

USING POST-EDITING IN TRANSLATION AND LSP COURSES

Natalia Udina

Assoc. Professor, RUDN University, Russia, udinann@gmail.com

Abstract

Teaching translation is getting important not only in language studies, but in professional development. Specialists are getting more involved in translation of scientific and specialized literature and often turn to machine translation. The latest development of machine translation improved the quality of translation nevertheless human involvement into editing is required. Machine translation skills include post-editing skills, which have to be developed

Post-editing (PE) of the texts, produced by Machine translation (MT) systems, has its specifics and perspectives in learning language for specific purposes and translation. Machine translation post-editing differs from human translation editing and becomes a new area of studies. As there is no model of teaching language and translation based on post-editing, the studies of post-editing impact on learning and teaching are necessary. The paper outlines the MT models, specifics of machine translation and post-editing and focuses on methodological and practical aspects of using post-editing in teaching. The paper deals with pre-editing and post-editing assignments for students of translation course. The analysis of post-editing monolingual and bilingual legal texts made by students revealed some problems they have while dealing with official language. Error correction and analysis and their classification are helpful in comprehension of the effectiveness of machine translation and post-editing. Students get new experience in identifying the main problem areas of translation and manage their knowledge, relevant information and skills to edit the translated text. The paper is aimed at considering the use of machine translation post-editing in teaching and learning translation and LSP.

Keywords: machine translation, post-editing, translation for specific purposes, skills development

1 INTRODUCTION

The field of machine translation includes special computer programmes for translation of natural languages, computer-based tools for human translators – electronic dictionary, terminology management systems, various translation databases etc. Machine translation development is intended to meet the growing demand of translation in various spheres.

Various models at early periods of machine translation (MT) worked out in the world such as 'direct translation' model, 'interlingua' model, the 'transfer' approach, 'meaning-text' model (developed in Russia) constituted the base of machine translation development. Rule-based approach and statistics based approach dominated in later programmes. The development of computers greatly influenced the machine translation. Corpus-based methods were introduced with further development of statistical and memory

based approaches. Statistical methods used no linguistic rules, the main operations were to align words and phrases of parallel texts and to calculate the probabilities of their correspondence. Memory based approach is based on the selecting phrases from the databank of parallel texts using either statistical, rule-based or semantic methods. The use of statistical methods improved greatly the quality of translation. The latest development of computer translation programmes resulted in introduction of Neuro Machine Translation (NMT) (Hutchins, 2003).

Progress in machine translation quality has been increased considerably, but nevertheless still requires human interaction. Post-editing (PE) is a new field of research, though it has been a long-time practice. The research projects on PE are mainly carried out in the area of new technologies and tools development (Koponen, 2016). But such issues as the specifics of pre-editing and post-editing strategies, the role of post-editor and the work status whether it is a translation or editing, cognitive aspects of PE require in-depth study.

Of special interest is the use of post-editing as learning tool for LSP and translator training.

2 THEORETICAL CONSIDERATION OF USING POST-EDITING IN LEARNING LANGUAGE AND TRANSLATION

Machine translation holds a distinct position in the field as there is a pressing need for reducing time and cost of translation that is why the quality of MT is in the focus. Two processes can be used to improve MT output: pre-editing and post-editing. Pre-editing is the text processing before translation such as grammar and spelling correction and eliminating vagueness of the text. Post-editing has two versions: light post-editing and full post-editing. Light or rapid post-editing is used in case the general information is required, full post-editing requires more time and effort on the part of the post-editor and the result is the text stylistically and linguistically appropriate. First studies of post-editing were initiated by Association for Machine translation in the Americas (AMTA) and the European Association for Machine Translation with the purpose of working out guidelines and training in this sphere.

Some researchers maintain the idea that the courses on teaching post-editing skills are also effective for training translation and linguistic skills. They assume that translation courses are structured differently at various universities, so the stage at which the post-editing course has to be introduced can differ, but it is important for students to acquire some practice in machine translation (O'Brien et al, 2014).

The studies of MT post-editing use in language class include computer-aided error analysis of error patterns, found in translation and post-editing to get insight into problems the students might have while using PE and the effectiveness of using PE (Niño, 2008).

