Potentials of Post-MT Editing as a Teaching Method to Promote Translation Quality & Time Spent

Mohammed Mohammed Al-Aqeeli
Dr. B. S. Jadhav

===

Abstract

The current study aims at providing a substantial knowledge that helps integrating and streamlining the post editing process of the entire machine translation production flow. It attempts to quest for the potentialities that may show how to strengthen the students/translators' abilities of using MT. Thus, students and instructors in a translation class can benefit from using this technology to know more about the differences between any two languages, making this area of research worthy of investigation. It investigates whether MT is useful to produce an eligible translation when students/translators come to post edit and correct errors of a given text or not. It aims at gathering an understanding of what happens when students come to post edit the raw MT output: How do they deal with it; do they get less errors; do they save time; and what results they generally gain when using MT in translating texts.

Key words: Machine Translation (MT), Post-editing, Raw MT output, With Ready Machine Translated Texts (WRMTT), No Ready Machine Translated Texts.

Introduction

Issues of utilizing technology in the field of language have become a vital aspect and many researchers of computational linguistics are opting for it as it provides an incredible assistance for humans in various aspects of life. Therefore, they are rapidly moving towards discovering and suggesting ways of increasing the whole translation productivity process, including both human and MT. They applied themselves to incorporate the human correction activities within the translation process itself, thereby shifting the MT model to that of computer-assisted translation, considering the essential role of the human in such a process and not neglecting him since s/he is the inventor of the MT itself.

Building a fully automatic machine translation that did not require any intervention by the human was the primary objective of machine translation. At a 1953 conference, however, Yehoshua Bar-Hillel, one of the first MT researchers, reported that building a fully automatic translation system was improbable and years later still remained convinced that the fully automatic high-quality machine translation system was essentially unattainable (Bar-Hillel 1960/2003: 45). As a replacement for fully automatic translation, machine translation has emerged in its place, which is placed between Fully Automatic High Quality Translation FAHQT and Human-aided Machine Translation HAMT. The main aim of machine translation is still to generate translation automatically, but it is no longer required that the output quality is high, rather than it is high, rather that it is fit-for- purpose (C.K. Quah, 2006: 7). In fact, discussing the issue of machine translation is opening a number of issues relating to translation and technology. These issues have

Language in India www.languageinindia.com **ISSN 1930-2940 15:1 January 2015**
Mohammed Mohammed Al-Aqeeli and Dr. B. S. Jadhav   13
Potentials of Post-MT Editing as a Teaching Method to Promote Translation Quality & Time Spent

confirmed that translation theory has been highlighted by linguistic theories and various formalisms in the development of rule-based machine translation systems.

Researchers need to realize that talking about MT is a try of discovering or suggesting ways of increasing the productivity of the whole translation process (MT plus human work) and to incorporate the human correction activities within the translation process itself, thereby shifting the MT paradigm to that of computer-assisted translation. It is not neglecting the essential role of the human since s/he is the inventor of the MT itself.

A. Background

The first real concern about Machine Translation (MT) started after the Second World War. At that time, Alan Turing's team in the United Kingdom invented the first computer as part of the new famous code-breaking operation at Bletchley Park (Hinsley and Stripp 1993). With the beginning of the Cold War in the late 1940s, one finds significant investment for the cause of MT by the US government in automatic Russian–English translation systems for the military purposes. Similar but small programs were invented by France, Japan, and the Russian. These first-generation systems were known as 'direct' systems since they were basically word-based 'direct replacement' systems; each SL word would be looked up and replaced by a corresponding TL term. Word-for-word substitution is not a solid base for translation. Without significant progress, MT's reputation fell very low in the 1960s following damning criticism by Yehoshua Bar-Hillel in his Report on the State of Machine Translation in the United States and Great Britain (1959) and in the report published in 1966 by the Automatic Language Processing Advisory Committee (ALPAC). Instead the focus shifted to more basic questions of language processing, the field that became known as computational linguistics (Jeffery Davis 2007).

Utilizing computers to translate text from one natural language to another with or without human intervention was the notion argued by many researchers, Hutchins was the famous one. But it was until the 1970s that various research projects led to working, commercially available systems (Arnold et al., 1994).

