Pavlíková, Diana. 2022. Comparing Human Translation and Machine Translation Post-Editing. In: L10N Journal. 1(1), pp. 65–100.

Comparing Human Translation and Machine Translation Post-Editing

Diana Pavlíková

Comenius University in Bratislava

dianapavlikova@protonmail.com

Translation: Alex Barák

Abstract

As machine translation becomes a part of the translation market around the world, post-editing appears to be an increasingly used alternative to human translation. The aim of this article is to compare human translation with the post-editing of machine translation through an experiment on students at master's level in the specialization of Translating and Interpreting in a language combination with English at Comenius University and to discover which process provides better and more time-efficient translations. The results acquired through a quality assessment of the submitted translations were analyzed based on a temporal aspect of effort and quality. The findings of the analyses are interpreted and the results are explained to reveal the strengths and weaknesses of human translation and the post-editing of machine translation.

Keywords: machine translations, post-editing of machine translations, human translations, post-editing, LQA

Introduction

Machine translation has become a part of the translation market worldwide. Combined with post-editing, machine translation can be an alternative to human translation in certain language combinations and particularly in branches of specialized translation (Tatsumi 2009; Plitt & Masselot 2010; Green et al. 2013). This article will compare human translation with machine translation post-editing and find out which of the two processes is better in terms of the speed of work and the quality of the final translation.

It will summarize the basic theoretical foundations of machine translation and machine translation post-editing before introducing Memsource Translate, Memsource's feature[1] for managing machine translation tools, and translation quality assessment. The main part of the article will focus on an experiment conducted with students from Comenius University enrolled in the master's degree program in Philology with a focus on translation and interpreting in English in combination with another language and on a qualitative and temporal analysis of the results.

1 Machine translation

Nowadays, we can no longer imagine a translator who does not work on a computer. Technologies such as CAT tools, terminology databases, and memories facilitate the translation process; indeed, machine translation, "the mechanical and automatic process of translating a text from one natural language to another" (Munková 2013, 16), is being increasingly mentioned. The main objective of the complete computer automation of the translation process is to speed up translators' work and meet the growing demand in the translation market (Hutchins & Somers 1992). Over the last ten years, the demand for translation services has grown so significantly that it exceeds the supply of translation service providers (Hudecová et al. 2021). Machine translation is generally considered fast, cheap (Google Translate is even free for ordinary users), and accessible; in addition to online translators such as Google Translate, machine translation can also be found in unpaid versions of some CAT tools.

Despite these advantages, machine translation is not perfect and cannot always be relied upon. The present author compared human translation and Google Translate's neural machine translation in four areas of expertise; although positive results were recorded, there was also an error rate (10.5% on average), especially when translating polysemous terms (Pavlíková 2021). The conclusion was that in order to achieve high quality translation using machine translation, post-editing is a necessary process (Pavlíková 2021).

[1] Memsource is currently called Phrase, and the Memsource Translate tool is called Phrase Translate. Since the renaming occurred only shortly before the article was published, the old names are still used.

Pavlíková, Diana. 2022. Comparing Human Translation and Machine Translation Post-Editing. In: L10N Journal. 1(1), pp. 65–100.

1.1 Neural machine translation

Neural machine translation has the most potential in the translation market, as it is a technology that, compared to older types of machine translation, has the "ability to learn directly, in an end-to-end fashion, the mapping from input text to associated output text" (Wu et al. 2016, 1). Gene (2019) presented several reasons why neural machine translation is more efficient compared to other types of machine translation. A neural machine translation system:

- is able to learn about the complex relationships between the two languages it works with
- consider whole sentences, understanding the relationships between words despite the greater distance within the sentence unit
- assess fluency at the level of the whole sentence

Despite its efficiency, neural machine translation has some shortcomings, such as the incompleteness of translations for longer texts, the slow training of the system on a large corpus, and problems when translating less frequent words (Wu et al. 2016).

1.1.1 Neural machine translation and machine translation post-editing

Machine translation is often combined with machine translation post-editing. Indeed, "MT output today still needs to be post-edited by humans in order to produce publishing quality translation" (Tatsumi 2010, 3). Machine translation post-editing has been used with neural machine translation, and recent studies have demonstrated productivity gains for post-editors using neural machine translation post-editing compared to those just using translation memory (Sánchez-Gijón, Moorkens, & Way 2019; Läubli et al. 2019).

Nonetheless, the post-editing of neural machine translation can create problems. As the output of neural machine translation becomes more fluent and natural, the post-editor needs to be more careful in properly identifying and correcting errors. In addition, the neural machine translation system works within the context of a single sentence. When working on a longer text, post-editors must look for and edit for consistency across the entire work (Gene 2019). Despite these problems, the combination of post-editing with neural machine translation seems to be advantageous since neural machine translation has significant advantages over older types of machine translation. Post-editing will help to raise the output text to a publishable level comparable to human translation.

