying and/or correcting machine-translated texts. Senez (1998) refers to it as a term used for the correction of machine translation output by human linguists/editors, and Wagner (1985, cited in Allen, 2001, p. 26) says it is “the correction of a pre-translated text rather than translation from scratch”.

Training in post-editing first of all involves the detection of translation errors that arise from the limitations of the MT system. Sometimes post-editors have to cross-reference with the source text,

consult reference materials or resort to other tools to find solutions to the MT-produced errors (Shih, 2007). In this regard, MT PE is similar to human translation. However, having a gist translation for reference can indirectly affect the translation process and lead to different translation performance.

In general, MT PE is a type of translation service that is offered as an optional parallel process to HT. The term post-editing in this study is limited only to the task of editing, modifying and/or correcting raw MT output. MT does not equal to HT, but it adds an assortment of options that should not be underestimated. A translator is, more often than not, accustomed to a certain freedom in word choices when translating, but a post-editor is restricted by the words provided by the MT system, many of which could be wrong and misleading. Therefore, HT can be seen as a creation as opposed to PE as a correction mechanism, as proposed by Guerra (2003), who adds that PE is an ongoing exercise of revising relatively predictable inaccuracies, much like a discovery process.

In a survey on the MT experience of UK-based freelance translators, Fulford (2002) reported that over half (53%) of the translators in the sample had undertaken post-editing work of documents translated using MT systems for clients, and 27% of those translators undertook such editing work on a frequent and regular basis. This post-editing service where efficiency is a priority could be successfully fulfilled by those with good bilingual and linguistic skills, e.g. marketing and product specialists (Allen, 2001), students (Somers, 2003) or trained specialists (Allen, 2003). Moreover, Fulford (2002) suggests that, although the uptake of MT has, up to now, been slightly low in the freelance translation community, many translators express a keen interest to learn more about MT and to explore the prospect of using it in their translation assignments.

### *Error Analysis for Translation Accuracy Assessment*

Since the present study explores MT with post-editing, which entails human involvement in the translation process, approaches to evaluating human translation should also be weighed, e.g. translation competence. This study uses Pym's (1992) classification of binary and non-binary errors for translation accuracy assessment. As de-

defined by Pym, binary error means “It’s wrong.”, while non-binary error stands for “It’s correct, but...”. A binary error opposes a wrong answer to the right answer, i.e. there is only right and wrong. Meanwhile, a non-binary error suggests that there is at least one more possible TT (Target Text) which could also have been selected other than the TT actually selected, i.e. there are at least two right answers and then the wrong ones. The former is attributed to the lack of language proficiency, whereas the latter is the result of poor translation ability. Pym’s error classification, i.e. binary and non-binary errors, can be more or less seen as errors in fidelity and readability, respectively, as the former implies a deviation from the true meaning of the source text, while the latter entails the target text is correctly rendered in terms of meaning but with blemishes that could be otherwise substituted by other more polished alternatives.

With only two metrics, the translation evaluation process is easier. The results can be concrete and offer insights into the problems encountered in performing a translation task. Table 1 shows the definition of the two types of translation error used in this study.

Table 1. Error Classification (translated from 賴慈芸, 2003)

	Type	Definition
Binary error	Misunderstanding of ST (English)	Misinterpretation of words, meanings
	Faulty rendition of TT (Chinese)	Syntactic errors, omissions
Non-binary error	Improper word use	Improper collocation, register mismatch, vague expression, inability to convey the message
	Insufficient transfer competence	Over-long modifier, logic inconsistency, over-use of pronouns, superfluous words

Pym (1992) also pointed out that the ratio of binary and non-binary errors can be used to tell apart beginner and advanced student translators; that is, binary errors tend to decrease in number, accompanying by more non-binary errors, with improving translation competence. In the end, students should be able to translate a text

without making any binary errors. Measuring students' ratio of binary and non-binary errors is also one of the aims of this study.

### *MT in the Classroom*

The use of MT technology in the classroom is on the rise. According to Bowker (2002), this increase is needs-driven, an outcome of globalization resulting in the ever-expanding volume of translation along with increased pressure to produce translation quickly. Sufficient knowledge with MT technology has become a prerequisite for translation students if they are to be able to front the challenges and survive in the competitive market.

For example, Somers (2003) finds MT very useful for advanced students, where the students are asked to use software to produce a first draft into their native language and then produce an improved post-edited version, along with a commentary. These students have had classes on the commonly seen difficulties and problems of MT, and are asked to relate errors in the MT text to problems previously discussed in class. This way, they have hands-on practice on classifying errors made by the MT system on a linguistic and pragmatic basis. As Shih (2007) suggests in her study where students learn about translation of different text types through post-editing of MT text, students respond that they have learned the different contextual elements of the different text types, including organizational patterns, syntactic modes and lexical categories through the practice, confirming the value of teaching translation with MT error analysis and post-editing.

In a study by Kliffer (2005) aiming to introduce MT via post-editing to a class of third-year French majors, it is found that the weaker students appreciated the exercise more than the stronger ones, some of whom found MT to contain too many stupid errors. Interestingly, the weaker students expressed that having an MT translated text and correcting its mistakes is less stressful than doing a translation entirely by themselves without any reference.

Richmond's (1994) use of MT that allows students to pay more attention to the contrasts between two languages, English and French in this case, has been proven to be an enjoyable exercise for some students, in which they find working with the MT software challeng-

ing and worthwhile. In Richmond's study, students are asked to type in the source text in English and note that the MT sometimes gets the translation in French wrong. They are then asked to modify the English sentence and retranslate it using the MT system, continuing to do so until they obtain a more appropriate target text. The distinctive aspect of this lies in the fact the MT system tends to produce relatively literal translations, and for a more desirable output, the original English text has to be revised and twiddled to make it more like the French target text, in which he calls this "doing it backwards" or, in professional terms, pre-editing. Richmond says the pedagogical implication is that the students learn the differences between French and English, and they are acquainted with how to express meanings in French through the process of modifying English.

Furthermore, Yuste (2001) indicates the more students learn about the functions of MT systems, the more they are eager to use MT in the future, particularly speediness as one of the most attractive aspects. When the students were asked to compile a list of potential users and situations where MT could be particularly of use, some came up with the idea that freelance translators could benefit from MT to produce more translations in less time, and, consequently, make more money. Gaspari (2001) also suggests that the more the trainee translators become familiar with MT, realizing its potential and limitations, the less afraid they are to use it.

The aforementioned studies support the necessary integration and introduction of MT into the translation curriculum, because computer technology may directly or indirectly affect professional human translators.

## Research Method

This study investigated the translation processes of human translation and MT with post-editing. The time spent on completing the translation and the accuracy of the translated texts were calculated in order to probe into the discrepancies between with or without MT use in performing a translation task. The percentage of words taken from the machine translated text and used in students' translation was also computed. In addition, qualitative examinations would further shed light on possible reasons for the similarities and differences

of the two translation modes.

## Research Subjects

The subjects of this study were 140 translation students recruited from three colleges in Taipei. Of them, 58 students came from a top national university, 14 were first-year graduate students in a graduate institute of translation and interpretation, and 68 students were from an institute of technology. Preliminary results showed that there was minute performance difference between the students from the prestigious university and those from the graduate institute. Hence, in this research, students from the national university and the graduate institute were classified as Group I, and those from the institute of technology as Group II.

The differentiation of the subjects' English proficiency was determined by the percentage of students passing the High-Intermediate level of the General English Proficiency Test (GEPT) in Taiwan. The statistics showed that 82% of students in Group I passed this level, while only 2% of Group II students reached the equivalent English proficiency.

## *Data Collection Procedures*

Prior to the experiment, a pilot was conducted with two students at the graduate institute of translation and interpretation. They were invited to translate the text used in this study, one receiving only the source text, while the other had the MT text for reference. Both students took about 30 minutes to complete the translation task without the use of computers. When asked about the difficulty of the text, both responded that the text was of medium difficulty, in that they did not need a dictionary to look up words, but efforts were still needed to digest the source text and produce a publishable translation. The student with MT text for reference commented that the MT text was helpful in terms of providing word choices, but was somewhat slowing down the translation process from thinking of better ways to revise the MT text into a refined target text. Neither of them noted they had trouble comprehending the source text.