Machine translation refers to automated process that utilizes computers to translate text from one natural language to another with or without human intervention (Hutchins, 1992). The internet boom in the 1990s ushered MT into a new era, as the numerous amount of available multilingual documents enabled corpus-based MT techniques, consequently motivating ever more research effort into online MT systems. The successful results of these systems should mean that MT is established, in both areas of legitimate research and useful applications of technology. Nonetheless, MT is still heavily debated regarding its usefulness and efficacy in the real translation world.

As a technology, free online MT is also an important application in the issue of translation technology. It is considered as the most accessible form of MT, although it is not the most representative sample of MT performance and is mostly used for assimilation purposes, as pointed out by Gaspari (2007), to get the gist of what a foreign text says, but not so much for dissemination purposes where output of a higher linguistic quality is often required. Unfortunately not all students are aware of these two main uses of MT and of the fact that free online MT is not particularly suited for the latter because free online MT constitutes a 'black-box' system that cannot be customized according to the language pair, the type of text to be translated, its complexity and purpose. Even so many students, especially weak students or students with

Language in India www.languageinindia.com **ISSN 1930-2940 15:1 January 2015**
Mohammed Mohammed Al-Aqeeli and Dr. B. S. Jadhav   14
Potentials of Post-MT Editing as a Teaching Method to Promote Translation Quality & Time Spent

a low command of the target language, use free online MT output as a language resource for their FL written production assignments, mainly FL writing and translation into the FL.

Increasingly, post editing the output of MT is becoming the common use by many translators. Upon this, many studies and experiments have been conducted to measure and estimate post editing effort and time spent in comparison with the output of translation from scratch. These studies concluded that post editing sufficiently high-quality MT produce benefits for translation productivity, whereas post-editing poor machine translation can require more effort than translating from scratch. In addition, an increasingly common use for machine translation is producing texts to be post-edited by translators. While sufficiently high-quality MT has been shown to produce benefits for productivity, a well-known problem is that post-editing poor machine translation can require more effort than translating from scratch. Measuring and estimating post-editing effort is therefore a growing concern addressed by Confidence Estimation (CE) (Specia 2009b). Time spent on post-editing can be seen as the most visible and economically most important aspect of post-editing effort (Krings, 2001); however, post-editing effort can be defined and approached in different ways. Maarit Koponen (2012) points out the division of post-editing efforts as discussed by Krings (2001). She cites:

"Krings 2001 divides post-editing effort into three types: 1) Temporal, 2) Cognitive and 3) Technical. Temporal effort refers to post-editing time. Cognitive effort involves identifying the errors in the MT and the necessary steps to correct the output. Technical effort then consists of the keystrokes and cut-and-paste operations needed to produce the post-edited version after the errors have been detected and corrections planned." (2012: 181)

These different aspects of effort are not necessarily equal in various situations. In some cases, the errors may be easy to detect but involve several technical operations to be corrected. In other cases, parsing the sentence and detecting the errors may require considerable cognitive effort, although the actual technical operations required are quick and easy. According to Krings (2001), temporal effort is a combination of both cognitive and technical effort, with cognitive effort being the decisive factor. Assessing and reducing the cognitive effort involved in MT post-editing would therefore be important but the task is far from simple. Past experiments have involved cognitive approaches such as think-aloud protocols (Krings, 2001; O'Brien, 2005; Carl et al., 2011)

In her study, Chung-Ling Chih (2007) concluded that MT errors refer to the inappropriate translations at the lexical, pragmatic and syntactic levels in the MT output. She continued illustrating that, human editing is required to improve the quality since there is a lack of producing a satisfactory MT output by the MT systems and this process hereafter is known as post-MT editing. Post-MT editing, as defined by Juan C. Sager, is "the adaptation and revision of output of a machine translation system either to eliminate errors which impede comprehension or to make the output read like a natural- language text" (1994: 327). Melby defined this term as "the process of revising a translation after the draft translation has been completed" (1987: 146). In short, all mentioned researchers tend to employ human editing in developing MT raw output quality, so they propose to say that, post-MT editing can be mainly undertaken to improve the quality of the MT output for publication purposes.

