science, engineering and health studies

Short Communication

Science, Engineering and Health Studies
https://lio1.tci-thaijo.org/index.php/sehs
ISSN (Online): 2630-0087

Source text pre-editing versus target text post-editing in using Google Translate to provide health services to culturally and linguistically diverse clients

Yingping Liang[1] *and Weifeng Han*[2*]

[1] College of Foreign Languages, Taiyuan University of Technology, Taiyuan 030024, China
[2] Institute of Health and Wellbeing, Federation University, Melbourne 3000, Australia

***Corresponding author**:
Weifeng Han
w.han@federation.edu.au

Received: *30 June 2021*
Revised: *07 July 2021*
Accepted: *06 September 2021*
Published: *28 September 2022*

Citation:
Liang, Y., and Han, W. (2022). Source text pre-editing versus target text post-editing in using Google Translate to provide health services to culturally and linguistically diverse clients. Science, Engineering and Health Studies, 16, 22050009.

ABSTRACT

The study reports early-stage research on the efficacy of using the source text pre-editing (STPE) method to improve translation accuracy and cost-efficiency in conjunction with Google Translate as compared to the traditional target text post-editing (TTPE) method. Based on fluency, accuracy, cultural appropriateness and error severity, preliminary results show that STPE significantly increased the meaning adequacy and accuracy in translation as compared to TTPE. STPE also saved significant time, and, therefore, was more cost-efficient, as compared to TTPE. The results suggested a fundamentally new and more efficient method to the better employment of machine translation that differed from existing approaches. Governments and health providers may use the STPE plus Google Translate method more widely to reduce translation inaccuracy as well as to increase cost-efficiency, and provide more accessible information to culturally and linguistically diverse clients.

Keywords: culturally and linguistically diverse clients; Google Translate; source text pre-editing; target text post-editing

1. INTRODUCTION

Providing health services to culturally and linguistically diverse (CALD) communities, including providing consultation, assessment/test information, discharge instructions, etc., sees increasing use of machine translation (MT) tools, such as Google Translate. In 2018, Anil Sabharwal from Google Photos and Communication showcased in Sydney how AI could automate laborious tasks (Bhatt, 2018). This includes using Google Translate in multilingual and multicultural communications. Doctors are using Google Translate to overcome language barriers in client consultations (Moberly, 2018), and Google Translate is encouraged in remote/telehealth contexts where there is a language gap between the practitioner and the client (Wade, 2011). In the context of the Covid-19 pandemic, an unprecedented amount of government health-related information and advice needs translating and communicating to CALD speakers, usually on short notice, which provides an opportunity to promote the use of Google Translate due to it being the most easily available and free initial mode of communication to the CALD community (Patil and Davies, 2014). For example, the Australian government has been employing Google Translate for its public Covid-19 updates, guidelines, and notices, especially considering that the Department of Home Affairs alone had spent more than $2 million on translating Covid-19 related materials using trained human

translators (Dalzell, 2020).

While MT is both time- and cost-saving, it appears to be more suitable for short and simple phrases than complex sentences (Beh and Canty, 2015). It is less accurate and reliable on non-medical terms than medical terms (Patil and Davies, 2014), and it does not adequately address the cultural aspects of translation (Guo, 2016). Simply relying on MT in the public health service sector may result in providing the CALD community misleading information, e.g., mistakes in the Spanish version of the Obamacare (Affordable Care Act) (Abdullah, 2014; Petri, 2014), as well as the CALD community losing trust in the government, e.g., mistranslation in the Chinese (Mandarin) version of Covid-19 updates by the Department of Home Affairs, Australia (Dalzell, 2020). Although the MT plus target text post-editing (TTPE) service is proposed to be effective in mediating mistranslation (Turner et al., 2015), the MT+TTPE model is as costly (both in time and money) as employing human translators (Ponce et al., 2006; Raynor, 2016; Sentell et al., 2015).

The early 1980s saw the emergence of implementing pre-editing and post-editing of raw materials (e.g., the European Commission). However, the new millennium has seen MT plus post-editing as the mainstream of translation workflow (Lommel and DePalma, 2016). Post-editing is defined as a process "to edit, modify and/or correct pre-translated text that has been processed by an MT system from a source language into (a) target language(s)" (Allen, 2003). That is, post-editing is an error correction process under specific quality criteria of the MT output against the original text. Therefore, data-driven approaches, for example by using a set of regular expressions based on error patterns found in the MT output, are recommended in MT post-editing (Chatterjee et al., 2015). Pre-editing, on the other hand, involves the use of a set of terminological and stylistic guidelines or rules to prepare the original text before translation automation to improve the raw output quality (Arenas, 2020). It focuses on the controlled language of the source text where vocabulary and grammar are intended for very specific domains, as well as purposes (O'Brien, 2003).