In the formal experiment, prior to handing out the source text, one of the researchers stated the purpose of this study to the participating students in three different schools, explaining that translation time and accuracy were the two factors to be measured. Using a PowerPoint presentation, the researcher also gave a few examples of post-editing techniques, such as reordering of words or phrases (See Appendix 4). Then, participating students in both groups either received a clean hard-copy with only the source text (No MT) or one with source text plus machine-translated text (With MT) (see Appendix 1 and 2). They were asked to hand-write the translation in a blank paper provided to them. No dictionary or Internet access was permitted. However, a vocabulary list with technical terms or words was made available to the students (see Appendix 3). Also, their translation effort was timed. A digital stopwatch for timing purpose was projected on-screen in the classroom. Students began their translation once the stopwatch was clicked, and they were to record the time took to finish the translation on top of the paper upon completing the translation. Each translated copy was then numbered for record keeping and further analysis.

### *Translation Material and Text Readability*

It is widely accepted that MT systems are more appropriate for domain-specific text in controlled environment, e.g. technical manuals and user guides (Hutchins, 1995). With vast amounts of text of instructional and descriptive nature that are utilitarian and intended to serve a relatively finite purpose, a cell phone care instruction guide was taken from *Caring for your mobile phone webpage*<sup>1</sup> of Nokia Europe and used for the experiment in this study. Minor revisions were made, such as taking out the brand name, but with most of the text remaining intact. The final version contained a total of 208 words (see Appendix 1). The machine translated text was obtained by feeding the source text to Google Translate<sup>2</sup>, setting the direction to be English to Traditional Chinese, and used directly in the experiment without any modification (see Appendix 2).

Moreover, as Hale and Campbell (2002) propose, there exists an interaction between text difficulty and translation accuracy. One of

the approaches to determining text difficulty is to measure its readability. In order to make sure the translation material was not too simple or too difficult for the subjects to comprehend, the text was subjected to a set of readability tests. The text readability scores were acquired by using the measurement tool at <http://www.addedbytes.com/code/readability-score/>. This tool reported that the text used in this study contained 13 sentences, with 208 words (16.00 per sentence) made up of 305 syllables (1.47 per word), with a readability index of 9.84, implying 10 years of formal education would be needed to comprehend this text. This means a 10th grader in the U.S. would have little problem understanding the passage.

### *Translation Tool—Google Translate*

The MT tool used in this study is Google Translate (see Figure 1), which is powered by its own statistical translation system. It is easily accessible online all the time free-of-charge, and it is able to translate a lengthy piece in the matter of seconds. On top of that, according to an MT test held by the National Institute of Science and Technology (NIST) in the United States, Google Translate was rated number one in the English-Chinese language pair in 2008, besting Systran (IBM), Microsoft, and other competitors (NIST, 2008).

Google Translate is an automatic translator, using state-of-the-art technology. Instead of employing a rule-based approach in the previous MT systems that require tedious analyses of vocabularies and grammars, Google Translate feeds the computer billions of words of text, both monolingual text in the target language and aligned text consisting of examples of human translations between the languages. Statistical learning techniques are then applied to build a translation model. Google Translate currently supports translation between 57 languages and has a large corpus of billions words' worth of human translations (Tanner, 2007).



Figure 1. The Interface of Google Translate

Nonetheless, Google Translate, like any other machine translation systems, has its limitations. It can help readers understand the gist of a foreign language text, but it does not always produce accurate translations. This is part of the reason why post-editing is needed to alter and refine the machine-translated text if it is to be published.

### *Quantitative and Qualitative Data Analyses*

The quantitative aspect of data analysis consisted of calculating the time spent on the translation and errors that would impair the accuracy of the translation. The time calculation was taken directly from the time recorded by the students. The error calculation was carried out by assessing each student's 13 translated sentences and determining whether the student had made binary, non-binary, or no errors in 11 sentence units, some of which might contain two smaller sentences. If a sentence unit had both binary and non-binary errors, it was tagged with binary errors. The number of each type of error

in each sentence unit would not affect the error tagging, i.e. if a sentence had two binary errors and one non-binary error, it would still be tagged with binary error, as exemplified in Appendix 4. For each student, the ratio of binary to non-binary errors was calculated and abbreviated as B/N ratio in later sections.

For the benefit of statistical significance analysis, all students' errors were converted into a numerical score. There were a total of 11 sentence units in the translated text, and two points were assigned to each of these sentence units. For the score conversion, two points would be deducted for each binary error count in each sentence unit, and one point for each non-binary error count. Hence, the highest possible score a translated piece could receive was 22, in which no errors were found in any of the sentences, and the lowest score would be zero. The translation time and score data were both subjected to independent *t*-test using SPSS 17.0 to see if any significant differences existed between the two student Groups and the No MT and With MT sets.

Since no computer was allowed in the experiment, all the target texts were only accessible in hard-copy. To explore the percentage of MT words used in students' translation, the number of borrowed words in each With MT translation was hand counted by the authors. A converted score was calculated by counting the number of binary and non-binary errors made in each sentence by individual students to be summed up for deduction, where binary errors were subjected to a deduction of 2 points and 1 point for non-binary ones. The sentence with the most deducted points was assigned to have a score of zero, while the remaining sentences received scores based on their deducted score differences with the score-zero sentence. The correlation between the percentage of MT word use and translation score was analyzed using SPSS 17.0. Lastly, the survey results based on subjects' responses in percentage were also computed.

The qualitative analysis mainly involved the examination of different translation products and processes between With MT and No MT students in both Groups I and II. Categories of investigation included comparing students' register and word use, how differences in language proficiency lead to differences in performance, how the students utilize the MT text, what translation problems students seem to have, as well as what constraints the MT system may face.

## Results and Discussion

### *Quantitative Findings*

The following will report the statistical results of this study, which are divided into sections of time spent on translation, translation accuracy, ratio of binary to non-binary errors, percentages of MT text used in translation, the relationship between the percentage of MT use and translation accuracy, as well as results of the feedback survey.

### *Time Spent on Translation*

After the experiment, a total of 140 student translations were obtained. Out of 72 translations collected from Group I, 69 were effective for analyzing time spent on translation (3 students in the “With MT” group did not record the time). Of the 68 translations gathered from Group II, 66 were effective for the time analysis (1 student in the “No MT” and 1 student in the “With MT” group did not put down the time). In Group I, the No MT set averaged 32.18 (SD=10.06) minutes in completing the translation, while the With MT set averaged 33.23 (SD=7.82) minutes. An independent *t*-test was conducted and the result showed that there was no significant difference between the two sets in terms of translation time.

The No MT set from Group II averaged 26.80 (SD=6.06) minutes in completing the translation, while the With MT set averaged 28.91 (SD=6.16) minutes. An independent *t*-test of these two means indicated no significant difference between the two sets in terms of translation time. These results are shown in Table 2.

Table 2. The Mean Time Spent on Translation in No MT and With MT Sets of Each Group

Groups	No MT: Mean time spent (min)	With MT: Mean time spent (min)	P
Group I	32.18 (n=36)	33.23 (n=33)	0.63
Group II	26.80 (n=33)	28.91 (n=33)	0.17

Students in both No MT sets finished a little faster than their counterparts who had MT text for reference. The data collection procedure did not allow the test subjects to use a computer, due to mainly resources limitations and the fact that these students were from different schools and it would add even more variables to conduct the experiment in different computer labs. For instance, Internet access, processor speed, different versions of word processing software, etc., may contribute to variations in the outcomes that are not part of the variables this study aims to explore. As a consequence, the measured time spent on translation may not completely reflect the use of MT PE (post-editing) in a real setting. If we borrow the results from previous studies (e.g. Allen, 2003; Guerra, 2003), it can be indirectly inferred that the With MT groups would be faster if they were to perform the translation task on a computer.

