POST EDITING: SUGGESTIONS FOR TRANSLATION COURSES
SUB-CONTENT

Alvin Taufik*

English Language and Culture Department, Bunda Mulia University

Received on 25 March 2019 / Approved on 8 April 2019

ABSTRACT

The practice of post-editing has become so common these days with the introduction of Machine Translation (MT). Because of that, the author believes post-editing should be included in all types of translation activities. Subsequently, in the teaching of translation, post editing should be considered as one of the course to be taught. In translation teaching, however, especially in Indonesia, post-editing has not yet been discussed specifically. Most post-editing activities in the classroom in Indonesia is not taught specifically as one of the courses as a part of a minor. Rather, they are integrated as a sub-course of a particular course, such as literary translation. In this research, I would like to propose several inputs which can be considered when teaching post-editing as a part of a bigger course. This research is conducted by reflecting on the actual process of post-editing conducted by the researcher when translating a text. It is expected that the result of the research is to be used by translation trainers, teachers, and practitioners in training their students.

Keywords: Post-editing, translation course, sub-course

INTRODUCTION

With a steadily increasing demand of translators, (in contrast to the popular belief that translators are losing their jobs due to the existence of Machine Translation (MT)) faster work process of translating is required to fill in the demands. That is where the MT comes into play. To be clear, in this research, what is meant by MT are those which still requires the interference of human (Human assisted (aided) Machine Translation or HAMT); one example of which is Google Translate. For this reason, a clear understanding of post-editing process and activities is necessary. However, post-editing activities in actual translation or in translation training has always been considered as an integral and ‘natural’ process; not many research has been done to investigate what actually happens in that process. For that reason, I conduct this research. This research is significant because it will inform trainers or students alike to ponder on the process of post-editing, and take appropriate actions to deal
with the problems found. To do that, the researcher needs to answer the following questions:

a) What is it that we do when we are post-editing a text translated by a machine translation?
b) What should be taught in a post editing class/subcourse?

LITERATURE REVIEW

2.1. Post Editing as a Standalone Activity

To start off, we need to see different phases of translation. Nida (1969) coined the different phases of translation into a ‘process of translating’ which includes analysis, transfer, and restructure. From Nida’s conception, is it fair to say that post-editing is included as a part of a translation activity? Some might argue that it (post-editing) can and should be included as a part of translation activity given the fact that it is the same as the ‘restructure’ phase in Nida’s conception. However, it should be noted that, at the time Nida created this conception, all translation activity was still conducted by Human as the translator. Post-editing in this research refers to activity that the translator/translator trainee does after the product of translation has been given by the machine – or PEMT (Post-Editing Machine Translation). In this research, Google Translate is the machine, and henceforth, it will be mentioned by its abbreviation, GT.

Some researchers in translation study seem to have the same opinion, Kring and Koby (2001) states that Source Text (ST) comprehension (perhaps the ‘analysis’ phase in Nida’s conception) and post editing is indeed different. O’Brien (2014) has also conducted a research on the similarity/dissimilarity of translation or translating and post editing. The result states that the two are actually different from a practical point of view. She added that while translation involve a transfer of ONE text from one language to another, post-editing involves two texts; the ST and the raw MT/GT output. Therefore, it is safe to say that post-editing is a standalone activity. As such, it should have its own approaches.

2.2. Post-Editing Approaches

There have been many research on translation post-editing. One prominent research was by Krings (2001, quoted from Koponen, 2016). In his book, he introduced five post editing approaches. The first relates to the modification of the ST to be translated into a specific domain and types. This is to have the text features to be defined precisely by the machine. In my opinion, in this approach, the editor needs to alter some vocabularies or terminologies in the ST so that they confirm to the often-used vocabularies or terminologies which are specific to one domain so that the MT can categorize the text into a specific domain and, subsequently, be translated in the norms of that particular domain.

The second approach still involves the ST; the editor should conduct pre-edit the text in accordance to a specific simplification rule. Kring’s third approach mentions the use of interactive MT system. Unfortunately, the author has no adequate explanation on this. Koponen only mentions that this is related to human translator involvement in solving problems during the machine translation.