Although both pre-editing and post-editing have co-existed for over 30 years and have shown advantages of one over the other in different contexts, there has not been enough discussion on the efficacy of source text pre-editing (STPE) and MT to provide unbiased and effective health services to CALD clients in the health sector. Therefore, in this study, we investigated whether STPE outperforms TTPE in using Google Translate to provide health services to culturally and linguistically diverse clients.

2. MATERIALS AND METHODS

The study consisted of both the development of the STPE+MT method and a user study of stakeholders' perception of STPE+MT in comparison to MT+TTPE. The development of the STPE+MT method is because MT is effective for words and phrases, but not suitable for more complex texts, e.g., those involving complex grammar or word order (Lear et al., 2016). The truth is around two-thirds of MT translation errors in the health service sector are from complex grammar and/or word order (Turner et al., 2015). Another consideration is the fact that many erroneous MT translations are results of inaccurate, idiosyncratic, or highly abbreviated English source texts (Khoong et al., 2019). In the case of Google Translate, the likelihood of incorrect translation sees an increase when the original English texts involve complicated sentence structures and vocabularies that require higher grade levels of comprehension (Chen et al., 2016). The STPE method helps to reduce the grammatical complexity in the source text while keeping the core meaning of the original message. This is rooted in the bottom-up translation theory (Munday, 2016) that prioritizes the semantics of key lexicons over complex grammatical rules and/or word order in the source language and then transfers them into those in the target language.

As for the user study, first, from the Australian state and national public health websites, an initial of 25 public health documents (around 25,000 words in English) including health guidelines, and Covid-19 related instructions/updates were collected. The English documents were then translated into Chinese (Mandarin) through both the MT+TTPE method and the STPE+MT method.

2.1 The MT+TTPE method

The English documents were first translated into Chinese using Google Translate. The results were then post-edited by two native Chinese speakers with health backgrounds. Specifically, the post-editing ensures that the translation is consistent with Chinese grammar and adequately represents the meaning of the English text while being culturally appropriate (Turner et al., 2015). Post-editing time was measured.

2.2 The STPE+MT method

The English documents were first pre-edited by two native English speakers with a background in the health sector. Specifically, keywords of each complex sentence/phrase are highlighted, and the order of the keywords is (re-)arranged as per the two most frequent sentence types in health documents, i.e., the declarative (in the noun+verb word order) and the directive (in the verb+noun word order). Pre-editing time was measured and compared against the post-editing time. The texts were then translated into Chinese using Google Translate. The example below illustrates how an original complex sentence is pre-edited and rewritten into shorter sentences with simpler grammatical structures and vocabularies (see 2.3 and 2.4). An example of Google Translation of the pre-edited text follows that (Figure 1).

2.3 Original text

Here is an example of the original text: "The Government's focus on containing the virus through testing, tracing and border and travel measures, building and investing in the capacity of our health system, and collaborating with states and territories, the health sector and others, has helped stop the spread of COVID-19" (Office of the Prime Minister of Australia, 2021).

2.4 Edited text

Here is an example of the pre-edited text: "The government has three focuses. First, the government contains the virus through testing measures, tracing measures, border measures, and travel measures. Second, the government

invests in our health system capacity. Third, the government collaborates with states and territories. It also collaborates with the health sector and other sectors. All these focuses have helped stop the spread of COVID-19".

2.5 Google translation of the pre-edit text

The MT result in Figure 1 was highly consistent with the pre-edited text regarding its content. There are no noticeable grammatical errors, and the translation was culturally appropriate to Chinese speakers and, therefore, highly acceptable.

The STPE+MT translations were blindly rated and compared against that of MT+TTPE by 6 English-Chinese bilingual speakers with a health background. Both the STPE+MT and MT+TTPE translations were based on an evaluation rubric, including four scales of quality: fluency, accuracy, cultural appropriateness, and error severity (Chen et al., 2016) (Table 1).