2 Machine translation tools

Translators encounter machine translation when using CAT tools. In addition to the built-in translators in CAT tools, there are also freely available translators for ordinary language users. The experiment discussed here used two machine translation tools: Memsource Translate (Memsource's tool that helps the user to select the most suitable translator for a particular project) and Google Translate as a neural machine translation tool set for the language combination of English and Slovak.

2.1 Memsource Translate

Memsource Translate is used to manage machine translation tools. Memsource runs an online cloud-based CAT tool, which is one of the top three most used translation systems, mainly due to its accessibility via a web browser and its simple user interface (Dengová 2020). The goal of Memsource Translate is to make it easier for users to choose from the large number of machine translation tools available on the market. Memsource Translate selects from three online translators: Amazon Translate, Microsoft Translator, and Google Translate. The Memsource Translate algorithm selects the most suitable of these three tools primarily based on the language pair; thanks to updates in 2020, it can also distinguish between several specialized fields based on keywords and thus select the appropriate translator according to whether the text is a legal, technical, or medical one (Frívaldský 2020; Pavlíková 2021).

For projects with English as the source language and Slovak as the target language with texts from multiple fields, Memsource Translate primarily chooses Google Translate, which works on the basis of neural machine translation, meaning that it primarily evaluates it as the most suitable for this language combination.

3 Machine translation post-editing

Machine translation post-editing is one process used to edit the output of a machine translation. In Slovak, Absolon (2018) identifies three terms associated with this process: *posteditácia* (post-editing) represents the final product, *posteditovanie* (post-editing) is the actual process of using machine translation to produce a suitable translation in the target language, and *posteditácia strojového prekladu* (machine translation post-editing) is a unifying term for the product and the process of post-editing. This terminology is often not used, as the term "machine translation post-editing" is lengthy. "Post-editing" will therefore be used in the sense of "machine translation post-editing"

in the rest of this article. Machine translation post-editing has several possible terms as well as definitions. Generally speaking, machine translation post-editing is the inspection of a pre-translated text by a machine translator, after which the post-editor corrects possible errors in order to meet specified quality criteria with the fewest possible edits (Mesa-Lao 2013). The establishment of adequate quality criteria and the adherence to principles (such as the aforementioned insistence on as few edits as possible) mentioned in the definition are crucial in post-editing but are still inconsistent in the translation world (Hu & Cadwell 2016).

3.1 The advantages and disadvantages of machine translation post-editing

Machine translation post-editing appeared on the translation market as a faster, cheaper, and more efficient substitute for human translation. The effectiveness of post-editing versus translation has been confirmed in several studies. According to Robert (2013), post-editing can increase the average number of words translated per day from 2,000 to 3,500 words. Guerberof Arenas (2010) even reports an increase up to 5,000 words; however, she adds that the increase in word count can vary depending on the type of text and the experience of the post-editor. Faster post-editing is also in demand due to the growth of the localization industry and globalization, and many businesses are now moving towards a multilingual expansion. Thanks to this, localization service agencies are experiencing a 30% to 50% annual increase in the number of orders (Allen 2003). This growing demand for post-editing was confirmed by the Common Sense Advisory research agency in 2016, when it predicted that demand for post-editing services would grow faster than any other segment of the language industry. Through previous research, they found that language service providers who used machine translation between 2013 and 2015 grew almost 3.5 times faster (Common Sense Advisory 2016). Last but not least, the growth of post-editing is helped by the change in requirements for the target text from clients. Nowadays, translation quality requirements have relaxed and many companies commission translations for internal purposes; these are not published anywhere and do not require high quality human translation (Bubnic 2022).

Despite the advantages and popularity of machine translation post-editing, it is still a new process in the translation industry, and it has its shortcomings. According to Gene (2019), there are several challenges that are yet to be resolved in machine translation post-editing:

- the absence of general principles and rules for post-editing
- the inconsistent competences and requirements for the post-editor
- payment for post-editors

In addition, Doherty and Gaspari (2013) mention the lack of teaching and training materials for post-editors as another drawback. As machine translation post-editing is rarely taught as a separate discipline, there is a lack of teaching materials and general principles for post-editing that could better establish uniform competences and requirements for post-editors and address the issue of how much to pay them.

3.2 Types of post-editing

Post-editing is generally faster than human translation; however, there are several types of post-editing that can be influenced by factors other than speed, such as the number of edits or price. Two basic types of post-editing are usually mentioned. Allen (2003), however, introduced three types:

1. **Light post-editing** (also known as Light/Rapid/Fast MTPE) is used for texts written for internal purposes. The main task of the post-editor is to make as few edits as possible and preserve the meaning of the text without having to notice stylistic imperfections.
2. **Full post-editing** (also known as Full/Conventional MTPE) is used when translating texts that will be published. It requires a target text of high quality comparable to human translation. Several changes are therefore expected, and, in addition to preserving the meaning of the text, it is important to achieve correct syntax, grammar, and punctuation.
3. **Minimal post-editing** is like light post-editing, where the post-editor tries to make as few edits as possible, but it works with texts that are also used for external communication.