The fourth approach relates to the production of necessary general information which does not enforce high quality text. To my understanding, this approach emphasize the understanding of a text in a whole to be translated based on the general information in the content of the text. Finally, the fifth approach, and the one which becomes the main focus of this research, is the production of raw translation to be edited by human. This is the practice I have been using a lot in my translation activities. In this case, the MT produce raw translation which sometimes needs improvement. The nature of the improvement and the process in improving the raw data is the focus of this research.

2.3. Post-Editing Process

Before moving further to the process of Post-Editing, as mentioned in Koponen’s
dissertation (2016) on post editing, we need to decide whether this post editing process involves the checking of TT acceptability against the ST, or, as often said, is monolingual, i.e. the process does not involve the checking of ST. In my previous experiences, the checking of ST was necessary because sometimes the TT was so obscure that I need to check what is actually meant in the ST. For that reason, in this research, the term ‘Post-Editing’ refers to the checking of the TT against the ST.

Another aspect which needs consideration is the editor; are they professional translator or are they trained to be editor. In this research, the aim is to provide feedback to the students in regard with the Post-Editing activities. Thus, it is safe to say that the editor/s are not professional translator. This setting of the participant will further impact the lessons to be taught in the sub-course.

Of yet other aspects related to Post-Editing are related to the amount of Post-Editing conducted on the MT results. On this, we come to know the terms coined by Löffler-Laurian, (1986), as quoted in Koponen, of partial and complete Post-Editing, as well as the leveling based on the Post-Editing activities, namely light, medium, and full by O’Curran (2014). In this research, both or all three activities in Post-Editing could be adopted. Therefore, all process mentioned in those distinctions are probably going to be the basis for the Post-Editing process involve in this research’s practice.

In relations to the process itself, in general, it can be seen from three different distinctions of planning, text generation, and revision (Englund Dimitrova, 2005). This was originally meant to differentiate the three different cognitive processes involved in making a translation. More on this process and its implementation in the reflective practice conducted in this research will be mentioned in the latter part of this research.

2.4. Translation Training

To successfully convey the feedback to the students, I also pay attention to how post-editing is taught in educational situations. The following are principles pointed out by Schäffner (2016) in conducting translation training.

a. constructivist approaches to translator training: teaching how to learn
b. student-centered approach to translator training
c. competence-oriented teaching methods
d. professionally-oriented learning models
e. preparing students for working in a real-life setting
f. collaborative learning, team-work
g. task-based learning / teaching
h. authentic assignments
i. pedagogical tools (corpora, portfolios, TAPs, Translog, screen recordings, etc – research-led)

RESEARCH METHODOLOGY

As mentioned previously, this research is a reflective practice. Lambert mentions the following on reflective practice and its relation to translation training:

“Reflective practice, for me, represents one of the best working methods available for the practicing translator as well as providing a rare opportunity to integrate theory and practice, which remains a key issue in the discipline. Reflecting on your own work as you translate forces you to question and to justify your translatorial decisions in relation to an end goal (a translation brief in this case) and this in turn allows a deeper understanding of your own working process and a better grasp of your areas of strength and weakness.” (Lambert, 2013)

However, this does not mind that it is a baseless practice, in the sense that practitioner can decide and justify their reflective practice without theoretical bases. Lambert further mentions that reflective practice should be based on an underlying principle in translation study. For that reason, in conducting reflective practice in this research, I use the principles of
post-editing, particularly, post-editing principles which are based on MT outputs.

As stated in the previous section, Englund Dimitrova (2005) has pointed out the three main cognitive process of translation. Although they are initially meant to differentiate three cognitive processes of translation, Koponen (2016) has already mentioned that they are also applicable to PEMT. The processes are influenced by task environment, which includes the translation, topic, audience, the text produced so far, long term memory of the writer, and the writing plan. Included in those aspects are the ST and TT knowledge. In the context of PEMT, MT is added to task environment, and ‘more general bilingual PE scenario, SL and TL knowledge, and perhaps some specific knowledge of how PE is carried out’ is included in the long term memory of the writer (Koponen, 2016). These additions might change the overall process of Post-Editing.