For instance, Allen (2003) demonstrated that the MT PE cycle of a magazine article took one third of the time than a human translator to complete the translation. Guerra (2003) also showed that MT PE was 40-60% faster than a human translator. These findings suggest that the With MT groups in this study would have probably only used approximately half the time to complete the translation, provided they had done the task with a computer. This can be further corroborated by calculating the percentage of words borrowed from the MT text in the With MT sets. The resulting numbers illustrated that Group I used 58.5% and Group II used 66.3% of the MT text on average, which means around 60% of the words in their translation can be directly used without typing them down that would have accounted for a substantial portion of time needed to complete the translation task.

### *Translation Accuracy Assessment*

All 140 translations were examined for their translation accuracy. As shown in Table 3, in Group I, No MT set had on average 3.53 binary and 4.80 non-binary errors in their translation task, or 10.14 converted points. With MT set had on average 2.33 binary and 5.33 non-binary errors, or 11.94 converted points. The B/N ratio was 1.06 (SD=1.01) for No MT and 0.55 (SD=0.51) for With MT.

Also as shown in Table 3, in Group II, No MT had on average 7.59 binary and 2.47 non-binary errors in their translation task, or 4.35 converted points. With MT had on average 4.21 binary and 5.32 non-binary errors, or 8.26 converted points. The B/N ratio was 4.65 (SD=3.38) for No MT, and 1.19 (SD=1.56) for With MT.

Table 3. The Mean Errors Made by No MT and With MT Sets in Each Group

Errors	No MT (I)	With MT (I)	No MT (II)	With MT (II)
Binary errors	3.53	2.33	7.59	4.21
Non-binary errors	4.80	5.33	2.47	5.32
Converted score	10.14	11.94	4.35	8.26
B/N ratio	1.06	0.55	4.65	1.19

Next, the accuracy of the translations was compared between different student groups. Both in Groups I and II, With MT set performed better than their counterparts without MT as a reference, where the number of binary errors was approximately halved in the With MT sets. The results of independent *t*-tests in the following Tables 4 and 5 also show that, in terms of converted scores, there was a significant difference between the No MT and With MT sets in both Groups I and II ( $p < 0.05$ ), especially in Group II where the With MT set scored almost twice than the No MT set did. This indicates that MT is indeed helpful in improving translation accuracy. It is also interesting to note that by applying the scoring method to the target text produced by Google Translate, it would receive a score of

5, which is slightly higher than the No MT set in Group II, with 6 binary errors and 5 non-binary ones.

Table 4. The Results of *t*-test between Group I and Group II

<i>No/ With MT</i>	<i>Group I: Converted score</i>	<i>Group II: Converted score</i>	<i>P</i>	<i>t</i>	<i>Std. Error Difference</i>
No MT	10.14 (n=36)	4.35 (n=34)	.000*	9.51	0.61
With MT	11.94 (n=36)	8.26 (n=34)	.000*	5.48	0.67

Table 5. The Results Sets of *t*-test between the No MT and With MT sets

<i>Groups</i>	<i>No MT: Converted scores</i>	<i>With MT: Converted scores</i>	<i>P</i>	<i>t</i>	<i>Std. Error Difference</i>
Group I	10.14 (n=36)	11.94 (n=36)	.008*	2.72	0.66
Group II	4.35 (n=34)	8.26 (n=34)	.000*	6.39	0.61

Students' translation accuracy improvement with the use of MT can be further validated through looking at the performance between Groups I and II. Both the No MT and With MT sets between Groups I and II exhibited significant difference. However, the mean difference of 3.68 between the With MT sets in Groups I and II was much smaller than that of 5.79 between the No MT sets (see Table 4), implying that MT is bridging, albeit not entirely, the gap between students whose language proficiency may not be on the same level.

### *Ratio of Binary to Non-binary Errors (B/N ratio)*

An independent *t*-test was once again conducted to see if there was a difference in the ratio of binary to non-binary translation errors (B/N ratio) between the No MT and With MT sets in both Groups I and II. The statistical results are presented in Table 6 below.

Table 6. The Results of *t*-test for B/N ratio between Groups I and II

No/ With MT	Group I: B/N ratio	Group II: B/N ratio	P	t	Std. Error Difference
No MT	1.06 (n=36)	4.64 (n=34)	.000*	-6.26	0.59
With MT	0.55 (n=36)	1.19 (n=34)	.022	-2.35	0.27

Table 6 shows that there was a significant difference in B/N ratio between the No MT sets in Groups I and II, but no significant difference between the With MT sets in the two Groups. Given Pym's (1992) suggestion that the ratio of binary to non-binary errors decreases as students advance in their translation ability, Group I performed significantly better than Groups II in both No MT and With MT sets, indicating Group I is likely to be more ahead in their translation training than Group II. However, it is interesting to observe that the performance of With MT set in Group II was more on par with that of Group I; that is, no significant difference was observed, whereas a significant gap was observed between the two Groups in the No MT set (see Table 6). On another note, the B/N ratio of the With MT set in Group 2 was very close to that of the No MT set in Group 1, suggesting that with machine-translated text at hand can substantially assist students lacking sufficient language proficiency to perform a translation task more successfully. Meanwhile, the B/N ratio for the target text produced by Google Translate was 1.2.

Another *t*-test for B/N ratio between the No MT and With MT sets was performed. Statistical results in the following Table 7 show that the With MT sets in both groups outperformed their No MT counterparts, highlighting once again the assistance of MT in carrying out a translation task, whilst confirming its value in improving the accuracy of students' translation.

Table 7. The Results of t-test for B/N ratio between the No MT and With MT Sets

Groups	No MT: B/N ratio	With MT: B/N ratio	P	t	Std. Error Difference
Group I	1.06 (n=36)	.55 (n=36)	.008*	-2.71	0.19
Group II	4.64 (n=34)	1.19 (n=34)	.000*	-5.57	0.64

### *Percentages of MT Text Used in Translation*

This section presents the results of the percentage of MT text used in student's translation. The number of words used was calculated as percentages for the With MT sets, as shown in Fig. 2. From the bar graphs, it can be observed that, on average, Group II students (66.3%) relied more heavily on the MT for their translation than Group I (58.5%). Among the total of 11 sentence units, sentence 8 was the most borrowed, which was correlated to fewer With MT students making binary errors when translating this sentence. Meanwhile, sentence 3 was the least borrowed, which had comparatively more students making binary errors in their translation.

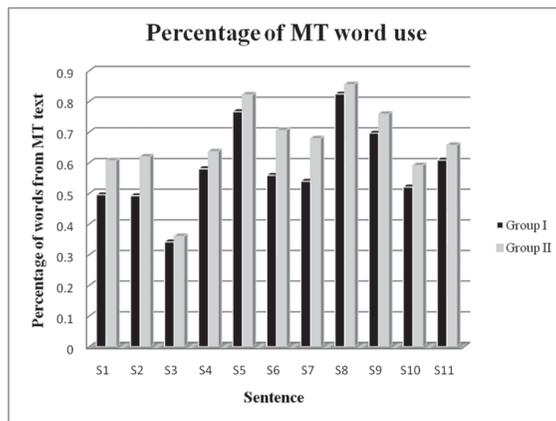
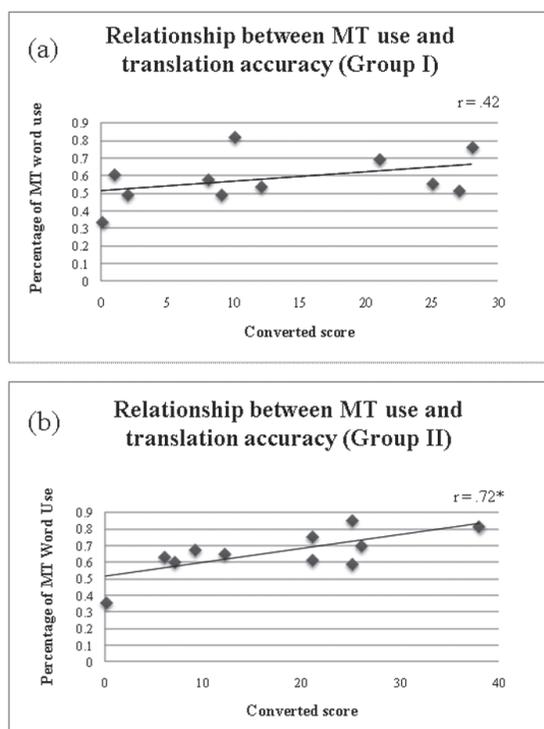


Figure 2. Percentages of MT Text Used in Students' Translation

### *The Relationship between the Percentage of MT Use and Translation Accuracy*

Pearson correlation analysis was conducted to show the relationship between the percentage of MT words used in the translation and the converted scores of the two Groups, and the results are presented in Fig 3. No significant correlation was observed in Group I (n=36, r=0.42), but a highly significant positive correlation was seen in Group II (n=34, r=0.72).