The experiment in this article worked with full machine translation post-editing. The respondent sample tried to produce an output translation of comparable quality to a human translation.

3.3 The post-editing process

Doherty and Gaspari (2013) state that post-editing, specifically full machine translation post-editing, works with three texts:

1. A source text
2. A raw machine translation output
3. A post-edited machine translation output/target text

As this involves working with three texts, the overall post-editing process involves up to five steps according to the TAUS (2010) model:

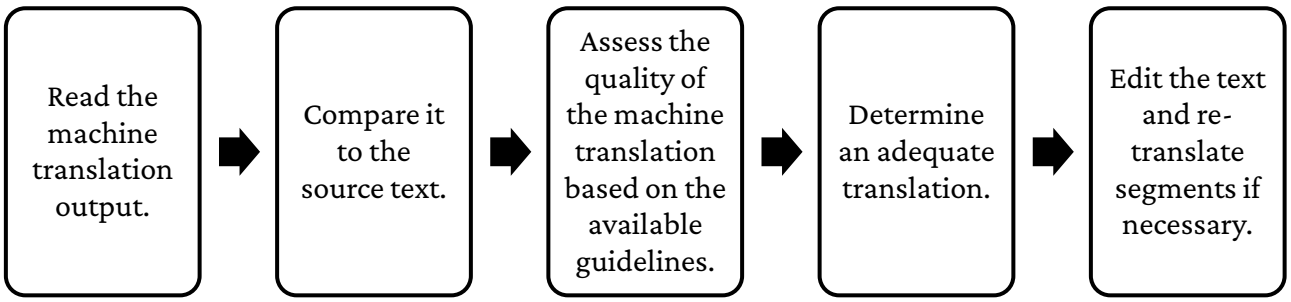

Figure 1 – ***A model of the post-editing process***

According to this model, a post-editor must complete four steps before editing the text itself. The whole process takes place at the level of segments. Post-editors usually work with CAT tools, which segment the text based on predefined rules. Although segmentation helps to keep the text clear, the post-editor must be careful to perceive the text as a whole and adjust for any inconsistencies at the segment level and throughout the text.

According to Krings (2001), a post-editor performs several types of tasks, such as those related to the source text (reading the source text and noticing its elements), machine translation, working with sources, and typing on a computer. Most of the tasks are related to the production of the target text, where the post-editor works with the available post-editing guidelines in order to produce an adequate target text. Gene (2019, 10) summarized the most common changes a post-editor makes to the output of machine translation: "correcting punctuation, adding capital initial letters, changing word order, adjusting grammatical gender and number correspondences, formatting changes, adding omitted words or removing redundant words, and re-translating words or phrases unsuitable for the target text."

3.4 The post-editor

Machine translation post-editing has created new jobs on the translation market within the group of language service providers. Mesa-Lao (2013) reports that several translation agencies still use freelance translators for post-editing; however, not every translator can handle the job of a post-editor. This is because experienced translators find it more difficult to get used to producing a target text with a lower level of quality compared to translation (Allen 2003). The post-editing and translation processes are different from each other since in translation one works with only two texts instead of three; post-editors need different competences and put a different level of effort into their work. This is because post-editors need to have knowledge about machine translation and be able to identify the errors that machine translation makes. Compared to translators, they also need to be more decisive and more easily overcome uncertainty in order to be able to produce the target text more quickly and according to the client's requirements and expectations (Torrejón & Rico 2012).

3.5 Post-editing guidelines

Post-editing guidelines help post-editors meet clients' expectations and create a target text of adequate quality. There is an inconsistency in these guidelines on the market, as each agency and organization tends to create their own guidelines for internal use (Hu & Cadwell 2016). As part of the experiment discussed in this article, the freely available full machine translation post-editing guidelines of the TAUS organization were used.

TAUS Machine Translation Post-editing Guidelines (2016)[2]

- Aim for grammatically, syntactically, and semantically correct translation.
- Ensure that key terminology is correctly translated and that untranslated terms belong to the client's list of "Do Not Translate" terms.
- Ensure that no information has been accidentally added or omitted.
- Edit any offensive, inappropriate, or culturally unacceptable content.
- Use as much of the raw MT output as possible.
- Basic rules regarding spelling, punctuation and hyphenation apply.
- Ensure that formatting is correct.

[2] The post-editing guidelines from TAUS were translated into Slovak by the author of the paper. The original TAUS Machine Translation Post-editing Guidelines are available at: https://www.taus.net/academy/best-practices/postedit-best-practices/machine-translation-post-editing-guidelines.