With the addition of MT, which affects task environment and memory, according to Krings (2001), there are two processes affected, they are the reading pattern which were broken up by the defect in the MT outputs, and the act of searching equivalency, as interfered with the suggestions given by the MT. The breaking of reading patterns as stated by Krings might have two purposes; one is to check the ST in the case the NT is unclear, or to check general understanding of ST to make the MT as a part of his/her overall translation. Often, the defect in the MT might lower the reading level of the ST into smaller segments such as clauses or even words.

The previously mentioned process eventually limit the segments of the MT to be memorized by the Post-Editor. Thus, there is a need to read the result of the MT. Often times, there is another process which must be conducted, namely to ensure the coherence and the cohesion of the Post-Edited MT.

In this research, the method that I use to provide feedback is by identifying the way Post-Editor check the ST when dealing with defect MT, and whether I am using more of the checking the ST or understanding the overall idea of ST to make MT a part of my editing, and to what extent and frequency a check on the cohesion and coherence of the ST segments needs to be done.

Moving on to the actual cognitive process involved in both translation or post-editing, the planning part is also influenced by the addition of MT. With the MT, the planning is in the hands of the ST writer; MT suggest inputs for editors to accept or alter according to the needs. In this research, I aim to identify whether the planning stage has been completely reduced by the MT, or there are still some things that can be done in that stage. In terms of text generation, several studies have shown that PEMT’s results are just as good as the human translator’s result, with weaknesses in stylistic area and strength in terms of accuracy (Carl et al., 2011). It is interesting to see if these problems of stylistic can be moderated in this research, and what can be done to achieve successful moderations. Such is what I aim to do in this research.

In the revising, it is said that post-editors of MT would alter the result more freely (Krings, 2001; Koby, 2001). In a sense, Post-Editor tends to rewrite the defect segments of the MT productions as a whole. According to the experts, this is due to the fact that the Post-Editor was not worried that they might hurt other people’s feeling as is the case with revisers of Human Translator. In this research, in relation to revising, it is my intention to see what else is done when Post Editor is faced with text which requires extensive altering.

THE REFLECTIVE PRACTICE

4.1. Data Selection and Research Segmentation

The data selected based on two different classifications. The first data contains a very long word to be translated by the PEMT. Having found out the result of this, I am using the long word in its paragraph. The data used is taken from the Annual Report of PT Jamkrindo of 2018. No Non-Disclosure Agreement was breached in this research since

*Author(s) Correspondence:
E-mail: ataufik@bundamulia.ac.id
all annual report of Perum Jamkrindo is published online and in specific publications for all stakeholders to see. Such stakeholders include the general public.

I chose this data because it is an actual work that I am currently working on at the moment. Thus, it reflects the actual practice and usefulness of post-editing. For the segmentation of the research, it is based on the cognitive processing and their influencing factors as pointed out by Englund Dimitrova (2005). Another reason for the choosing of this data is to see whether there are different process involved in MT Generation of the texts, and in its post-editing process.

4.2. The Research Process

(1) On a sentence

In this section, I will explain what I do in my actual usage of MT to translate, and later, post-edit the results. As I have mentioned before, I start this by inputting a very long word. These are the kind of words which will generate the most defect. This is the word in question:

ST:
Sejalan dengan perkembangan dunia usaha, prinsip GCG turut mengalami perkembangan yang mencakup sistem pengawasan dan pengendalian yang mendukung etika kerja dan pengambilan keputusan yang bertanggung jawab, mendukung integritas dalam pelaporan keuangan, pengelolaan risiko yang layak, serta hubungan pemangku kepentingan dan Perusahaan yang berlandaskan etika.

At first, what I always do is to open three web windows; the GT itself, the google websites, which I always open to refer to the actual use of the word/sentence in the real world (this can actually be akin to Corpus-based referencing, but since I have not got enough background on the material, I am actually using web-referencing), and the second GT to check individual segments of the text. This is the result of the MT:

TT:
In line with the development of the business world, GCG principles support the development that regulates a system of supervision and support that supports responsible work, supports coordination in financial reporting, appropriate financial planning, and supports ethical and corporate interests.

