LATIP 2021
International Conference on Language and Technology in the Interdisciplinary Paradigm

PRINCIPLES OF TASK DEVELOPMENT FOR TRAINING POST-EDITING OF MACHINE TRANSLATION

Vlada Dorozhkina (a), Marina Ivleva (b)*
*Corresponding author

(a) Novosibirsk State Technical University, 630073 Pr. K.Marksa, 20, Novosibirsk, Russia,
hasanova_galieva@mail.ru
(b) Novosibirsk State Technical University, 630073 Pr. K. Marksa, 20, Novosibirsk, Russia, m.ivleva@corp.nstu.ru

Abstract

Quality of machine translation increased rapidly and this led to the emergence of the profession of a post-editor. Post-editing of machine translation (PEMT) is an activity of improving machine translation and is in high demand in the translation market. However, there is no PEMT training program at Russian universities, so some specialists post-edit intuitively and apply machine translation to inappropriate types of texts. We aimed at developing PEMT training tasks based on the criteria and requirements for post-editors that we compiled as a result of self-observation and a survey of post-editing professionals. Our action research employs several stages of task development and such methods as comparative analysis and statistical analysis. The proposed tasks were tested on third-year students of Novosibirsk State Technical University who major in translation. While creating tasks, we used text on topics in demand with language service providers (LSP) clients. Based on the results, we offered a list of principles, which would be helpful for producing further tasks and teaching PEMT program at Russian universities.

Keywords: Post-editing of machine translation (PEMT), translation training, teaching post-editing of machine translation, task development

1. Introduction

The education system does not always respond flexibly to the changes that constantly appear in the modern world. Russian educational paradigm, which relies heavily on theory and traditional training, is in the process of being replaced by a paradigm of useful knowledge that aims at practical results. Such a shift "from the theoretical to the practical" occurred owing to the changes in the requirements for new type of professionals in various spheres and the emergence of new professions.

Machine translation (MT) is a translation made by a computer program. With the appearance of neural machine translation systems, the quality of MT has significantly increased. That is why an increasing number of LSPs are using it to speed up the work and optimize the translation process. This has led to the emergence of the profession of a post-editor, a specialist who edits MT to achieve the quality of professional translation (as performed by a person).

Post-editing is a popular service that is now in high demand in the translation market. As well as abroad, in Russia post-editing is actively studied in its various aspects. At the theoretical level, a number of works are devoted to the architecture of MT systems (Nuriev, 2019), identification of typical errors of MT (Nechaeva & Svetova, 2018), research on the PEMT effectiveness (PROMT, 2020). However, post-editing is not taught in Russian universities. Post-editors are acquiring necessary skills attending specialized courses (for example, from SDL) and while working.

2. Problem Statement

Since there is no unified program for teaching PEMT, some post-editors process machine translation intuitively and apply machine translation to the texts that are suitable only for TEP (translation, editing, proofreading). Our aim is to develop a program and offer it for implementation at universities. Previously, we have compiled a list of criteria and requirements for post-editors based on the analysis of foreign PEMT training programs at universities and data from a survey of professional post-editors in Russia.

3. Research Questions

What kind of tasks is possible to develop based on list of criteria and requirements for post-editors? Would they increase the quality of full PEMT? What kind of principles should be used in further PEMT tasks?

4. Purpose of the Study

We aimed at developing a set of tasks for training PEMT based on list of criteria and requirements for post-editors compiled earlier. Based on the results, we offer a list of principles which would be helpful for producing further tasks and teaching PEMT program. The proposed tasks were to be tested on third-year students of Novosibirsk State Technical University who major in translation.

5.   Research Methods

In order to propose the principles of PEMT (2020) training activities, we analysed and compared some European PEMT training programs from different universities. They include: course at KU Leuven (Belgium), lectures at Universita di Bologna (Italy), additional course at University of Zagreb (Croatia), seminars at Universität Leipzig (Germany) and course at University of Helsinki (Finland) (PEMT, 2021). All these educational programs cover the structure of MT systems and their types, errors made by MT systems, and methods of evaluating MT (manual and automatic). At the practical level, different types of PEMT (light\full) and their differences were considered, and PEMT strategies were proposed.

In addition, we made survey intended for professional post-editors. It was taken by 26 respondents and most of them were working in TA *AUM* (Novosibirsk, Russia). The main aim of the survey was to determine necessity of teaching PEMT at universities and linguistic difficulties with PEMT. Results show that 75% of post-editors believe that PEMT needs to be taught in Russia. The most difficult components are phraseological units and culture specific terms (92%); genre features of the text (67%); stylistics (58%) and the choice of connotative meanings (46%).