Sun Dongyun (2017) claims that researches on using PE in language learning are few and further studies in this field are required. He outlines that translation and language teachers try to use post-editing in various programs on translation, Computer-Assisted Language Learning (CALL), and English writing course. He asserts that along with effective use of PE in language learning and teaching, there is no PE based teaching model and no consensus on how to teach PE in translation and language courses. Though there are some controversial issues, he states that MT and PE influence on translation teaching should not be overlooked.

The cognitive translation studies, relatively new interdisciplinary field of research, deal with various cognitive process of translation. It was reported that the cognitive processes of text comprehension while post-editing and translating differ. It was also revealed that human translation is less linear process than post-editing (Krings, Koby 2001). Researches into post-editing efforts, including temporal, cognitive and technical, give insight into cognitive process of post-editing (Daems, J., et al., 2015). These and other cognitive aspects such as attention and memory in post-editing are important and have to be considered.

3 PRE-EDITING AND POST-EDITING ASSIGNMENTS AND RESULTS ANALYSIS

A post-editing seems to have aroused considerable interest recently, but in fact, it has been of great concern since early time of MT development. There is a need of general guidelines for users and language service providers. Guidelines worked out by Translation Automation User Society (TAUS, 2010) outline two quality levels of PE. The first level is defined as comprehensible and accurate. At this level the task of the post-editor is to see whether the text is semantically correct and no information added or missed. The second level is of 'publishable' quality, that is comprehensible, accurate, grammatically and stylistically relevant (Cononen, 2016).

Post-editing studies usually involve post-editing of bilingual texts, apart from this, the monolingual post-editing was analyzed to see whether monolingual readers are able to understand the meaning of machine

translated text (Koehn, 2010).

The introduction of post-editing in teaching foreign languages and translation and PE impact on developing language and translation skills have been discussed by researchers and educators. The availability of MT systems make the use of MT post-editing possible in various environment.

The following PE assignment was given to the three groups of students, studying at the translation programme in their second year (1 group) and third year (2 groups), 20 students in all. They received monolingual and bilingual texts for post-editing. The assigned texts were translated by Google Translate system, which is one of the most widely used machine translation services, which completes billions of translation requests, involving over 100 languages each year. Google Translate can be used to translate websites, written texts, documents, and handwritten texts on the phone screens, it also can translate speech spoken into a device (Wikipedia).

The students were first assigned to read monolingual text without seeing the original text and identify the most obvious mistakes of any kind. This assignment was supposed to focus the students' attention on grammar and meaning of the text in general, to have some idea about parts of the text where correction is required. Assigned texts were translated from English into Russian, the themes of the texts related to migration legislation and migration issues and contained 60-80 words each,

While post-editing monolingual text (translated in Russian), the following types of mistakes were identified by students of the 2nd year:

The system missed the word and left it untranslated in the text

Wrong case used (endings of the words)

Synonyms or words close in meaning, translated by the same word, used twice

Wrong grammatical forms of the verb (perfective –imperfective aspect)

Name of the Agency not translated

Word order

Wrong choice of the word

the word used twice in the sentence, one after another

Change of the part of the speech, not fitting in the given context

Sentence translated not clearly

The students of the 3rd year were given the same texts and the following errors in MT were identified by them:

Wrong part of the speech used

Sentence structure

Wrong preposition

Sentence identified as wrong translated

Subject-verb plural coordination

After students looked through and post-edited monolingual texts, they were assigned post-editing of the bilingual parallel texts (the same text but in English- Russian). Two students of the 2nd year and 1 student of the 3rd year haven't found any further errors in post-editing bilingual text. Other students did more editing, but many mistakes remained uncorrected. It might be because they haven't mastered the language of legislation, the topic of migration was relatively new for them.

This type of work was new and more complicated for the 2nd year students, as they haven't developed translation skills, so post-editing was a difficult task for them. The students of the 3rd year did more corrections and after post-editing the quality of text translation improved. But there were also miscorrections and unnecessary word substitutions.

Post-editing assignments could be effective tools in developing translation and linguistic skills of students. In using monolingual post-editing the students got some idea of the text content and translation errors, they could think of some strategies of post-editing.