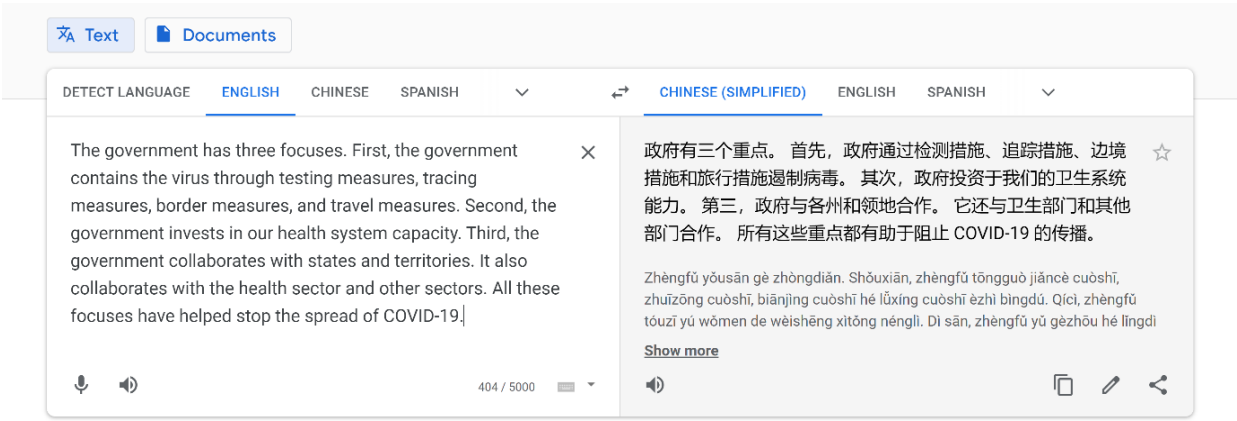

Figure 1. Sample of STPE Google translation: English to Mandarin

Table 1. Evaluation rubric

Score	Quality			
	Fluency	**Accuracy**	**Cultural appropriateness**	**Error severity**
5	High fluency. Almost no grammatical errors	All information conveyed from the original	Highly appropriate. The translation always tailors the linguistic and cultural norms of the target population	No negative effect on patient care
4	Good fluency. Minor grammatical errors	Most information conveyed from the original	Mostly appropriate. The translation, in most cases, tailors the linguistic and cultural norms of the target population	Unclear effect on patient care
3	Medium fluency. Occasional grammatical errors, but does not affect comprehension	Half information conveyed from the original	Occasionally appropriate. The translation occasionally tailors the linguistic and cultural norms of the target population	Delayed patient care
2	Marginal fluency. There are a few grammatical errors. The errors affect comprehension	Partial (less than half) information conveyed from the original	Mostly inappropriate. The translation usually does not tailor the linguistic and cultural norms of the target population	Negative effect on patient care
1	Low fluency. Full of grammatical errors. Not understandable	No information was conveyed from the original	Inappropriate. The translation does not tailor the linguistic and cultural norms of the target population	Dangerous to patient

3. RESULTS

3.1 Inter-rater reliability

Six English-Chinese bilingual speakers with a health background compared the quality of the STPE+MT translations and the MT+TTPE translations. Each rater reviewed the 25 sets of translations. Each set consisted of the original English text, the STPE+MT version, and the MT+TTPE version. The translations were not labelled as per which method was employed. The raters were asked to evaluate each set based on the four dimensions in Table 1.

Cronbach's alpha was used to assess the degree of agreement between the raters. Two sets of means from the MT+TTPE and STPE+MT methods were calculated to represent the scores in each of the four domains of fluency, accuracy, cultural appropriateness, and error severity. Cronbach's alpha value exhibited high degrees of agreement on the rating outcome of the MT+TTPE method (.903) and the STPE+MT method (.926).

3.2 Editing time and quality rating

The time taken for both pre-editing and post-editing was calculated (Table 2). Specifically, the number of edited words per minute (WPM) (Turner et al., 2015) was calculated for each document based on the recorded time of each pre-editor/post-editor. WPM was calculated as a metric to evaluate the quality of the translation (Van de Velde et al., 2015). The mean and SD in WPM for each editor were also calculated to understand the correlations between editing time and translation quality, as well as editor differences. On average, a pre-editor corrected around 42 WPM with a variation of 11 WPMs. In comparison, a post-editor corrected around 35.5 WPM with a variation of 10 WPMs.

As per the quality ratings, while there were no significant differences between the two methods in the domains of fluency, cultural appropriateness, and error severity, the 25 STPE+MT translations (Mean = 4.2, SD = 0.6) received significant higher ratings in accuracy than the MT+TTPE translations (mean = 3.5, SD = 0.5), $t(48) = 4.16$, $p < 0.01$ (95% CI, 0.34-0.96). Overall, the STPE+MT method was more time-efficient and yielded semantically more loyal translations based on the source texts (Table 3).