Combining survey data and comparative analysis of European PEMT training program we made the list of criteria and requirements for post-editors. The post-editors should possess the following knowledge and skills:

- Know text registers in order to be able to determine the PEMT (2020) applicability, since some types of texts are easier to translate "from scratch" (literary, journalistic);
- Understand the mechanisms of different types of MT systems, their advantages and disadvantages;
- Know errors that are typical for MT systems, be able to evaluate the quality of MT manually and automatically;
- Perform different types of post-editing (light and full) for different tasks;
- Be able to prioritize, starting with post-editing elements of cognitive information and ending with elements of emotional; adhere to certain strategies in their work;
- Know collocations, culture specific terms, idioms of the source language and the target language (Khasanova & Ivleva, 2019).

These requirements are the basis of our task development.

5.1. Developing PEMT tasks and preparing materials

First of all, we prepared a lecture based on theoretical studies and European programs listed above. It covers differences between translation and PEMT, types of MT systems and types of PEMT, areas of PEMT applicability and PEMT strategies.

Next, we developed tasks with the focus on three important aspects of the post-editing process: determining the type of PEMT (corresponds to criteria 4), determining PEMT applicability (criteria 1), and evaluating the results of the MT system (criteria 3) (Dorozhkina & Ivleva, 2020).

And finally, we chose a MT system as we wanted students to get the same MT output consistently. We recommended using "Yandex Translator" because it is neural MT system, which provides coherent

text for a wide range of text registers. According to the TA TranExpress study, Yandex system demonstrates the best of quality for the Russian language (TranExpress, 2020).

5.2. Working with group 1 and group 2

To formulate task-developing principles, we planned and conducted an experimental class for 2 groups of third-year students who major in linguistics. Each group was given both theoretical information and practical tasks. All of them were asked to use "Yandex Translator" (Mitchell, 2012). The action research was conducted with the group 1 of 14 students with subsequent statistical analysis. Processing the analysis results, we made necessary changes and repeated the updated PEMT training with group 2 of 13 students. The results of the two groups were compared as well.

6. Findings

Both groups of third-year students received an introductory lecture with all required theoretical information, and then group 1 was asked to do a set of 3 tasks. Having analyzed the results of this group, we updated the tasks and gave them to the second group. The obtained results and their evaluation helped formulate the principles of PEMT task development.

6.1. First task: types of PEMT

In some tasks here we include not all content from tasks as they are quite long.

Task 1. *Determine the type of post-editing (light or full).*

EN: The human brain is made of densely packed neurons. Some estimates place the number at more than 100 billion (about the number of stars in the Milky Way), each of which is capable of receiving and passing neural impulses to sometimes thousands of other neurons and is more complex than any other system known to be in existence (Solso, 2014).

PEMT: Головной мозг человека содержит огромное количество нейронов. По некоторым оценкам, их число составляет более 100 миллиардов (примерно столько звезд в Млечном Пути). Каждый из нейронов способен принимать нервные импульсы и и передавать…

EN: The study of sensation generally deals with the structure and processes of the sensory mechanism and the stimuli that affect those mechanisms. Perception, on the other hand, involves higher-order cognition in the interpretation of the sensory information. Basically, sensation refers to the initial detection of stimuli; perception to an interpretation of the things we sense (Solso, 2014).

PEMT: Изучение ощущений обычно связано с устройством и работой органов чувств и со стимулами, которые воздействуют на эти органы. С другой стороны, восприятие подразумевает участие высших когнитивных механизмов в интерпретации сенсорной информации...

The results of both groups are presented in the Table 1.

Table 1. Results of task 1

Correct answers	Group 1	Group 2
1) Full PEMT	78,6%	84,6%

2) Light PEMT	85,7%	92,3%
3) Full PEMT	**71,4%**	84,6%

Obviously, group 1 did not always manage to determine Full PEMT in the last text. The post-editor made minimal changes to the text, so the students were not able to recognize Full PEMT. Consequently, it is important to note in the rubric that the type of PEMT is determined not by the number of changes made, but by the final result. That is why we reformulated the rubric for the group 2 and the result was better:

Task 1. *Determine the type of PEMT (light or full). When completing the task, you should pay attention not to the number of changes made, but to the quality of the PEMT output.*

6.2. Second task: PEMT applicability

Task 2. *Evaluate the applicability of the PEMT for each text using pre-translation analysis. Determine the target audience, the communication task, the text register and its main type of information and make a conclusion about the applicability of the PEMT.*

Text 1. The first step in devising efficient and effective methods for teaching women equitation, is to make certain that both riders and their instructors know the differences between male and female anatomy, and understand how such important parts as the hip sockets, lower back and seat bones of each sex function in the saddle (Bennet, 2008, p. 1).