\* denotes statistical significance

Figure 3. Correlation between MT Word Use and Accuracy: (a) Group I and (b) Group II

Drawing from the word count analysis, there was a highly significant positive correlation ( $r = 0.72$ ) between the number of words borrowed from the MT text and the students' translation accuracy

in Group II. The finding demonstrates that the more words from the MT text a student uses, using sentence as a unit, the less likely a student would make a mistake in translating that particular sentence. This is possibly due to the fact that the students recognize they can use the MT directly if the meaning is intact, and they would only have to do a little tweaking. This finding echoes Guzmán's (2007) suggestion that "if it's not broken, don't fix it," which, when put in more professional terms, means to avoid unnecessary stylistic corrections during manual MT post-editing. This finding also suggests that the percentage of words borrowed from the MT can be seen as an indicator to locate problematic or difficult sentences in the passage, thus further attention can be paid toward them, much like checkpoints in translation assessment. In short, if the sentence has a lower percentage of borrowed words from the MT text, it will very likely be a problematic sentence for students.

### *Feedback Survey*

A short feedback survey with eight questions was administered to the subjects to understand their responses to the translation task. Both With MT and No MT students were required to answer the first three questions (questions 1 to 3), but only With MT students needed to answer the remaining five questions (questions 4 to 8). Table 8 depicts the survey results with responses of subjects from Groups I and II in percentages.

Table 8. Survey Results

Questions	Responses	Group I	Group II
1. What do you think of the difficulty of this text?	Extremely easy	1%	0%
	Quite easy	56%	13%
	Medium difficulty	36%	52%
	Rather difficult	6%	31%
	Extremely difficult	1%	4%
2. What do you think of your translation speed?	Extremely slow	12%	16%
	Quite slow	43%	28%
	Moderate	40%	52%
	Pretty fast	4%	4%
	Extremely fast	0%	0%

3. What do you think of the accuracy of your translation?	I guessed most of it	1%	10%
	I guessed some of it	20%	59%
	Mostly correct	42%	27%
	Almost no mistakes	33%	4%
	No mistakes	0%	0%
4. Do you think the MT text is helpful during the translation process?	Not at all	3%	0%
	A little bit	16%	3%
	Of some use	53%	66%
	Quite useful	25%	31%
	I copied all	3%	0%
5. Do you think the MT text is helpful in improving your translation speed?	It slowed me down	14%	11%
	Same without it	3%	3%
	A little bit	44%	57%
	I did it faster	31%	26%
	I did amazingly faster	3%	3%
6. Do you think the MT text is helpful in improving the accuracy of your translation?	Sometimes misleading	8%	3%
	Same without it	23%	11%
	A little bit	60%	63%
	Marked improvement	9%	23%
	Must have MT text to translate	0%	0%
7. If the MT text is useful, in what aspect is it the most helpful?	Vocabulary	24%	36%
	Terminology	41%	24%
	Comprehension	11%	26%
	Save time on thinking of words to use	14%	14%
8. Would you consider using MT for your future translation assignments?	Never	0%	0%
	Not so sure	8%	0%
	Only if I have the time	25%	29%
	A good reference	64%	63%
	I love MT	3%	8%

From the results, a discrepancy toward scoring the text difficulty between Group I and Group II was observed. In the subjects' responses to question 1, more than half of the students in Group I (57%) thought the text was quite simple, while 36% deemed it to be of medium difficulty. On the other hand, over half the students in Group II (52%) thought the text was of medium difficulty, and 35% believed it was rather or extremely difficult. This disparity was also reflected in their translation performance, as Group I outperformed Group II by a significant margin.

Another finding from this survey demonstrates that in Group II, of the 21 students who thought the text was rather difficult, the With MT set only accounted for 7 (33%), and of the 35 students who deemed the text to be of medium difficulty, 24 (69%) had MT as a

reference. The presence of MT at hand had boosted their confidence and helped them translate. Nonetheless, this phenomenon was not observed in Group I, as the numbers were more or less divided equally between the With MT and No MT sets.

Question 3 indicates that most students in Group I (75%) thought their translation was mostly correct or thought they had made almost no mistakes. In contrast, most of the students in Group II (69%) put down they guessed some or most of the translation, all of whom did not have MT for reference. A lot more students With MT in Group II thought their translation was mostly correct compared to their counterparts in the same group. These findings suggest that MT is indeed useful in helping students translate, reducing the frequency of guesses that would hamper the accuracy.

Question 5 shows that a small portion of With MT students, 11% in Group I and 14% in Group II, suggested the MT actually slowed down their translation, and some of them had commented that “it takes up more time having to revise the MT text” and “the MT text is sometimes misleading”. From question 6, 31% of subjects in Group I and 14% in Group II thought the MT text was misleading or they would have performed the same without it. But 23% in Group II said their translation markedly improved with the use of MT.

Though a portion of the students did not care for the use of MT while carrying out a translation assignment, almost all the students, regardless of Groups, said that MT was helpful during the process (question 4), while more than 60% of the students in both groups put down they would consider using MT for their translation tasks in the future (question 8). There was a slight difference in the aspect of which the students think MT was the most helpful between the Groups (question 7). Most students in Group I (41%) thought the MT helped them in translating the terminologies, but only 11% considered the MT was helpful in enhancing their comprehension. On the other end, most students in Group II (36%) found the MT text useful in that it offers the meaning to vocabulary words they might not know, and 26% thought MT text would help them comprehend the passage. From this, we can suggest that MT is useful for students, particularly those who are beginners in translation or language proficiency, to grasp the message of the text.

## *Qualitative Findings*

The following sections will provide more in-depth analysis to the translated texts created by both With MT and No Mt students in Groups I and II, on top of pointing out the helpful aspects of MT to students.

### *Word Use Consistency and Proofreading*

Maintaining a consistent translation of terminology or nouns is imperative, especially in technical translation, because it is critical to the readability of localized materials (Itagaki, Aikawa, & He, 2007). This is why translation agencies use a term base to ensure all the different translators translating the same text can use the same translation for the same term, preventing ambiguities and confusion for readers.

The term “mobile device” appeared six times in the source text, but inconsistent translations were observed, especially in the With MT sets. Since Google Translate rendered the term as 移動設備, many students only changed the 移動 into 行動, which is more commonly used in Taiwan. However, toward the end of the passage, some students changed the translation again to 行動裝置, which is actually the most suitable translation for the term. This phenomenon was seen in both the With MT sets in Groups I and II, where they might be constrained by the MT text in the beginning but later found that a better translation can be used. Another possible reason for this inconsistency is the With MT students did not read the entire text beforehand; hence they directly used the MT text without pondering for a better translation. The No MT sets, in contrast, might have read the passage first then carried on with the translation; therefore, they had fewer inconsistencies in word usage.

This disparity can be angled from a different perspective—proofreading. An interesting finding is the With MT students probably did not proofread their translations as carefully as the No MT students, possibly owing to the fact that they had something to reference by, so they felt more secure with their output. With a closer examination

of students' translations, the No MT students had made considerably more changes or revisions to their translations, seeing that there were a lot more caret marks, whiteout or eraser traces, cross-outs, arrows switching phrases, etc. on their papers. This implies these students either monitored their output more frequently or they had proofread their translation more carefully. The With MT students made noticeably fewer changes, as a good portion of their papers were clean with few revision marks.