B. Aims

Language in India www.languageinindia.com ISSN 1930-2940 15:1 January 2015
Mohammed Mohammed Al-Aqeeli and Dr. B. S. Jadhav   15
Potentials of Post-MT Editing as a Teaching Method to Promote Translation Quality & Time Spent

1-   Analyzing college students' translation production when they post edit ready machine translated texts or when they use MT to translate texts and comparing it with their translation from scratch for the same texts.

2-   Analyzing the raw MT output problems from many aspects of language structure for suggesting further potential development of the translation software applications.

C. Hypothesis

This paper presents a controlled experiment comparing post editing (hereafter "post-edit") to unaided human translation (hereafter "unaided") for two language pairs. We test two hypotheses: (1) post-editing reduces translation time, (2) post editing increases translation quality.

D. Procedures

Two phases were followed to conduct the current study; the first phase involved using in-depth interviews with the students. Interviews and tasks of filling the questionnaire relating to their reaction and reflections towards the use of machine translation were conducted after they had already finished the experiment of translating the texts using both tasks; (1) translation from scratch NRMTT and (2) translation with the use of MT WRMTT. For a critical analysis, researcher interviewed three teachers of translation about their familiarity and knowledge towards the use of machine translation and potentialities of using machine translation in their classrooms. They were asked to fill a small questionnaire aimed to measure the extent of familiarity and use of the machine translation by the tutors of translation.
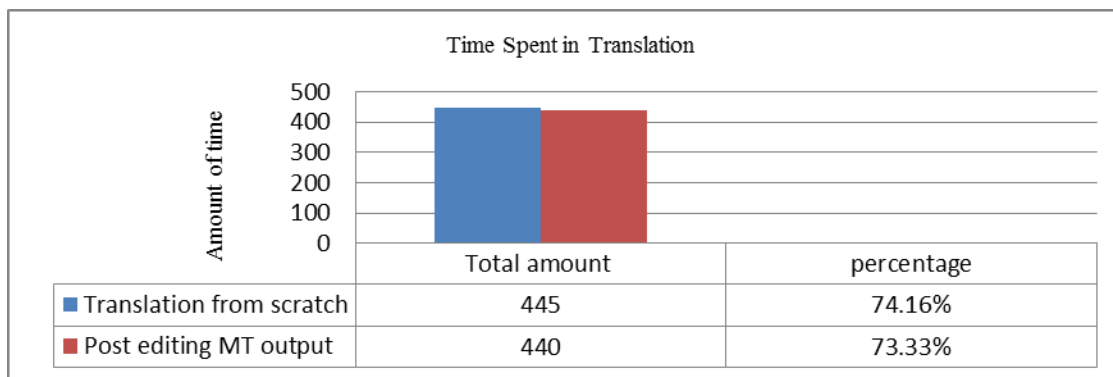
E. Sample

A group of 20 students were asked to translate a selected text from their scratch. Then they were asked to translate an equivalent text using ready machine translated texts in a try to rephrase and correct the errors produced by machine translation. This is what we actually know as post editing the outcome which aims at providing an eligible and appropriate translation. They were given an equal specific time in each task they had been asked to do.

F. Results and Discussion

Regarding the teachers of translation, researcher was contented with conducting some interviews and listening to them writing down many suggestions and modifications by them. In fact, an essential improvement had been made in many aspects of the questionnaires and tasks of translation experiments prepared to be given to the chosen sample of the study. Upon their suggestions, some new topics had been added and some had been omitted according to the suitability of the real field of application. Researcher also found that he has to change the way of introducing the experiments when he came to collect the data.

Figure 1 shows the total amount and percentage of the time spent by students to translate text in both tasks – translation from scratch and post editing the MT output.

Fig. 1 shows amount and percentage of the time spent in translation

Language in India www.languageinindia.com **ISSN 1930-2940 15:1 January 2015**
Mohammed Mohammed Al-Aqeeli and Dr. B. S. Jadhav   16
Potentials of Post-MT Editing as a Teaching Method to Promote Translation Quality & Time Spent

Time Spent in Translation	Total amount	percentage
■ Translation from scratch	445	74.16%
■ Post editing MT output	440	73.33%

With translation from scratch the total percentage is 74.16% whereas, the total percentage of time spent on translation using machine translation was 73.33%. This percentage clearly, indicates that the amount of time spent in translating texts has a slight statistical difference. From the point of view of the researcher, that slight difference occurred in favor of the task of translation with the use of machine translation.