From the MT result, at first, I check for the coherence of the result. My reading pattern at first was stopped by the repetition of the word ‘support’ as highlighted in the MT above. Of the second reading, I notice that the word ‘support’ is used quite frequently; my initial thought was these ‘supports’ can be made non-redundant. So, after finding out the problem, I continue by, first, reading from the ST. From the ST reading, I found out that the clause ‘prinsip GCG turut mengalami perkembangan yang mencakup sistem pengawasan dan pengendalian...’ does not have the same meaning as ‘GCG principles support the development that regulates a system of supervision and support...’. So what I did was, I rewrite the clause altogether into:

PEMT:
The GCG principles also experiences a development which covers a system of supervision and control...

The decision to use the the tenses was also a part of my process of post-editing. The way I decided the tense to use is based on my background knowledge of the text being translated. I took the decision to use present tense because this is related to the annual implementation; it is always done. To continue, I need to change the two ‘supports’ included in the MT. However, I no longer need to do that since the rewriting process has already dealt with the repetition. So, what I did was, I checked the rest of the sentence while still considering the ‘soundness’ of the translation. Once I continue with the
reading of the ST, I found out that there are more defects in the translation. The other defect which I found relate to inappropriate use of Noun Modifier as seen in the sentence ‘mendukung etika kerja dan pengambilan keputusan yang bertanggung jawab...’ which is translated into ‘supports responsible work...’. From this, I realized that I need to put in the missing (untranslated) words, and put the appropriate Noun Modifier into the proper Noun. The result of my modifications can be seen below:

PEMT: ‘...which supports work ethics and responsible decision making...’

Continuing on the reading, I was put on a dilemma of repeating the words ‘support’ in line 3 and 4. I was countered with the choices of using the word ‘support’ to clarify the clause ‘...co-ordination in financial reporting...’ and ‘...ethical and corporate interests...’. I finally took the decisions of not repeating the words ‘support’ since the clauses are in the form of Noun Clauses. In addition to that, I also made correction to the word ‘co-ordination’ with ‘integration’ as well as adding the word ‘stakeholder’ which is missing in the TT and altering the word ‘interest’ with ‘relationship’. Thus, the result of the PEMT is as follow:

PEMT: ‘...which supports work ethics and responsible decision making, integration of financial reporting, appropriate financial planning, and ethical and corporate relationship’

The final result of the PEMT will be as seen below:

PEMT: ‘The GCG principles also experiences a development which covers a system of supervision and control which supports work ethics and responsible decision making, integration of financial reporting, appropriate financial planning, and ethical and corporate relationship’.

To conclude, the followings are the process of Post-Editing conducted on a sentence level.

a) Pre-processing Stage:
In the pre-processing stage, as stated earlier, what I intend to do besides what have already been stated in literatures are to identify the way Post-Editor check the ST when dealing with defect MT, and whether I am using more of the checking the ST or understanding the overall idea of ST to make MT a part of my editing, and to what extent and frequency a check on the cohesion and coherence of the ST segments needs to be done.

Based on the process it can be said that when confronted with defect MT, I tend to check the ST more than to understand the whole idea of the ST to make MT part of my translation. The way I check the ST on the sentence level translation segments of this research is to check the meaning of each individual words contained in the ST and compared them to MT/TT. As for the checking of the cohesion and coherence, I did not do that until later stages.

b) The Planning
In the planning stage, I aim to identify whether the translation planning has been completely reduced by the MT, or there are still some things that can be done in that stage. It turns out that there are still planning to be done before the text generation. From this research, it was found that coherence planning was conducted after the reading of the results. Some coherence planning which I did was to avoid repetition and reduce redundancy.
c) Text Generation
In the text generation, in relation to the coherence section of the planning, I was occupied with finding the word which can avoid repetition and determining whether there is a need to change the TT to support the reduction of redundancy. Also in the text generation, it should be included that tense-usage decisions was also conducted. In the stage where I need to determine the tense which must be used, I refer to the overall context of the text being translated. Other grammatical adjustment was also conducted in this stage.