### *Omission*

It was also observed that the With MT students had fewer omissions than their No MT counterparts, especially in Group II. Omission here refers to unintentional omission; that is, words or sentences inadvertently overlooked and left out in the translation, as opposed to intentional omission, which may have happened in situations such as when a student does not know the meaning of a word and leaves a blank space in a sentence. Although intentional omission is also regarded as a binary error, it should not be confused with an inadvertent omission. This demonstrates that MT is of assistance in lowering potential unintentional omissions, since it offers a fully-translated template for a translator to work on.

Inadvertent omission is considered a serious error by translation standards, since it results in the failure to convey the message carried by the source text and may incur losses on the client's part. Shih (2007) showed that students with MT text for reference made fewer errors of missing words/sentences in comparison to those without MT use. Machine translation is faithful in a way that it translates everything in the source text without leaving out words or sentences, however unintelligible they may be once translated.

Then concerning intentional omissions, the With MT sets had significantly fewer ones than the No MT sets, where the discrepancy is strikingly evident in Group II. The crucial role of language proficiency is fully evidenced as a lot of the No MT students in Group II had trouble translating simply because they had too many unfamiliar vocabulary words. For example, the following are some of the words left un-translated, either left blank or written as it is: "mobile device", "detergents", "moisture", "mineral", and many more. Some of these

words have less of an impact on the entire message of a sentence, while some have an adverse effect. Take “mobile device” for instance, it is the core of the passage, so knowing accurately what it means would increase a student’s general comprehension of the text by making logical connections. Six No MT students in Group II failed to translate this phrase, and it can be deduced that they probably did not understand the passage very well, seeing these particular students received low scores for their work. On the other hand, these words were all correctly translated by MT; thus, the With MT students had little problem with them.

### *Register*

Register is defined as “a variety of a language or a level of usage, as determined by degree of formality and choice of vocabulary, pronunciation, and syntax, according to the communicative purpose, social context, and standing of the user” by the Oxford English Dictionary.<sup>4</sup> A good translator should always pay attention to the register of the source text and produce the equivalent in the target text. Hatim and Munday (2004) note that consistency of register is what makes a text hang together. Manipulation of the register is also a common exercise, sometimes even in exams, in translation and interpretation institutions (Sampaio, 2007). Thus, having good command of the register can ensure the level of formality is accurately depicted. The text used in this study is an instruction guide to cell phone care, with many of the sentences written in imperative. An experienced translator would most likely translate the passage with a tone of formality, maintaining a fairly high register in the wording. Google Translate seems to be doing a good job in this respect.

The divergence in register between the With MT and No MT students was evident, regardless of groups. For example, almost none of the No MT students used 您, a euphemized form of 你 (you), in their translations. Over half of the With MT students in Group I used 您, as it is translated by Google Translate, though only a few of them in Group II did the same. Another example is also observed with the translation of “saliva” as 口水 or 唾液, where the latter is considered more formal.

In another note, the majority of students in both Groups failed to avoid the use of 被 when translating the only sentence in the pas-

sage written in passive voice, i.e. “it will have to be replaced,” especially in the With MT sets, since the passive voice was translated into a 被 sentence by Google Translate. The use of passive voice is often used in formal writing in English (Robin, 2008), but not in Chinese, which does not exploit the passive voice frequently in general. Moreover, the word 被 used as a passive verb, more often than not, describes things of inflictive nature (Xiao et al., 2006), e.g. 他被霸凌毆打 (He was beaten up by bullies.) or 她被灼熱的水燙傷了 (She was scalded by the piping hot water.). This shows that perhaps the With MT students were still unable to refine the MT text into a quality translation.

Once again, a disparity between the Groups was detected, but less evident between the With MT sets, where the MT text may have bridged the gaps and led them to produce more similar translations. For instance, more Group I students used higher register wordings such as 切勿 (do not) compared to 不應該 (don't), which is more colloquial. As for the With MT sets, Group I students seemed to have more spare effort to improve on the MT text, such as shortening of words like 銳利的牙齒 to 利齒, and 電子線路 to 電路. Generally, the more concise the Chinese wording, the higher register the text carries. It may be safe to say that the MT text is helpful in enhancing the register of the translation.

### *Source Language Proficiency*

The researchers were astonished by the lack of language proficiency in Group II. The degree of poor vocabulary knowledge extends to not only a couple of students, but to a good portion of the student body. For example, five No MT students in Group II translated “cloth” as 衣物 (clothes) or 毛巾 (towel), which is probably caused by the confusion between clothes and cloth. And 10 No MT students in Group II, out of 34, translated “as well as” into 就像 or 如同 (as...as or such as), 3 of them translated “in dusty areas” as 工廠 or 產業 (industry), another 3, in addition to 2 in the With MT set, translated “knock” as 踢 (kick), which does not even include the many who simply left a blank space or copied the English text. These students obviously had little idea what that particular sentence was about, because the above examples would not fit into the passage

logically. It is also interesting to observe that their translation of the source text was not entirely random, but they opted for words that bear morphological resemblance to the source text, e.g. “in dusty” was mistaken as “industry”, and “knock” mistaken as “kick.” This might be attributed to how they memorized English words. These kinds of error were not at all observed in the With MT set, highlighting once again the significant aid of MT text to students.

Moreover, 8 out of the 34 No MT students in Group II failed to properly translate “mobile device,” either they copied the English in their translations or they put down words unrelated to cell phone, e.g. 自動化儀器 (automatic machine). A possible explanation for this is that students in Taiwan are mostly exposed to American English, while knowing little about other varieties like British English. Had the students known that a mobile phone is merely an alias of a cell phone, they would probably not have any trouble translating this particular term.

### *Target Language Proficiency*

In this study, all the students were translating into their native language, Chinese, but some of them faltered to produce fluent or quality Chinese or deliberately changed the adequate translation by Google Translate into wordy sentences that were not better than the MT text. One reason for this inability may lie in the fact that students were capable of comprehending the source text, but were unable to transfer that information into native equivalent in the target language, as can be seen in their literal translation output. For example, “try not to...” was translated “試著不要...” or “不要嘗試去...” by 2 No MT students in Group I and 17 in Group II, and 4 With MT students in both Groups. Another example was “with proper care of...” which was translated as “隨著(伴隨)...的適當照顧” by 3 No MT students in Group I and 1 in Group II, who actually crossed out 有了, a better translation, and changed it to 隨著. In these cases, the MT text was helpful for students to produce more native-speaker-like translations, instead of translating each word literally without realizing it might not be how we write and speak Chinese in real life.

## *Context-bound Word Choice*

Almost all translations are context-bound, in which the same word in the source text may have completely different meanings in varying contexts. For example, the word “strong” means physically powerful if talking about a bodybuilder, pungent if about smell, and convincing if about a persuasion. In this passage, “strong” is used to describe detergents, giving it the meaning of the power to remove stains, which can be translated as 強效 in Chinese. However, 5 No MT students in Group II translated this adjective into 硬 (hard) or 大力 (power), descriptive words associated with physical strength. It should not be a general trend to make this error because “strong detergents” is one of the three terms appearing side by side, i.e. chemicals and solvents, as things to avoid using on the device, and that, theoretically, would give students a clue to find the appropriate contextual translation for the term. Some other examples: “moisture” (濕氣) in “liquids or moisture can contain minerals” was translated as 滋潤液 or 乳液 by 2 female No MT students, 1 from each group; “break” (損壞) in “rough handling can break internal circuit boards” as 切斷線路 or 造成電路板斷線 by 3 No MT students in Group II. Several No MT students, including ones from Group I, rendered “if you need to clean your phone” as 若你需要清潔你的電話/聽筒, in which “phone” in the context of this passage, should be translated as 手機 (cell phone). Thanks to the statistical MT system employed by Google Translate, these context-bound terms are taken into account in the frequency calculation, as the system looks at chunks of phrases instead of individual words, thus giving the translation contextual meaning. Therefore, the With MT students might have no problem translating these words correctly.