With regard to error analysis, it is clearly shown in figure 2. Results showed 180 varied errors were detected in the translation of the students as they translated from scratch, whereas a number of 90 errors were detected in their translation when they post edited MT output, which clearly indicates the reduction of errors in the case of using machine translation tasks.

Figure 2 beneath shows the total account of errors committed by students in both translation tasks, from scratch and post editing MT output. It shows the frequencies of three types of errors grammatical, lexical and spelling, which were chosen for the current study.

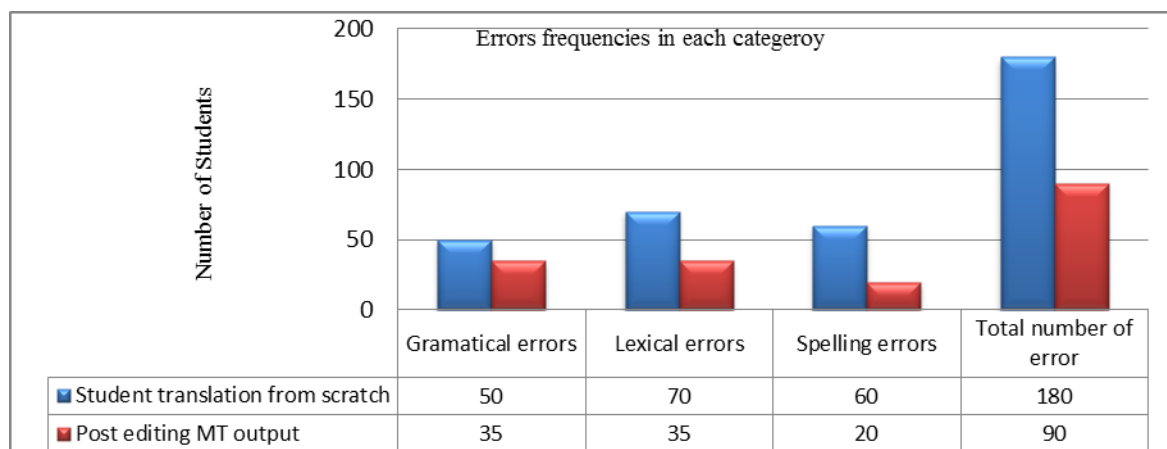| Errors frequencies in each categeroy	Gramatical errors	Lexical errors	Spelling errors	Total number of error
■ Student translation from scratch	50	70	60	180
■ Post editing MT output	35	35	20	90

Fig. 2 shows the errors in each category

It clearly shows that the occurrence of errors in the part of translation from scratch was double in size compared to post editing MT output, which gives a positive clue about MT post editing and pushes us to think seriously about those challenges and potentialities of utilizing MT

Language in India www.languageinindia.com **ISSN 1930-2940 15:1 January 2015**
Mohammed Mohammed Al-Aqeeli and Dr. B. S. Jadhav
Potentials of Post-MT Editing as a Teaching Method to Promote Translation Quality & Time Spent   17

in classrooms as a method to practice translation. Regarding the students' reflection toward the use of machine translation in translating texts, results showed that they were interested in using machine translation.

Conclusion

Results clarify the value of post-editing: it decreases time spent to some extent since a slight difference occurred in favor of the task of translation with the use of machine translation and, surprisingly, decreases errors and improves quality for each language pair. Results push us to think seriously about those challenges and potentialities of utilizing MT in classrooms as a method to practice translation. We find that MT suggestions prime translators but still lead to novel translations, suggesting new possibilities for re-training MT systems with human corrections.

===

===

References

[1]   Arnold, D. L. Balkan, S. Meijer, R.L Humphreys and L. Sadler. (1994). *Machine Translation: An Introductory Guide*. Oxford: NCC Blackwell.