Table 2. Editing time

Editor	Number of texts edited	Number of edited words per minute, mean (SD)
Pre-editor (A)	11	42.3 (10.2)
Pre-editor (B)	14	41.7 (11.9)
Post-editor (A)	12	34.6 (9.3)
Post-editor (B)	13	36.6. (10.8)

Table 3. Translation quality rating

Editor	Average fluency rating	Average accuracy rating	Average cultural appropriateness rating	Average error severity rating
Pre-editor (A)	4.3	4.0	3.9	4.2
Pre-editor (B)	4.1	4.4	4.2	3.9
Post-editor (A)	4.1	3.6	4.4	4.1
Post-editor (B)	4.4	3.5	4.2	4.3

4. DISCUSSION

4.1 Principal findings

Although MT is widely used to translate clinician-entered, patient-specific medical and health instructions, particularly to large CALD communities, such as the Spanish- and Chinese-speaking patients, significant inaccuracies, mainly arising from different grammatical structures, e.g., word order, between the source and the target languages are not uncommon (Khoong et al., 2019). The most widely used method to mediate inaccuracy resulted from MT is the employment of TTPE, which is as costly and less efficient as human translators.

The current study hypothesized that the STPE+MT method is more time- and cost-efficient than the MT+TTPE method and will result in more accurate and culturally appropriate translation while providing health services to CALD communities. The evaluation of stakeholder's attitudes demonstrates this hypothesis to be true. STPE showed a higher WPM as compared to TTPE. Therefore, it is less time-consuming. The Cronbach's alpha showed a high degree of agreement between the raters, and there were significantly higher ratings in meaning accuracy with STPE than TTPE. This shows that pre-editing complex sentences into simple grammatical structures helps to keep the core meaning of the original text. On the other hand, as TTPE focuses on the target text, the editors had the opportunity to refine the form of the target text as per the linguistic and cultural appropriateness. This is why there were no significant differences between the STPE and TTPE methods in the domains of fluency, cultural appropriateness, and error severity. However, the fact that the STPE+MT method, without further editing of the target text, received no poorer ratings in these domains as compared to the MT+TTPE method, showed that STPE+MT is as reliable in these domains.

Overall, the current study suggests a fundamentally new and more efficient method to the better employment of Google Translate that differs from currently proposed approaches.

4.2 Limitations

One limitation of the current study is it employed only Google Translate as the MT method. Considering there are other major online translation services, such as Bing Translator, YouDao FanYi, etc., the efficacy of STPE+MT needs widely testing on multiple platforms. Another limitation lies in the employment of untrained editors. While STPE+MT showed a better result in translation adequacy in this study, additional work is needed to improve the overall quality of the translations, as compared to human translators. Trained editors, in this sense, may be of help. Further to the ratings, this study did not involve CALD users' experience from the clients' perspective. Finally, statistically, there is a small sample size in this study. Future studies could increase the sample size to increase the statistical power for generalizability.

5. CONCLUSION

While there exist works on both the easiness of access and potential risks of using Google Translate in providing health services to CALD clients, previous research focuses on how the MT+TTPE method is used to mediate mistranslation by MT only, a method as costly as human translators. There is a lack of mature research investigating alternative methods in reducing translation inaccuracy while keeping cost-efficiency. This study trialled a novel STPE+MT method and the results showed it helped to reduce translation inaccuracy as well as to increase cost-efficiency in comparison to the MT+TTPE method. In the future, governments and health providers

may use the STPE plus Google Translate method more widely to reduce translation inaccuracy as well as to increase cost-efficiency, and provide more accessible information to culturally and linguistically diverse clients.

REFERENCES

Abdullah, H. (2014). Lost in translation: Obamacare en Español. *CNN Politics*. [Online URL: https://edition.cnn.com/2014/01/15/politics/obamacare-spanish-language-site/index.html] accessed on March 1, 2021.

Allen, J. (2003). Post-editing. In *Computers and Translation: A Translator's Guide* (Somers, H. ed.), pp. 297-317, Amsterdam: John Benjamins Publishing.

Arenas, A. G. (2020). Pre-editing and post-editing. In *The Bloomsbury Companion to Language Industry Studies* (Angelone, E., Ehrensberger-Dow, M., and Massey, G., eds.), pp. 333-360. London: Bloomsbury.

Beh, T. H. K., and Canty, D. J. (2015). English and Mandarin translation using Google Translate software for pre-anaesthetic consultation. *Anaesthesia and Intensive Care,* 43(6), 792-793.