In relation to stylistic adjustment, in my case, it was found that the MT actually has better stylistic results than the PEMT, however on the accuracy and consistency, I tend to believe that the PEMT is better. So, in this case I could not achieve the aim I have set previously since the results contradict with existing literatures.

d) Revising
What is interesting in this stage is that in my case it is not of separate stage. I was also doing stylistic and soundness adjustment of the text when I was generating the PEMT. Also in the revising stage, I rewrite most of the clauses to adjust to the meaning contained in the ST. Although it was included in the previous literatures on the revising stage, I could not help but notice that it was actually done mostly in the text generation stage. In addition, as I was rewriting the TT, I often had to add the missing words and change the incorrectly used words.

e) Conclusion
It can be concluded that, in the case of long sentence to be translated, it is very difficult to put the process accurately into separate stages; instead they are overlapping and contributes each other. Planning stage, in my case was conducted after the reading of the TT; this is unlike translation process in which planning is actually done in the writing stage. The text generation was conducted mostly by frequently checking the ST meaning, overall context of the whole text, and cohesion of the TT. Finally, in this research, the revising stage of rewriting the clauses of the sentence was conducted, actually, per segments in the text generation stage. This confirms my initial remark about the overlapping activities of text generating and revising.

(2) On a paragraph
Using the same data, I now include the paragraph as a part of the text to be generated by the MT. This is the text:

ST:
Sejalan dengan perkembangan dunia usaha, prinsip GCG turut mengalami perkembangan yang mencakup sistem pengawasan dan pengendalian yang mendukung etika kerja dan pengambilan keputusan yang bertanggung jawab, mendukung integritas dalam pelaporan keuangan, pengelolaan risiko yang layak, serta hubungan pemangku kepentingan dan Perusahaan yang berlandaskan etika.

Di lingkungan Perusahaan, evaluasi secara berkala dilakukan atas penerapan GCG untuk memastikan terwujudnya prinsip GCG secara komprehensif.

As with the sentence level, I open three web windows; GT, second GT, and Google. As I input the ST, the following is the result:

TT:
In line with the development of the business world, GCG principles have also experienced developments that include a
system of supervision and control that supports work ethics and responsible decision making, supports integrity in financial reporting, appropriate risk management, and ethical and stakeholder relations. Within the Company, regular evaluations are carried out on the implementation of GCG to ensure the realization of the principles of GCG comprehensively.

At a glance, we can already see the difference in the result. At first, I checked the correctness of the previously rendered sentence. In overall, this time, the choice of words or accuracy of the MT has improved. I have no idea if the second sentence influences the result of the MT. However, I also still see the repetition of the word ‘support’ in this MT. So my first plan is to, again, reduce the redundancy of that word. Continuing with the reading, I also still see some words are missing in the initial MT/TT. The missing word is ‘company’ which is supposed to be a translation of the word ‘perusahaan’. I also thought that the translation in that part can be modified quite moderately. Thus, for the first sentence, I post-edit it into a text which can be seen below:

PEMT:
In line with the development of the business world, GCG principles have also experienced developments that include a system of supervision and control that supports work ethics and responsible decision making, integrity in financial reporting, appropriate risk management, and ethical relations between the stakeholders and the Company.

Continuing on with the second sentence, I first checked whether it is understood.

Immediately, the word ‘comprehensively, did not sit right with me. So, I plan to either change the position and the word class of the word ‘comprehensively’ or rewrite that part only. To do that, I need to see whether what I was thinking was actually reflected by the ST. So, I now checked the ST to see the general idea, which then lead into the decisions that I am going to take to deal with the ‘comprehensively’ part.

After reading the ST, I decided to choose my first plan; change the position and word class. The result of that decision can be seen below:

PEMT:
Within the Company, regular evaluations are carried out on the implementation of GCG to ensure the comprehensive realization of the principles of GCG.