Another example is the translation of “rough handling”, which educated a variety of rendition from students, such as 握 (grip), 擠壓 (compress), 過熱 (overheated), 不當的控制 (inappropriate control), 不平順處理 (uneven handling), 粗糙的處理 (coarse handling), or 堅固的處理 (tough handling). This presents a combination of the lack of language proficiency and the attempt to produce contextually appropriate meaning to the text. The first is evidently a mistranslation of the contextual word “handling” as “grip,” while the subsequent two are the results of poor English comprehension, but the students were likely trying to give the word a contextual meaning by producing words

that may have some relation to the text, like overheating or crushing of cell phones. 不當的控制 (inappropriate control) can be seen as the inability to produce correct Chinese, even though the student could comprehend the source text. The last three are inappropriate contextual processing where both “rough” and “handling” were handled coarsely in generating a best-fit translation under the context.

### *Collocation*

Another problem many students had was the collocation of adjectives and nouns. For example, 乾爽 (dry) is used to describe the comfort of being dry, but a few students used it to describe keeping the cell phone dry; 使用期限 is used to describe the term of use or expiration date, but 10 and 15 No MT students in Groups I and II respectively, and 7 and 11 With MT students in Groups I and II respectively, translated or decided to present the term “lifespan” as 使用期限 (年限) instead of a more appropriate translation—壽命, as rendered by Google Translate. It can be evidently observed that, even with the MT text providing a usable translation, students were unable to recognize the improper collocation of 使用期限 with cell phone. Nonetheless, there were slightly more students with No MT who committed this error than students with MT text for reference.

Knowing commonly used collocations is conducive to producing fluent translations, sometimes even helpful in comprehending the source text, as certain words frequently appear close to another set of words. However, a collocation can be misused in the target text, often times due to an overly hasty glance of the source text, and other times caused by students’ lack of comprehension of the source text, so they opt for a collocation they are familiar with to complete the translation. In this study, a misuse of collocation was observed, or more precisely, a correct collocation was used but the meaning deviated from the source text: “sharp teeth and saliva” was translated as 尖牙利爪 (sharp teeth and claws) by 3 No MT students in Group II, where 1 student even went on to write 會傷到你的手 (will hurt your hands) which is not found anywhere in the source text. The word “saliva” was provided to all students in the form of a vocabulary list; therefore, it should not have been a comprehension difficulty. A possible reason would be they read the text in a hurry and failed to rec-

ognize “sharp teeth” was not collocated with claws in the passage, or they simply ignored the words they do not know and made up translations that might be otherwise correct, considering “sharp teeth and claws” are frequently collocated. Either way, the MT text faithfully translated the words, for no With MT students made this error.

### *Constraints by MT Text*

The examples given above have proved that the MT text is helpful in improving students’ translation output, but nothing is without faults. The imperfection lies mainly in the MT text constraining the students to translate freely as they, though not all, are often confined by the MT text, resulting in translations with rigid sentence patterns or unsatisfactory terminology when the MT text is incorrect to begin with.

The most evident example is the translation of “mobile device,” which was translated by Google Translate as 移動設備. This word choice was considered a non-binary error in this study since it is more commonly used in Mainland China, but not in Taiwan. By contrast, 19 out of the 36 With MT students in Group I and 27 out of the 34 in Group II translated the term as 行動/移動設備(儀器), where 設備 is more of an equivalent to equipment or facilities and 儀器 to apparatus, while most No MT students in both groups translated the term simply as 手機 (cell phone), which is fitting in this passage. As mentioned earlier, the more appropriate translation 行動裝置, which can be found toward the end of some students’ translation even though they did not use it the first time the phrase appeared, was only used by a handful of students. A possible reason for students’ confinement by the MT text is two-fold. Firstly, students might not know what mobile device means, which would make them rely on the MT text for all the information. Secondly, some of them might have known the term, but were unable to use the best words to translate it. The former entails that students were not familiar with simple technology terms in English, while the latter suggests that students were not acquainted with the term in Chinese. A combination of the two would lead the student to copy the MT text and use it directly in their translations.

Another instance is the overuse of pronouns in Chinese and the

adjective “的。” For example, the MT text reads 適當的照顧你的移動設備和它的電池 (with proper care of your mobile device and its battery). Many students copied the sentence almost verbatim, retaining all the pronouns that are grammatically necessary in English. In Chinese, however, dropping these pronouns would not affect the sentence at all, and would probably result in a better translation. In fact, this pattern is also observed in the No MT students, which suggests students are prone to translate an English text word(s) for word(s), much like how machine translation processes a text.

Another example is “here are some useful tips on proper mobile device maintenance” translated by Google Translate as 以下是一些有用的提示正確的移動設備的維修, where there are two adjectives in the sentence. The superfluous use of 的 in the MT text obviously affected how students translated the sentences because nearly all With MT students in Group II and half of those in Group I used at least two 的 in translating this segment. On the other hand, No MT students in Group I were able to use just one 的 in their translations, a more readable rendition in Chinese, e.g. 以下是一些實用的行動裝置維護小秘訣; most Group II students with No MT produced similar translations, such as 這裡提供你一些實用的保養秘訣, using just one 的, even if some of the translations were not entirely accurate. Clearly the use of 的 is often limited at one per sentence, but the MT text seems to misdirect the students in producing uncommon Chinese that encumbers the readability. The superfluous use of pronouns in Chinese was also observed by 賴慈芸 (2003), where the subjects did not have MT text for reference. It is therefore feasible to use MT texts as examples before giving out translation assignments to teach students not to make the same errors.

## *Conclusions*

This study has demonstrated that MT in combination with human intervention at the back-end of the automated process helps students produce more accurate translations than having to translate the source text from scratch, especially those who may lack the sufficient language proficiency to perform such tasks. The findings regarding the differences between the With MT and No MT sets, as well as that between the two Groups, are summarized as follows.

## *Quantitative Findings*

### *Translation Time*

Though the students in both No MT sets finished faster than their counterparts who had MT for reference, the difference was not statistically significant. In addition, if we borrow the results from previous studies such as Allen (2003) and Guerra (2003), it can be inferred that the With MT sets would be faster if they were to perform the translation task on a computer. From the survey results, around 80% of the With MT students in either group expressed that the MT was helpful in improving their translation speed, which if cross-referenced with the actual results would suggest that the MT text boosted students' confidence even if they did not perform faster in reality. This, unfortunately, cannot be proven in this study and merits further investigation in future studies.

### *Translation Accuracy*

MT text was proven to be helpful in improving translation accuracy. The With MT students in both groups performed significantly better than their counterparts without MT as reference, where the number of binary errors was nearly halved in the With MT sets. The converted scores also exhibit statistically significant difference between the With MT and No MT sets, especially in Group II where the With MT set scored almost twice than the No MT set did.

Regarding the ratio of binary and non-binary errors, a noteworthy result showed that with the help of MT, Group II's performance was more on par with that of Group I, that is, no significant difference was observed; whereas a significant gap was observed between the two Groups in the No MT set. The survey results would further back up this result by demonstrating that more With MT students, especially in Group II, thought their translation was mostly accurate, than No MT students.

## *Word Count Analysis*

A highly significant positive correlation ( $r = 0.72$ ) was obtained between the number of words borrowed from the MT text and the students' translation accuracy in Group II whom relied more heavily on the MT for their translations (66.3%) than Group I (58.5%). This finding demonstrates that the more words from the MT text a student used, using sentence as a unit, the less likely a student would have errors translating that particular sentence, which could be a potential indicator to locate problematic or difficult sentences in a passage.

## *Qualitative Findings*

Several qualitative observations and analyses were also made from this study. First of all, No MT students had better word use consistency than With MT students. On the other hand, having MT text as a template to work on helped With MT students to have fewer unintentional or intentional omissions, especially for Group II students.