These guidelines include several categories that the post-editor must know about. In addition to creating a grammatically, syntactically, and semantically adequate text, they include the choice of appropriate terminology, the use of proper punctuation, and correct formatting.

3.6 Teaching and the current use of machine translation post-editing in Slovakia

Machine translation post-editing is becoming increasingly popular in the world of translation. O'Brien (2002) states that studying machine translation post-editing allows students to learn how to work with machine translations and acquire post-editing skills and competences. In addition, teaching machine translation post-editing helps meet the growing demand for translation services by presenting a quicker alternative and helping standardize post-editing guidelines.

Universities are beginning to offer post-editing instruction in courses devoted to computer-assisted translation, machine translation, and localization (e.g., Kabát 2022). In Slovakia, however, only the University of Constantine the Philosopher in Nitra specifically mentions post-editing on its web pages. As for the use of machine translation post-editing, and despite the high demand for machine translation globally, in Slovakia "machine translation post-editing as a product still represents only a small percentage of the market" (Absolon 2018, 95).

Absolon (2018) foresees an increase in demand for post-editing in the domestic market, but he expects that this practice will not be openly presented as machine translation is still perceived negatively by many clients. This assumption is being seen in the services offered by translation agencies. Only a small percentage of them explicitly mention post-editing in their machine translation services, using instead terms such as "revision," which clients are more familiar with.

4 Translation quality assessment

Translation quality assessment (also known as linguistic quality assessment) is "the process of evaluating the overall quality of a completed translation by using a model with pre-determined values which can be assigned to a number of parameters used for scoring purposes" (Korkas n.d.). This quality assessment system is being increasingly used by translation agencies and large companies to assess the quality of translations and machine translation post-editing. This assessment is done by revisers or

proofreaders using a template (usually in an Excel document) where they enter translation errors, possible corrections, and comments (Kabát 2022). The main goal of translation quality assessment is to "identify quality gaps and propose solutions for translators to ensure that the translation meets the qualitative needs and expectations of the client" (Finnegan 2018).

4.1 The TAUS quality assessment template

The TAUS quality assessment template is used to evaluate human or machine translation and machine translation post-editing of any type of input text. This template helps the user to count, identify, and categorize errors in the translation; rate the quality of the translation with a grade; suggest possible solutions in the comments; and improve the overall quality of the translation (TAUS n.d.). It includes an introductory sheet that includes basic project information, instructions on how to use the template, a description of the error typology and error severity levels, and a space for listing errors in each segment. In this space, the reviser or proofreader lists the source segment and its translation, suggests a corrected segment, and categorizes the type of error and the severity level. If necessary, it is possible to leave a comment for the translator. Different error severities have different scores, which add up to penalty points. After completing the template, the proofreader sees the final score, which can be used to evaluate whether the translation was adequate (according to the number of errors and the resulting percentage of correctness) or whether it had numerous failings.

The typology of errors in the TAUS template contains eight categories in which thirty-three types of errors are included; they cover morphological, lexical, stylistic, syntactic, terminological, formatting, localization, and cultural issues.

Pavlíková, Diana. 2022. Comparing Human Translation and Machine Translation Post-Editing. In: L10N Journal. 1(1), pp. 65–100.

Table 1: ***Categories and the typology of errors in the TAUS quality assessment template*** [3]

Category	**Typology of errors**
Accuracy	Addition Omission Mistranslation Over-translation Under-translation Untranslated text Improper exact TM match
Fluency	Punctuation Spelling Grammar Grammatical register Inconsistency Link/cross-reference Character encoding
Terminology	Inconsistent with termbase Inconsistent use of terminology
Style	Awkward Company style Inconsistent style Third-party style Unidiomatic
Design	Length Local formatting Markup Missing text Truncation/text expansion
Locale convention	Address format Date format Currency format Measurement format Shortcut key Telephone format
Verity	Culturally specific references
Other	Other

The severity of quality assessment interventions is indicated by five levels:

[3] The typology of errors was translated into Slovak by the author of the paper.

1. **Critical:** errors that may carry health, safety, legal, or financial implications; violate geopolitical usage guidelines; damage the company's reputation; cause the application to crash or negatively modify/misrepresent the functionality of a product or service; or that could be seen as offensive
2. **Major:** errors that may confuse or mislead the user or hinder proper use of the product/service due to a significant change in meaning or because errors appear in a visible or important part of the content
3. **Minor:** errors that do not cause a loss of meaning and would not confuse or mislead the user, but which would be noticed; would decrease stylistic quality, fluency, or clarity; or would make the content less appealing
4. **Neutral:** used to log additional information, problems, or changes to be made that do not count as errors (a reviewer's choice or preferred style, repeated errors or instructions/glossary changes not yet implemented, or a change to be made that the translator is not aware of)
5. **Kudos:** used to praise an exceptional achievement

The TAUS template is one example of what a quality assessment template looks like. Some translation agencies customize their templates based on the types of texts they translate most often or on the type of provided language service (e.g., translation, machine translation post-editing, or localization).