There is a saying ‘we are taught by our mistakes’ which is up to the point in the case of post-editing. Students have to identify mistakes and analyze them, thus practicing grammar rules, vocabulary, words coordination and collocations, and language use in the text in general. This type of exercise helps stimulate cognitive process and improve learning.

Monolingual post-editing can be effective at the beginning of teaching post-editing skills, as it gives students possibility to focus on the text and not split their attention between two texts, which is important cognitive aspect. Students got general ideas of the text which help them to start thinking about possible strategies of post-editing.

The following three factors have to be considered while assigning post-editing: time factor, topic relevance and the text/passage length.

Time provided for monolingual text should not be very short. Though it is kind of text scanning, students have to be allowed more time to get general comprehension and think of editing strategies.

The level of complexity of the text is important. For learning purpose the text should not be too difficult, the complexity level could be increased in the process of training.

While post-editing bilingual text, the sentences of parallel texts have to be aligned, to facilitate the mental processes. The students may have an access to dictionaries, the databases to get information they require while post-editing.

4 CONCLUSION

In the world of globalization, new technologies and mobility, the role of translation becomes manifold. The new technologies influenced MT systems development; many companies offer on-line translation services, though the quality of translation required human involvement. Post-editing is widely used in language industry and has been introduced into foreign language and translation courses.

Post-editing seems to be an effective tool in teaching students language and translating skills as they are taught to analyze the translation, to pay attention to grammar and structure of the sentence, while making the decision on editing strategies, students manage their knowledge and skills and gain useful experience.

Pre-editing text often means making text more comprehensible, this assignment can be very useful from the part of grammar and vocabulary learning. The students could be trained in making the text more comprehensible, using simple structures and avoiding complicated terminology. This aspect of language use is particularly pressing in the sphere of law and administration as formal language is claimed to be difficult and needs clarity.

The factor of novelty and practical usefulness make post-editing assignments more motivating, which contributes to language acquisition and translation skills development.

REFERENCE LIST

Daems, J., Vandepitte, S., Hartsuiker, R., Lieve, M. (2015). The Impact of machine Translation Error Types on Post-Editing Effort Indicators. In O'Brien Sh. & Simard M., (Eds) Proceedings of 4th Workshop on Post-Editing technology and practice.p.31- 45. https://amtaweb.org/wp-content/uploads/2015/10/MTSummitXV_WPTP4Proceedings.pdf .

Hutchins, J.W. (2003) Machine translation: half a century of research and use. <https://pdfs.semanticscholar.org/d8d9/c068dd2e0dd438e5d5d01356a24bba977dd4.pdf>

Koehn, Philipp (2010). “Enabling monolingual translators: Post-editing vs. options.” Ron Kaplan et al. (eds) (2010) Human Language Technologies: The 2010 Annual Conference of the North American Chapter of the Association for Computational Linguistics. Stroudsburg PA: Association for Computational Linguistics, 537–545.

Koponen, M.(2016). Is Machine Translation post-editing worth the effort? A survey of research into post-editing and effort. The Journal of Specialised Translation. Issue 25 https://www.jostrans.org/issue25/art_koponen.pdf

Krings, H. P. and Koby G. S. (eds) (2001), Repairing Texts: Empirical Investigations of Machine-Translation

Post-Editing Processes, Kent State University Press, Kent, Ohio.

- Niño, A. (2008). Evaluating the use of machine translation post-editing in the foreign language class. *Computer Assisted Language Learning*. Volume 21, issue 1. Taylor& Francis Online.
<https://eric.ed.gov/?id=EJ786006>
- O'Brien, Sh., Winther Balling, L., Carl, M., Simard M., and Specia L. (Eds) (2014). *Post-editing of Machine Translation: Processes and Applications*. Newcastle: Cambridge Scholars Publishing.
- TAUS (2010). *Machine Translation Post-editing Guidelines*. <https://www.taus.net/think-tank/best-practices/postedit-best-practices/machinetranslation-post-editing-guidelines>
- Sun Dongyun (2017). Application of Post-Editing in Foreign Language Teaching: Problems and Challenges. *Canadian Social Science*. Vol.13 №.7. p1-5.
<http://www.cscanada.net/index.php/css/article/view/9698/10391>
- Google Translate. Wikipedia. Accessed September 2018, https://en.wikipedia.org/wiki/Google_Translate