Text 2. Join some of your "Frozen" favorites—Anna, Elsa, Olaf, and Kristoff—around the campfire as they tell legends exploring the world of "Frozen 2." You can hear these stories on Google Assistant-enabled Android and iOS phones, smart speakers and Smart Displays. To get started, just say, "Hey Google, tell me a 'Frozen' story" and you can pick which character you'd like to narrate (Byer, 2020).

Text 3. To better handle that kind of nuance, a group led by researchers at MIT's Computer Science and Artificial Intelligence Lab (CSAIL) has developed a machine learning model that can look at an X-ray to quantify how severe the edema is, on a four-level scale ranging from 0 (healthy) to 3 (very, very bad). Results are presented in Table 2:

Table 2. Results of task 2

Correct answers	Group 1	Group 2
Yes	85,7 %	84,6 %
No	57,1%	84,6 %
Yes	78,6 %	76,9 %
No	85,7 %	92,3 %
Yes	92,8%	92,3 %

For the second text, only slightly more than half of group 1 chose the correct answer. This text is an advertisement for the Google Assistant feature. Notably, 83.3% of students who gave the wrong answer identified the type of text as an instruction. According to Alekseeva (2004), the advertising text and the instructions have a common feature: the predominance of operational information. It is often expressed by a large number of imperatives, which can be seen in text 2. The difference is that the

459

advertising text combines operational and emotional types of information, while the instructions contain operational and cognitive.

In addition, the students of group 1 did not pay attention to critical meaning errors in MT output. The output may seem logical to a post-editor who has no experience in the IT field, and affect the assessment of PEMT applicability accordingly. Text 4 is also an advertisement (KitKat advertisement), and a larger number of students gave the correct answer. This is due to the fact that the topic of chocolate was more understandable to the participants of group 1. So for group 2, text 2 was replaced with text on a familiar topic – the advertisement of sport app without IT-specific descriptions – and the result was better.

6.3. Third task: MT output errors

Task 3. *Specify the type of MT in each case. Use the LISA classification. Note that there may be several errors in a single fragment. The categories of errors can also be duplicated (for example, if you see that a sentence has both a typo and a punctuation error, put LA twice).*

1. There is another side to the sensory and perceptual process that is supported by studies… (Solso, 2014) Есть другая сторона к сенсорному и перцепционный процессу, которая поддержан исследованиями...

2. The pelvic construction of men and women is different (Bennet, 2008, p. 2). Тазовое строительство мужчин и женщин отличается.

3. How can Britons simultaneously be both self-controlled and prone to rip our clothes off in a drunken haze? Как британцы могут одновременно быть самостоятельным и склонны рвать одежду в пьяном тумане?

4. Heaven forfend–remember the Monkey's Paw! Небеса защитят – помните Лапу Обезьяны!

5. Many health issues are tied to excess fluid in the lungs. A new algorithm can detect the severity by looking at a single X-ray (Conner Simons, 2020). Многие проблемы со здоровьем связаны с избытком жидкости в легких. Новый алгоритм может определить степень тяжести, посмотрев на один рентгеновский снимок.

Results are shown in Table 3:

Table 3. Results of task 3

Correct answers	Group 1	Group 2
ME, LA, LA, LA	64,2%, 85,7 %, 85,7%, 78,6%	84,6 %, 92,3%, 84,6%, 76,9%
TE, ST	78,6%, 57,1%	100%, 69,2%
ME, ST, LA	64,2 %, 85,7%, 92,8%	69,2%, 84,6%, 76,9%
LA, ST	78,6%, 57,1%	92,3%, 100%
ME, ST	42,8%, 78,6%	76,9%, 84,6%

Sentence 1. The semantic error was revealed by a smaller number of group 1 students. There is ambiguity in the sentence: it is not clear whether studies support the side or the process. It is recommended to draw the attention of students to the fact that ambiguity is also a meaning error. We did so with group 2 and the result was better.