5 A comparison of human translation and machine translation post-editing

To compare human translation and machine translation post-editing, there was an experiment where twenty translation and interpreting students were divided into two groups. One group translated the text and the other post-edited it, and both groups were familiarized in advance with the machine translation post-editing process.

5.1 Methodology

Twenty students from the first and second years of the master's program in Philology with a specialization in translation and interpreting in a language combination with English took part in the experiment. They already had experience with translation, so the focus was on machine translation post-editing when preparing students for the experiment.

The experiment took place in the Localization of Game and Non-game Software course at the Department of British and American Studies, Faculty of Arts, Comenius University in Bratislava, under the supervision of the present author via MS Teams. (This was due to the pandemic situation during the summer semester of the 2020/2021 academic year.) Students had studied the translation and post-editing guidelines that had been provided in advance to become more familiar with the translation and post-editing requirements. In addition, they were given general instructions describing the procedure of the experiment.

On the day of the experiment, the twenty students were divided into two groups of ten students based on alphabetical order by last name. The first group translated the text and followed the translation instructions. The second group post-edited the text and followed the post-editing instructions. For the machine translation post-editing, students performed full post-editing and followed the 2016 post-editing guidelines from TAUS.

Both groups were provided with files via the school's Moodle platform, which they uploaded to Memsource after starting the experiment. The first group received a .docx file for translation and the second group received an .xliff file for post-editing, which contained the raw machine translation extracted from Memsource using the selected Google Translate tool. In addition to the translation and post-editing files, the students were provided with terminology in .tbx format, which they had to upload to their own terminology database. The terminology included basic terms for the chosen topic to make the students' work of searching easier.

The translation/post-editing text was an excerpt from an American brochure on composting; specifically, it was a manual on how to build a home vermicomposter. It was a specialized marketing text, and it addressed the reader directly. The manual had 262 words, and the Memsource tool divided it into twenty-six segments. After uploading the files to Memsource, students worked on post-editing or translation. When finished, they exported the file and uploaded it back as a bilingual file via the Moodle platform. Students were also advised that their time would be measured, so they were encouraged to hand in the final file as soon as possible to avoid further distortion of the time results.

5.2 The objective and the research questions

The objective of the research was to compare human translation with machine translation post-editing on a given sample of students, and to find out which process was better in terms of speed of work and the quality of the final translation. During quality assessment, the following questions were looked into:

- Which group was more correct based on the average error rate?
- Which group was more correct based on the average number of penalty points?
- What types of errors were predominantly made by translators, and what types of errors were made by post-editors?

In addition to the qualitative analysis, the temporal aspect of effort for both groups was also examined:

- Which group did the translation faster?

Answering these questions in a qualitative and temporal analysis would indicate how human translation and machine translation post-editing differ from each other and what their strengths and weaknesses are.

5.3 Qualitative analysis

The quality of the translations submitted by the students was assessed using the TAUS translation quality assessment template. As part of the penalization process, a scoring system from the template was used.

Table 2: ***The penalty point system based on the TAUS template***

Severity level	Number of penalty points
Critical error	10
Major error	5
Minor error	1

The correctness of the submitted translation was evaluated by the number of penalty points the translator/post-editor received. The qualitative analysis focused on how correct the translators and post-editors were based on the penalty points, how many and what types of errors were made (and at what levels of severity), and how they differed from each other in these categories, while also giving specific examples of the most common errors made by the two groups.

5.3.1 The translation error rate

As part of the qualitative analysis, the error rate of the translators and post-editors was examined. The errors were then categorized according to their severity level. The critical errors were mainly errors in numerical data and unit conversion which could lead to the failure of the manual. Major errors included those errors that changed the meaning of the source text and could also lead to the non-functionality of the manual. (These were mainly in the categories of accuracy and fluency of translation, terminology, and localization.) Minor errors were mostly related to the categories of style, design, and fluency of translation, which also included errors in the use of punctuation and mathematical symbols.

Table 3: ***The average error rate based on error severity level***

	Average error rate	
Severity level	**Translators**	**Post-editors**
Critical error	0.4	0
Major error	4.7	7.2
Minor error	2.7	2.8
Total	**7.8**	**10**

After counting the errors for each translator and post-editor and creating an average, there was a higher error rate for post-editors, who made an average of ten errors. Translators had an average error rate of 7.8. A closer analysis of these results revealed that in the critical error category, which influences the correctness of the translation the most, the post-editors had an average error rate of 0 whereas the translators had an average error rate of 0.4. The translators made more serious errors compared to the post-editors, which could have negatively affected the quality of the translated manual. However, post-editors made significantly more major errors, averaging 7.2, while translators had an average error rate of 4.7. The post-editors also made errors that negatively affected the functionality of the vermicomposter manual. In the minor error category, which mostly evaluated inconsistencies in tags, punctuation, and mathematical symbols, post-editors and translators performed similarly.