The MT text was a significant aid to Group II students, making up for their lack of language proficiency to translate parts of the passage. The MT text also helped students produce better Chinese on several occasions and produce context-bound translations. In addition, MT also helped avoid unintentional errors caused by imprinted collocation use in Chinese, since the MT always translates all the words faithfully. However, MT text sometimes confined the sentence patterns, such as the overuse of 的 in Chinese.

There were differences between the two groups of students. Group I students translated the text with a higher register than Group II to allow the text to carry a formal tone. Also, some students were unable to produce native Chinese translations, and were confined by the English source text to produce literal translations.

Lastly, by examining the similarities between the two different sets of students, there were several highly usable parts of the MT text, meaning much time can be saved from needing to type out words.

## *Research Implications*

The findings suggest various benefits for the use of MT, such as facilitating source text comprehension and reducing translation errors. The researchers believe one of the most beneficial aspects of MT lies in its facility to level the playing field among students of dissimilar levels in the same class, especially for beginners in translation. This study has provided empirical evidence that those students who may be deficient in their language proficiency perform considerably better when they have access to an MT text. Using MT in the classroom can boost the confidence of those students who have trouble understanding, let alone translating, the source text. It also allows the instructor to point out things to heed when translating by utilizing the MT text as cases in point. The revision process can also be rewarding because it gives students the opportunity to correct errors, which can be seen as a reversal of role in the classroom where the students are granted with the autonomy to correct the MT text in place of having their translations corrected by the instructor in a conventional classroom setting.

In effect, learning MT may serve as an alternative approach to translation training, as it helps students ease into the process of translation where MT acts as an intermediary that offers possible rendition of the target language, akin to the transition from riding a bicycle with trainer wheels to riding the bike on one's own, with the MT being the trainer wheels. Although the machine translated text is unlikely to be perfect, it could help students save time by automatically translating certain technical terms and repetitive sentences, so more of their cognitive attention could be paid toward refining the translation proper. Both students and the instructor in a translation class may benefit from using this MT technology to learn about the differences between English and Chinese, a win-win situation that should be looked into for future curriculum design.

Furthermore, MT training could lessen student translators' innate skepticism towards MT, and equip them with the necessary knowledge and skills to become a professional translator in tomorrow's market. Future translators are faced with the need to develop new work skills conforming to the multitasked translator profile, which is increasingly becoming a must-have in the translation market. Exposing student translators to MT can help raise this profile.

On the other hand, allowing translation instructors to learn more about MT may help create additional and diverse approaches to translation teaching.

### *Limitations of this Study and Suggestions for Future Works*

Even though this study involved 140 students, a reasonable sample size for statistical analysis, only one relatively short text was used; hence the conclusions in this study cannot be generalized to other text types. Different text genres of longer length may provide more in-depth insights into MT use. In addition, due to resource constraints, the participating students were not allowed access to computers or the Internet, which may be considered a common workstation setup for a professional translator. The best scenario would be for the participating students to receive some sort of MT and post-editing training and conduct the experiment with the use of computers. Since students' translations were only available in hard-copy form, it was difficult to compare the texts in excruciating detail. It would be a lot easier to perform the cross-examinations among students' translations if the target texts were in electronic forms, including the calculation of word use frequency.

Furthermore, Google launched the Google Translator Toolkit in 2009, which integrates its very own Google Translate and the use of translation memories and multilingual glossaries. The interface allows translator to do post-editing works as well as dictionary look-ups online, which would facilitate the translation process. It would be very interesting to train a class of students to use this new tool and compare their translation products with those of a control group who only have access to the Internet, to see how a proper MT and TM tool can help complete a translation task.

Teaching MT PE is necessary for prospective translators to develop this new work mode that caters to the growing translation demands. We cannot undervalue its advantages, because it might present the solutions to deal with the escalating pressure of increased volume of translation jobs that call for shorter translation turn-around cycles to sustain multilingual content for simultaneous disseminations. The current market trend has made MT a demand, and

the productive results have step by step clear the bad name MT once endured. One should learn to embrace it, and make the best out of it.

## Notes

1. <http://europe.nokia.com/support/learn-how/caring-for-your-mobile-phone/proper-care> (assessed on June 20, 2009)
2. <http://translate.google.com/#>(accessed on September 22, 2009)
3. <http://translate.google.com/support/>(Last accessed April 24, 2011)
4. <http://oxforddictionaries.com/definition/register> (Last accessed July 11, 2011)

## References

- 賴慈芸 (2003)。他們走了多遠？——大學部學生、翻譯所學生與專業譯者的翻譯表現比較。第八屆口筆譯教學研討會，臺北：國立臺灣師範大學。
- Allen, J. (2001). Post-editing or no post-editing? *International Journal for Language and Documentation*, 8, 41-42.
- Allen, J. (2003). Post-editing. In H. Somers (Ed.), *Computers and translation: A translator's guide*. (pp. 297-317). Amsterdam & Philadelphia: Benjamins.
- Allen, J. (2004). *Case study: implementing MT for the translation of pre-sales marketing and post-sales software deployment documentation* (pp. 1-6). Paper presented at the 6th Conference of the Association for Machine Translation in the Americas, AMTA 2004, Washington, DC, USA
- Arnold, D. J., Balkan, L., Meijer S., Humphreys, R.L., & Sadler, L. (1994). *Machine translation: An introductory guide*. London: Blackwells-NCC.
- Bowker, L. (2002). Why do translators need to learn about technology? *Computer-aided translation technology: A practical introduction* (pp. 11-21). Ottawa: University of Ottawa Press.
- Chan, Y. S., & Ng, T. H. (2008). *MAXSIM: A maximum similarity metric for machine translation evaluation*. Paper presented at the Association for Computational Linguistics, Columbus, Ohio, USA.
- Chang, P.C., Galley, M., & Manning, C.D. (2008). *Optimizing Chinese word segmentation for machine translation performance*. Paper presented at the Third Workshop on Statistical Machine Translation, Columbus, Ohio, USA.
- Chomsky, N. (1975). *The logical structure of linguistic theory*. Chicago: University

of Chicago Press.

- Fulford, H. (2002). *Freelance translators and machine translation: An investigation of perceptions, uptake, experience and training needs*. Paper presented at the 6th European Association of Machine Translation Workshop, UMIST, Manchester.
- Gaspari, F. (2001). *Teaching machine translation to trainee translators: A survey of their knowledge and opinions*. Paper presented at the MT Summit VIII Workshop on Teaching Machine Translation, Santiago de Compostela, Spain.
- Gross, A. (1992). Limitations of computers as translation tools. In J. Newton (Ed.), *Computers in translation — A practical appraisal*. London: Routledge.
- Guerra, L. (2003). *Human translation versus machine translation and full post-editing of raw machine translation output*, Unpublished (master's thesis), Dublin City University, Dublin.
- Guzmán, R. (2007). Manual MT post-editing: "If it's not broken, don't fix it!". *Translation Journal*, 11, 4.
- Hale, S., & Campbell, S. (2002). The interaction between text difficulty and translation accuracy. *Babel*, 48(1), 14-33.
- Hatim, B., & Munday, J. (2004). *Translation: An advanced resource book*. London and New York: Routledge
- Hutchins, J. (1995). Machine translation: A brief history. In E. Koerner (Ed.), *Concise history of the language sciences: From the Sumerians to the cognitivists* (pp. 431-445). Oxford: Pergamon Press.
- Hutchins, J. (2001). Machine translation. In S. W. Chan & D. E. Pollard (Eds.), *An encyclopaedia of translation*. Hong Kong: The Chinese University Press.
- Hutchins, J. (2003). Computer-based translation systems and tools. Retrieved October 15, 2009 from <http://www.hutchinsweb.me.uk/BCS-NLT-2003.pdf>
- Hutchins, J. (2005). Example-based machine translation: A review and commentary. *Machine Translation*, 19, 197-211.
- Hutchins, J., & Somers, H. L. (1992). *An introduction to machine translation*. London: Academic Press.
- Itagaki, M., Aikawa, T., & He, X. (2007). *Automatic validation of terminology translation consistency with statistical method*. Paper presented at the MT Summit XI, Copenhagen, Denmark.
- Kliffer, M. D. (2005). *An experiment in MT post-editing by a class of intermediate/advanced French majors*. Paper presented at the EAMT 10th Annual Conference, Budapest, Hungary.