Two factors decided my decision. First, I was not sure that the word comprehensively was supposed to be used at the back of the sentence like so, and; Second, I was so accustomed of using the word ‘comprehensive’ before the Noun. It can be said that my post editing of the second sentence was based on preferences. However, That is not the aim of this research. This research is aimed at giving feedback and suggestion to translator trainee when conducting post-editing, with added considerations that the trainee are not professionals as have been stated before. So, what I can suggest here is, when dealing with such problem, students can always check the Corpus, or Google to find out whether such use of the word ‘comprehensively’ is actually acceptable in the real situation. With that, the following is the PEMT complete results:

PEMT:
In line with the development of the business world, GCG principles have also
experienced developments that include a system of supervision and control that supports work ethics and responsible decision making, integrity in financial reporting, appropriate risk management, and ethical relations between the stakeholders and the Company. Within the Company, regular evaluations are carried out on the implementation of GCG to ensure the comprehensive realization of the principles of GCG.

To conclude, the following are the process of post-editing on a paragraph level (above sentence level).

a) Pre-processing Stage:
   In the pre-processing stage, I still see the ST only when I need to, sometimes even just to check whether the already ‘accurate’ translation actually relates strongly to the TT. This is different to the process that I did with sentence level TT.

b) The Planning
   As can be seen in the explanation of the process, I was still planning on reducing the redundancy of the TT, adding missing words, and light modification.

c) Text Generation
   In the text generation, still related to the coherence section of the planning, I erase the redundant use of the word ‘support’ from the text, add a missing word and restructure the clause based on that additional word, and conduct some light modification based on stylistic preference.

   In relation to stylistic adjustment, it can be said that the adjustment was made based on experience and not on accuracy. One way to avoid this is by asking for third reference (Google, or other relevant search engine, or Corpus Analysis software).

d) Revising
   Finally, in the revising, as has been done previously, the process is still interwoven with the text generation stage. One interesting aspect which can be considered is the amount of revising which must be done for a paragraph-level MT; It is actually a lot less than in the sentence level, at least in this research. It remains to be seen whether that can be recurring in other part of the text. Also in the revising stage, I still do rewriting activity, only now is based on preference.

e) Conclusion
   Post-editing a paragraph level text has been proven to be less complicated than the sentence-level one. One thing that can be taken from this is that when including the text to be generated by the MT, it is better to include larger portion of the text. What is causing the phenomenon to happen remains to be known and identified.

CONCLUSION AND SUGGESTION

The conclusion of this research is going to be presented in three parts; the pre-processing and planning stage, text generating stage, and revising stage.

In the pre-processing and planning stage, what I aim to do was to identify the way Post-Editor check the ST when dealing with defect MT, and whether Post-Editor use more of the checking the ST or understand the overall idea of ST to make MT a part of my editing, and to what extent and frequency a check on the cohesion and coherence of the ST segments needs to be done.

To answer that question, from this research, I tend to check the ST when dealing with defect TT. I check the ST not to get the general idea of the ST but rather to check for the accuracy of the TT. In terms of frequency, The frequency of ST checking is actually
higher when I included smaller segments of the text to be generated by MT.

In the planning stage, it turns out that it was not reduced by the MT at all. There was still a lot of planning to be done. What I did was a lot of planning on textual cohesion and coherence.

In the text generation stage, I found that it is related or overlap both the revising and planning process. This is in line with Koponen’s idea (2016) of overlapping stages. In generating the text, stylistic adjustment was not needed in the sentence level, and adjustment made to paragraph level was based on the preference. So, to the translator trainee using MT as platform, I would suggest to do no stylistic adjustment to the TT result by MT.

Finally, in the revising stage, for defect text, it turns out that re-writing is often conducted. One of the reason is the MT lack of background knowledge of the overall context and sentence stringing. This often happen to connective and conjunctive sentences. So, in applying this part as a sub-course, we can ask the students to pay extra attention to the revision when dealing with such types of sentences. Lastly, it turns out that the revising activities was less when the segments inputted int othe MT is larger in size.

One suggestion which I can give here is to conduct further research using more data since this research was only based on two data.

REFERENCES