5.3.2 Translation correctness

We evaluated the translation correctness in both groups based on the number of penalty points.

Table 4: ***The average number of penalty points***

Group	Average number of penalty points
Translators	30.2
Post-editors	38.8

An average in the two groups was created for comparison. Since the penalty points were closely related to the translation error rate, the translators were also better off in terms of translation correctness. They had an average of 30.2 penalty points, and the post-editors had an average of 38.8 penalty points.

5.3.3 Error typology

The last part of the qualitative analysis is the evaluation of the number of errors based on the TAUS error typology template along with the error subcategories and individual examples. The TAUS template contains eight basic error categories; however, there were no errors in the verity and "other" categories, so only the six main categories of errors shall be looked into.

Table 5: ***The average error rate based on error typology***

	Average error rate	
Error typology	**Translators**	**Post-editors**
Accuracy	2.8	3.5
Fluency	3.4	4.4
Terminology	0.7	0.9
Style	0.1	0
Design	0.2	0.9
Locale convention	0.6	0.3
Verity	0	0
Other	0	0

Of the six analyzed categories, post-editors showed higher error rates in four of them: accuracy, fluency, terminology, and design. The biggest difference was seen in the category of fluency, where translators had an average error rate of 3.4 and post-editors had an average error rate of 4.4. This category encompassed a wide range of errors represented by each subcategory, so the evaluation of specific errors will be analyzed in more depth; however, it appears that the translators' work proved to be more fluent.

In the remaining two categories, translators showed a higher average error rate. In the style category, there was an average error rate of 0.1 for the translators and no errors for the post-editors. The post-editors also did better in the locale convention category, with an average error rate of 0.3, while translators had an average error rate of 0.6. To see how and in what ways the errors made by post-editors and translators differed, the next section of the article looks at the subcategories of each error category and specific examples.

5.3.3.1 Accuracy

Accuracy draws attention to cases in which the translation does not reflect the source text and does not correctly transfer its meaning or purpose. This category encompasses seven types of error: addition, omission, mistranslation, over-translation, under-translation, untranslated text, and improperly exact TM matches.

Table 6: ***The average error rate in the accuracy category***

	Average error rate	
Typology of errors	**Translators**	**Post-editors**
Addition	0.1	0
Omission	0.4	0.3
Mistranslation	1.7	3.2
Over-translation	0	0
Under-translation	0.3	0
Untranslated text	0.3	0
Improperly exact TM match	0	0

A closer analysis of the accuracy category reveals that although the overall average error rate for this category is higher for post-editors, in the individual error types this is only the case for mistranslation, where post-editors had an average error rate of 3.2 and

translators had one of 1.7. Translators reported higher average error rates in four subcategories: addition, omission, under-translation, and untranslated text. The post-editors made errors only in the subcategories of omission and mistranslation. In the accuracy category, the translators showed a more varied error rate.

5.3.3.1.1 Addition

Table 7: ***Examples of errors from the text in the addition subcategory***

Group	**Translators**	**Segment**	**3**
Source	You can compost food scraps indoors using a worm bin!		
Target	Pomocou vermikompostéru môžete kompostovať zvyšky z jedál môžete aj v interiéri!		
Suggested target	Pomocou vermikompostéru môžete kompostovať zvyšky z jedál aj v interiéri!		

When assessing the quality of the translation, there was only one addition error. It was made by a translator in the third segment, and it was a duplication of a word in one sentence rather than added information. The translator had the word *môžete* present twice in the translation when it should have been present only once. This error was presumably made out of carelessness, when the translator changed the word order and forgot to delete the previously used word. This type of error was not observed with the post-editors.

5.3.3.1.2 Omission

Table 8: ***Examples of errors from the text in the omission subcategory***

Group	**Translators**	**Segment**	**8**
Source	Dimensions should be approximately 12 x 12 x 12 inches (one cubic foot) but do not have to be precise.		
Target	Rozmery by mali byť približne 30 x 30 x 30 cm, no nemusí sa to presne zhodovať.		
Suggested target	Rozmery by mali byť približne 30 x 30 x 30 cm (s objemom zhruba 28 litrov), no nemusí sa to presne zhodovať. (Dimensions should be approximately 30 x 30 x 30 cm (with a volume of roughly 28 liters), but do not have to be precise.)		
Group	**Post-editors**	**Segment**	**15**
Source	These holes will provide oxygen to the worms and other decomposer organisms in the bin.		

Target	Tieto otvory poskytnú kyslík dážďovkám a iným organizmom v koši.
Suggested target	Tieto otvory poskytnú kyslík dážďovkám a rozkladacím organizmom v nádobe.