[2]   Chung-Ling Shih. (2007). 'Teaching Translation of Text Types with MT Error Analysis and Post-MT Editing'. Translation Journal pp 3-4

[3]   Arnold, D. L. (1994). *Machine Translation: An Introductory Guide.* Oxford: NCC Blackwell

[4]   Gaspari, F. (2006). 'The added value of free online MT services: confidence boosters for linguistically-challenged Internet users, a case study for the language pair Italian-English'. Proc. 7th AMTA Conference, 46-55.

[5]   Gaspari, F. (2007). The Role of Online MT in Webpage Translation, PhD Thesis: The University of Manchester.

Language in India www.languageinindia.com **ISSN 1930-2940 15:1 January 2015**
Mohammed Mohammed Al-Aqeeli and Dr. B. S. Jadhav
Potentials of Post-MT Editing as a Teaching Method to Promote Translation Quality & Time Spent   18

[6]   Hinsley, F. H. and A. Strip. (1993). *Codebreakers: The Inside Story of Bletchley Park.* Oxford: Oxford University Press.

[7]   Hutchins, J. A. (1992). *An Introduction to Machine Translation.* London: Academic Press Limited.

[9]   Hutchins, W. J. (2002, April). The IAMT Certification Initiative and Defining Translation System Categories, Retrieved

[10]   Jeffery Davis, *Basil Hatim and Jeremy Munday (2007). Translation: An Advanced Resource Book.* "The Sign Language Translator and Interpreter" (SLTI): Volume 1, Number 1:157-162

[11]   K. Knight and I. Chander. (1994). 'Automated Post editing of Documents.' In Artificial Intelligence Conference, Seattle USA, 779–784

[12]   Krings, H. (2001). *Repairing Texts: Empirical Investigations of Machine Translation Post-Editing Processes.* Kent, Ohio: The Kent State University Press

[13]   Melby, A. K. (1987) 'On Human-machine Interaction in Translation'. In S. Nirenburg (Ed) *Machine Translation: Theoretical and Methodological Issues* (pp. 145-154). New York: Cambridge University Press.

[14]   Maarit Koponen (2012). 'Comparing human perceptions of post-editing effort with post-editing operations'. In Proceedings of the 7th Workshop on Statistical Machine Translation, pp 181–190 Montr´ et al (Eds). Canada, June 7-8, 2012. @ 2012 Association for Computational Linguistic operations

[15]   O'Brien, S. (2005). 'Methodologies for Measuring the Correlations between Post-Editing Effort and Machine Translatability'. In Machine Translation, 19, 37-58

[16]   Quah, C. K. (2006). *Translation and Technology.* London: Gunilla Underman & Margaret Rogers Great Britain

[17]   Somers, H. F. (2006). 'Detecting inappropriate use of free online machine translation by language students – a special case of plagiarism detection.' Proc. 11th EAMT Conference, 41-48

[19]   Somers, H. (3003b). 'Introduction'. In *Computer and Translation: A Translators Guide.* Amsterdam: John Benjamin 13-30

[20]   Sager, J. C. (1994). *Language Engineering and Translation: Consequence of Automation.* Amsterdam: John Benjamin

Language in India www.languageinindia.com **ISSN 1930-2940 15:1 January 2015**
Mohammed Mohammed Al-Aqeeli and Dr. B. S. Jadhav
Potentials of Post-MT Editing as a Teaching Method to Promote Translation Quality & Time Spent   19

[21]   Yehoshua Bar-Hillel, (1959). *Report on the State of Machine Translation in the United States and Great Britain.* Jerusalem: Israel, Hebrew University

==

Mohammed Mohammed Al-Aqeeli
PhD Research Scholar
School of Language, Literature & Culture Studies
SRTM University
Nanded 431606
Maharashtra
India
alakili7@yahoo.com

Dr. B. S. Jadhav
Associate Professor
School of Language, Literature and Culture Studies
Director of International Students' Center
SRTM University
Nanded 431606
Maharashtra
India
bhagwansjadhav@gmail.com

Language in India www.languageinindia.com **ISSN 1930-2940 15:1 January 2015**
Mohammed Mohammed Al-Aqeeli and Dr. B. S. Jadhav
Potentials of Post-MT Editing as a Teaching Method to Promote Translation Quality & Time Spent   20