Bhatt, N. (2018). Google showcases Australian AI successes: conservation, language and health. *Ausdroid*. [Online URL: https://ausdroid.net/2018/05/31/google-showcases-australian-ai-successes-conservation-language-health/] accessed on March 1, 2021.

Chatterjee, R., Weller, M., Negri, M., and Turchi, M. (2015), Exploring the planet of the apes: a comparative study of state-of-the-art methods for MT automatic post-editing. In *Proceedings of the 53rd Annual Meeting of the Association for Computational Linguistics and the 7th International Joint Conference on Natural Language Processing (Volume 2: Short Papers)*, pp. 156-161, Beijing, China.

Chen, X., Acosta, S., and Barry, A. E. (2016). Evaluating the accuracy of Google Translate for diabetes education material. *JMIR Diabetes,* 1(1), e3.

Dalzell, S. (2020). Federal government used Google Translate for COVID-19 messaging aimed at multicultural communities. *ABC News*. [Online URL: https://www.abc.net.au/news/2020-11-19/government-used-google-translate-for-nonsensical-covid-19-tweet/12897200] accessed on March 1, 2021.

Guo, J. W. (2016). Is Google Translate adequate for facilitating instrument translation from English to Mandarin? *Computers, Informatics, Nursing,* 34(9), 377-383.

Khoong, E. C., Steinbrook, E., Brown, C., and Fernandez, A. (2019). Assessing the use of Google Translate for Spanish and Chinese translations of emergency department discharge instructions. *JAMA Internal Medicine,* 179(4), 580-582.

Lear, A., Oke, L., Forsythe, C., and Richards, A. (2016). "Why can't I just use Google Translate?" a study on the effectiveness of online translation tools in translation of Coas. *Value in Health,* 19(7), A387.

Lommel, A. R., and DePalma, D. A. (2016). Post-editing goes mainstream: how LSPs use MT to meet client demands. Cambridge MA: Common Sense Advisory.

Moberly, T. (2018). Doctors choose Google Translate to communicate with patients because of easy access. *BMJ Clinical Research,* 362, k3974.

Munday, J. (2016). *Introducing Translation Studies: Theories and Applications*, 4th, pp. 141-168. London: Taylor and Francis Group.

O'Brien, S. (2003). Controlling controlled English: an analytical of several controlled language rule sets. In *Proceedings of EAMT-CLAW 2003*, pp. 105-114, Dublin, Ireland.

Office of the Prime Minister of Australia. (2021). Over $1.1 billion to extend Australia's covid-19 health response. [Online URL: https://www.pm.gov.au/media/over-11-billion-extend-australias-covid-19-health-response] accessed on 15 March 2021.

Patil, S., and Davies, P. (2014). Use of Google Translate in medical communication: evaluation of accuracy. *BMJ,* 349, g7392.

Petri, A. (2014). Spanish version of healthcare.gov apparently used computer translation? *The Washington Post*. [Online URL: https://www.washingtonpost. com/blogs/compost/wp/2014/01/13/spanish-version-of-healthcare-gov-apparently-used-computer-translation/] accessed on March 1, 2021.

Ponce, N. A., Hays, R. D., and Cunningham, W. E. (2006). Linguistic disparities in health care access and health status among older adults. *Journal of General Internal Medicine,* 21(7), 786-791.

Raynor, E. M. (2016). Factors affecting care in non-English-speaking patients and families. *Clinical Pediatrics,* 55(2), 145-149.

Sentell, T. L., Tsoh, J. Y., Davis, T., Davis, J., and Braun, K. L. (2015). Low health literacy and cancer screening among Chinese Americans in California: a cross-sectional analysis. *BMJ Open,* 5(1), e006104.

Turner, A. M., Dew, K. N., Desai, L., Martin, N., and Kirchhoff, K. (2015). Machine translation of public health materials from English to Chinese: a feasibility study. *JMIR Public Health and Surveillance,* 1(2), e17.

Van de Velde, S., Macken, L., Vanneste, K., Goossens, M., Vanschoenbeek, J., Aertgeerts, B., Vanopstal, K., Vander Stichele, R., and Buysschaert, J. (2015). Technology for large-scale translation of clinical practice guidelines: a pilot study of the performance of a hybrid human and computer-assisted approach. *JMIR Medical Informatics*, 3(4), e33.

Wade, R. G. (2011). Try Google Translate to overcome language barriers. *BMJ*, 343, d7217.