When assessing the quality of the translation, there were four errors in this subcategory that the translators made; three of them are shown in the selected examples. Some translators omitted the information in the parenthesis in the eighth segment. Since the dimensions were already listed, they probably figured that the volume was redundant information. One must be careful with numbers in manuals; it is not advisable to remove, for example, the volume of the container (as happened in this case) without consulting the client. For post-editors, there were three errors in this subcategory; in addition to the same errors made by the translators, one post-editor omitted the translation of the term "decomposer," which was an error also made by some of the translators. In the omission category, the translators and post-editors differed little from each other and made the same errors.

5.3.3.1.3 Mistranslation

Table 9: ***Examples of errors in the mistranslation subcategory***

Group	**Translators**	**Segment**	**10**
Source	Fine screen to keep out pests (optional)		
Target	Jemné sito proti škodcom (nepovinné)		
Suggested target	Jemná sieťka proti škodcom (nepovinné)		
Group	**Post-editors**	**Segment**	**12**
Source	\{b\>Tray (optional)\<b\}		
Target	\{b\>Zásobník (voliteľný)\<b\}		
Suggested target	\{b\>Podnos (voliteľný)\<b\}		

In the accuracy category, translators made the most errors of the mistranslation type. Most errors were mainly related to the mistranslation of the term "screen" in the context of the vermicomposter manual. Many translators translated it as *sito* (sieve), and others translated it as *plocha* (surface), *triedič* (sorter), or *filter* (filter); it was supposed to be *sieťka* (screen) to refer to an insect screen that is meant to serve as a protection against pests. The post-editors also had problems translating some of the terms. While

translators could not deal with "screen," there was a more widespread problem among post-editors. Many words were mistranslated in context; "bin" was translated as *kôš* (basket), but the meaning was *nádoba* (bin) or *vermikompostér* (vermicomposter, depending on the context). "Directions" was translated as *smer* (meaning "direction" but referring to the orientational meaning of the word) when it was meant to be "instructions" or a procedure on how to make a vermicomposter. Last but not least, the example above shows that "tray" was translated as *zásobník* (bin) when it was supposed to be *tácka* or *podnos* (meaning "tray") as something used to catch worm tea from the vermicomposter. Most of these mistranslations were caused by raw machine translations that were not corrected by the post-editors.

5.3.3.1.4 Under-translation

Table 10: ***Examples of errors in the under-translation subcategory***

Group	**Translators**	**Segment**	**26**
Source	Worms can tolerate temperatures a bit beyond this range, but they will be less active.		
Target	Dážďovky dokážu znášať aj teploty mierne pod 13°C, no v tom prípade nebudú také aktívne.		
Suggested target	Dážďovky dokážu znášať aj teploty mierne mimo tento rozsah, no v tom prípade nebudú také aktívne.		

Errors in the under-translation category were only made by the translators. This is because the post-editors worked with a machine translation that translated every piece of information in the text. In the above example, the translator erroneously only pointed out that earthworms can tolerate lower temperatures than recommended, omitting the information that earthworms can also tolerate temperatures slightly above the highest recommended value.

5.3.3.1.5 Untranslated text

Table 11: ***Examples of errors in the untranslated text subcategory***

Group	**Translators**	**Segment**	**24**
Source	Worms prefer temperatures between 55°F and 80°F (13°C and 27°C).		
Target	Dážďovky uprednostňujú teplotu medzi 55 °F a 80 °F (13 °C and 27 °C).		

Suggested target	Dážďovky uprednostňujú teplotu medzi 55 °F a 80 °F (13 °C a 27 °C).

In this subcategory, only the translators made errors because the post-editors were working with a machine translation which translated all the information in the text. It was impossible to find an untranslated part of the text without manually inserting it in. Some translators left untranslated text in the translation, and it was the same error in every case. In the twenty-fourth segment, they did not translate the text that was in parentheses. It is likely that they copied it from the source text without noticing that there was an "and" conjunction. In the final translation, the conjunction remained in the original language.

5.3.3.2 Fluency

The fluency category highlights problems with both the content and the form of the translation. This category encompasses seven subcategories: punctuation, spelling, grammar, grammatical register, inconsistency, link/cross-reference, and character encoding.

Table 12: ***The average error rate in the fluency category***

	Average error rate	
Typology of errors	**Translators**	**Post-editors**
Punctuation	1.9	1.7
Spelling	0.4	0.1
Grammar	0.9	2.5
Grammatical register	0.2	0.1
Inconsistency	0	0
Link/cross-reference	0	0
Character encoding	0	0

The average error rate in the fluency category is higher for post-editors. Looking at the subcategories, only the grammar subcategory showed a higher average error rate for post-editors. Post-editors had an average error rate of 2.4, while translators had an average error rate of only 0.9. In the subcategories of punctuation, spelling, and grammar, translators made more errors. In this category, translators also showed a greater variety of errors.