Sentence 2. A stylistic error ("отличается – различается") was noticed by a little more than half of the students of both groups. This issue should be elucidated within the existing course *Practical course of translation*. We should explain the significance of connotative meanings of words. Importantly, 3 out of 8 students who answered correctly, indicated that they were not sure whether the error was connected with style or meaning. In this connection, a survey of 11 professional editors working in different organizations was conducted, and 9 out of 11 (81,8 %) specialists identified the categories of edits as stylistic, arguing that the wrong choice of prefix leads to a distortion of the text logic, while the reader is more likely to understand the meaning.

Sentence 3. The meaning error ("self-controlled – самостоятельным") was noticed by a little more than half of the students of both groups. The following tendency is revealed: if several errors are made in the text, the meaning errors are harder to see. It is highly recommended to add more tasks on searching for meaning errors made by the MT system and to focus on them in complex tasks.

Sentence 4. Notably, some of the remaining students who did not identify the stylistic error indicated it as a meaning one (4 people from group 1 and 3 people from group 2). The meaning of the set expression *Heaven forfend* ("Боже упаси") is indeed distorted by the machine translation system, but the error is still stylistic, since it appeared because of missing cultural adaptation and literal translation. Approaches to these error categories differ and it is worth analyzing the differences. It is also recommended to focus students' attention on the fact that only one category can be assigned to an error. In our case, if the meaning is distorted due to the word-to-word translation of an original set expression, the error is stylistic.

Sentence 5. Participants of group 1 did not identify ambiguity: it is not clear whether the *severity* refers to health problems or to the filling of the lungs with fluid. As for group 2, putting ambiguity in a separate subcategory affects the result.

7. Conclusion

According to analysis of results, we compile a list of principles:

- For the PEMT type recognizing tasks, we recommend to emphasize that the type of PEMT should be determined not by the number of changes made, but by the quality of the final result.
- Preferably, beginning students should be familiar with the subject of the text.
- When forming tasks for assessing PEMT applicability, we recommend clarifying that the density of terminology is an important, but not the only factor to determine the PEMT applicability. Students should simulate "client-customer" communication and evaluate PEMT applicability, taking into account not only the linguistic features of the text, but also the glossary, the MT type, PEMT type, etc.
- For teaching PEMT, gradually increasing complexity of the texts for post-editing is important: fragments containing one error should be replaced by texts with several errors.
- Attention should be paid to differentiating the errors (for example, stylistic error from meaning or terminological error). Different approaches are applied to different types of errors.

- Meaning errors made by the MT system should be considered on numerous occasions, especially in complex tasks. Meaning errors are the most serious and they should be eliminated first.
- It is worth encouraging students to use additional fact checking sources when performing PEMT. It is necessary to teach how to search for information effectively.

References

Alekseeva, I. S. (2004). *Vvedenie v perevodovedenie: Ucheb. posobie dlya stud. filol. i lingv, fak. vyssh. ucheb. zavedenij* [Introduction to Translation Studies: Textbook]. Academia.

Bennet, D. (2008). Who's built best to ride? Happy Horse training.

Byer, N. (2020). *Need a dose of Disney+? Just ask!* Retrieved on April 2021 from: https://blog.google/products/google-nest/need-dose-disney-just-ask/

Dorozhkina, V. A., & Ivleva, M. A. (2020). Post-editing of machine translation in teaching students majoring in linguistics. *Teaching Methodology in Higher Education, 9*(33), 38-45.

Khasanova, V. A., & Ivleva, M. A. (2019). *Analysis of requirement for post-editors of machine translation*. Novosibirsk State Technical University, 695–698.

Mitchell, D. (2012). *Cloud Atlas*. Sceptre.

Nechaeva, N. V., & Svetova, S. Y. (2018). Post-Editing Machine Translation as a New Activity for Teaching Translation at Universities. *Teaching Methodology in Higher Education, 7*(25), 64–72. https://doi.org/10.18720/HUM/ISSN2227-8591.25.07

Nuriev, V. A. (2019). Architecture of a machine translation system. *Informatics and Applications, 13*(3), 90–96.

PEMT (2020). *PEMT educational courses*. Retrieved on April 2021 from: http://pemt.ru/study/st-education-%D1%81ourses/

PROMT (2020). *Types of machine translation systems*. Retrieved on April 2021 from: http://www.promt.ru/company/technology/machine_translation/index.php?sphrase_id=21480169

Solso, R. L. (2014). *Cognitive Psychology. 6th edit*. Pearson.

TranExpress (2020). *The best machine translation system 2020*: testing results. Retrieved on April 2021 from: https://www.tran-express.ru/blog/luchshiy-online-perevodchik-dlya-angliyskogo-2020-rezultaty-testirovaniya