- Liu, Q., Hou, H., Lin, S., & Qian, Y. (2005). *Introduction to China's HTRDP machine translation evaluation*. Paper presented at the MT Summit X, Phuket, Thailand.
- Mauser, A., Hasan, S. & Ney, H. (2008). *Automatic evaluation measures for statistical machine translation system optimization*. Paper presented at the Language Resources and Evaluation Conference Marrakech, Morocco.
- Newton, J. (1992). Introduction and overview. In J. Newton (Ed.), *Computers in translation — A practical appraisal*. London: Routledge.
- NIST. (2008). NIST 2008 open machine translation evaluation — official evaluation results. Retrieved October 15, 2009, from [http://www.itl.nist.gov/iad/mig//tests/mt/2008/doc/mt08\\_official\\_results\\_v0.html](http://www.itl.nist.gov/iad/mig//tests/mt/2008/doc/mt08_official_results_v0.html)
- Pym, A. (1992). Translation error analysis and the interface with language teaching. In C. Dollerup, Loddegaard, Anne (Ed.), *Teaching translation and interpreting: training, talent and experience* (pp. 279-288). Amsterdam & Philadelphia: John Benjamins.
- Richmond, I. M. (1994). Doing it backwards: Using translation software to teach target-language grammaticality. *Computer Assisted Language Learning*, 7, 65-78.
- Robin, T. (2008). *Practice makes perfect: Intermediate English grammar for ESL Learners*. New York: McGraw-Hill.
- Sampaio, G. R. L. (2007). Mastering sight translation skills. *Tradução & Comunicação*, 16, 63-69.
- Senez, D. (1998). *Post-editing service for machine translation users at the European Commission*, In Proceedings of the ASLIB conference on Translating and the Computer 20, London, UK.
- Shih, C. (2006). *Helpful assistance to translators: MT & TM*. Taipei: Bookman Books, Ltd.
- Shih, C. (2007). Mapping out students' translation process: An MT-specific comparative study. *翻譯學研究集刊*, 10, 163-190.
- Somers, H. (2003). *Computers and translation: A translator's guide*. Amsterdam: Benjamins.
- Stix, G. (2006). The elusive goal of machine translation. *Scientific American*, 294, 92-95.
- Tanner, A. (2007). Google seeks world of instant translations. Retrieved November 21, 2009 from <http://www.reuters.com/article/idUSN1921881520070328>
- Wong, T. M. (2008). *Machine translation and evaluation: Online systems, Unpub-*

lished (master's thesis). City University of Hong Kong.

Xiao, R., McEnery, T., & Qian, Y. (2006). Passive constructions in English and Chinese: A corpus-based contrastive study. *Language in Contrast*, 6, 109-149.

Yuste, E. (2001). *Making MT commonplace in translation training curricula — Too many misconceptions, so much potential!* Paper presented at the MT Summit VIII Workshop on Teaching Machine Translation, Santiago de Compostela, Spain.

## Appendix 1: Text for No MT and With MT Sets

Proper care is about appropriate storing, cleaning, charging and usage, as well as recycling after the device lifespan has reached its end. With proper care of your mobile device and its battery, you can lengthen its lifespan, so take good care of your mobile device. Here are some useful tips on proper mobile device maintenance that will help you get the best out of your mobile companion. You should not use or store your device everywhere, for example in dusty, dirty areas as these can damage its moving parts and electronic components. If you need to clean your phone or its camera lenses, use a soft, clean, dry cloth. Do not use harsh chemicals, cleaning solvents, or strong detergents on your mobile device! Charge the battery properly to extend its life although eventually it will have to be replaced. Keep the device dry. Precipitation, humidity, and all types of liquids or moisture can contain minerals that will corrode electronic circuits. Do not let your pet play with your mobile device. Sharp teeth and saliva can seriously harm the device and the battery. Try not to drop, knock, or shake your mobile device. Rough handling can break internal circuit boards and fine mechanics and dropping can harm the battery.

## Appendix 2: Text for With MT Set

適當的照顧是關於適當儲存，清洗，充電和使用，以及回收設備後，壽命已經走到了盡頭。適當的照顧你的移動設備和它的電池，可以延長其壽命，所以要照顧好您的移動設備。以下是一些有用的提示正確的移動設備的維修，幫助您獲得最佳的出您的移動伴侶。你不應該使用或存放設備無處不在，例如在有灰塵或骯髒的地方，因為這些會損害它的可拆卸部件和電子元件。如果您需要清潔您的手機或相機鏡頭，使用柔軟，

乾淨，幹布。不要用烈性化學製品，清洗劑或強洗滌劑在移動設備上！電池充電適當延長其生命，但最終它必須被替換。保持設備乾燥。雨水，濕氣和各種液體或水分可能含有礦物質，會腐蝕電子線路。不要讓您的寵物玩您的移動設備。銳利的牙齒和唾液會嚴重損害設備和電池。盡量不要扔放，敲打或振動您的移動設備。粗暴地對待手機會毀壞內部電路板及精密機械與下跌可能傷害電池。

### Appendix 3: The Vocabulary List

- circuit board 1. 線路板, 電路板
- component 1. ( 機器、設備等的 ) 構成要素 ; 零件 ; 成分
- corrode 1. 侵蝕 ; 損害
- lifespan 1.( 生物的 ) 壽命 , 預期生命期限 2.( 物的 ) 使用期限
- maintenance 1. 維持 , 保持 2. 維修 , 保養 3. 堅持 , 主張 4. 扶養 ; 生活費
- mechanics 1. 力 ( 學 ) ; 機械 ( 學 ) 2. 技術性的部分 ; 技術 ; 技巧
- precipitation 1. 猛然落下 ( 或摔下 ) ; 猛衝 2. 沈澱 3. 雨 ; 雪 4. 降雨 ( 雪 ) 量
- saliva 1. 涎 , 唾液
- solvent 1. 溶媒 , 溶劑 2. 解決方法

### Appendix 4: Examples of post-editing and binary or non-binary errors

- a. (ST) Charge the battery properly to extend its life although eventually it will have to be replaced.

(MT) 電池充電適當延長其生命，但最終它必須被替換。

For this particular sentence, the MT-produced 電池充電適當延長其生命 is considered a binary error, since a syntactic error is evident in

that the adverb 適當 precedes the verb 充電 ; while 但最終它必須被替換 is deemed to have non-binary errors, because of the use of the pronoun 它 and the passive construction of 被 . Given that a binary error is present, this sentence is flagged to be one with binary error.

(PE) 電池適當充電以延長其壽命，但最終仍必須更換。

Here, word reordering is used by swapping 適當 and 充電 . A preposition 以 is added to explain the connection between proper recharge and battery life. A word change of 生命 , a non-binary error if not corrected, into 壽命 is performed. The pronoun 它 and passive voice 被 are deleted, while 仍 is added to emphasize “it will have to be.” Another word change is in place to replace 替換 with 更換 , which is a better word choice in describing battery replacement.

b. (ST) Do not let your pet play with your mobile device. Sharp teeth and saliva can seriously harm the device and the battery.

(MT) 不要讓您的寵物玩您的移動設備。銳利的牙齒和唾液會嚴重損害設備和電池。

This sentence is correct in meaning and in syntax, with minor word choice problems that can be easily corrected; hence, this sentence is considered to have only non-binary error.

(PE) 不要讓寵物玩您的行動裝置，銳牙和唾液會嚴重損壞裝置和電池。

Here, the excessive possessive pronoun 您的 is dropped. A few changes in word choices are made, such as editing 移動設備 into 行動裝置 and 損害 into 損壞 . Meanwhile, 銳利的牙齒 is streamlined into 銳牙。