5.3.3.2.1 Punctuation

Table 13: ***Examples of errors in the punctuation subcategory***

Group	**Translators**	**Segment**	**22**
Source	If you drilled holes in the bottom on the bin, place a tray under the bin to catch any "leachate"—this is a waste product of the composting process made of excess moisture.		
Target	Ak ste do dna nádoby vyvŕtali otvory, umiestnite pod nádobu podnos, aby ste zachytili dážďovkový čaj – odpadový produkt kompostovacieho procesu, ktorý je výsledkom prebytočnej vlhkosti.		
Suggested target	Ak ste do dna nádoby vyvŕtali otvory, umiestnite pod nádobu podnos, aby zachytil dážďovkový čaj – odpadový produkt kompostovacieho procesu, ktorý je výsledkom prebytočnej vlhkosti.		
Group	**Post-editors**	**Segment**	**24**
Source	Worms prefer temperatures between 55°F and 80°F (13°C and 27°C).		
Target	Dážďovky dávajú prednosť teplotám medzi 13 ° C a 27 ° C.		
Suggested target	Dážďovky preferujú teploty medzi 13 °C a 27 °C.		

In addition to the usual problems with commas, the translators also made errors in writing units and hyphens. In the example above, the translator incorrectly used a hyphen instead of a dash. Indeed, every translator who wrote this sentence in the same order as the original text used a hyphen instead of a dash. This error may also have been due to students not being used to checking for dashes. This is because if you type a hyphen in a Word document with spaces around the hyphen, the hyphen automatically changes to a dash; however, this is not how it works in CAT tools.

Both the post-editors and translators had problems with typing commas and dashes. No one from the post-editors group used the dash correctly either. Clearly punctuation is a problem for students, but this is something a proofreader can fix when revising. In addition to the correct writing of hyphens, the post-editors had a problem with the correct writing of units. In the example above, there is an incorrect notation of the Celsius unit which was caused by the machine translation and was not corrected.

Pavlíková, Diana. 2022. Comparing Human Translation and Machine Translation Post-Editing. In: L10N Journal. 1(1), pp. 65–100.

5.3.3.2.2 Spelling

Table 14: ***Examples of errors in the spelling subcategory***

Group	**Translators**	**Segment**	**10**
Source	Fine screen to keep out pests (optional)		
Target	Tekná plocha, ktorá zabráni prístupu škodcom (nepovinné)		
Suggested target	Jemná sieťka, ktorá zabráni prístupu škodcom (nepovinné)		
Group	**Post-editors**	**Segment**	**15**
Source	These holes will provide oxygen to the worms and other decomposer organisms in the bin.		
Target	Tieto otvory poskytnú kyslík dážďovkám a iným heteretrofným organizmom vo vermikompostéri.		
Suggested target	Tieto otvory poskytnú kyslík dážďovkám a iným heterotrofným organizmom vo vermikompostéri.		

The spelling subcategory included errors related to the incorrect spelling of words, such as typos, the incorrect spelling of i/y, and the incorrect change of prepositions in assimilation. In this case, it was mainly related to typos, which were more often made by translators. In the example above, there is the misspelled word *tenká* (thin), which could have been corrected by a spellcheck. Post-editors made only one typo since the raw machine translation did not contain any. The word *heterotrofný* (decomposer) was misspelled by a post-editor when transcribing the machine translation.

5.3.3.2.3 Grammar

Grammar encompasses syntax and morphology. Grammatical errors include the incorrect use of cases, a wrong inflectional form, and incorrect sentence construction. Within this subcategory, there was a check to see if the words were from the standardized variety of the language. Grammatical errors were mainly associated with incorrect prepositional phrases and the use of non-standard words or expressions. In the given example, there is the non-standardized form of the word *zohnať* (purchase). In other cases, translators used phrases that are considered to be Czechisms in Slovak, such as the phrase *o veľkosti* instead of the phrase *s veľkosťou*.

Pavlíková, Diana. 2022. Comparing Human Translation and Machine Translation Post-Editing. In: L10N Journal. 1(1), pp. 65–100.

Table 15: ***Examples of errors in the grammar subcategory***

Group	**Translators**	**Segment**	**11**
Source	Purchase at a hardware store.		
Target	Zohnáte ho v železiarstve.		
Suggested target	Zoženiete ho v železiarstve.		
Group	**Post-editors**	**Segment**	**22**
Source	If you drilled holes in the bottom on the bin, place a tray under the bin to catch any “leachate”—this is a waste product of the composting process made of excess moisture.		
Target	Ak ste do spodnej časti koša vyvŕtali otvory, umiestnite pod kôš podnos, aby zachytili všetok „dážďovkový čaj“ – ide o odpadový produkt z procesu kompostovania vyrobený z prebytočnej vlhkosti.		
Suggested target	Ak ste do dna nádoby vyvŕtali otvory, umiestnite pod nádobu podnos, aby zachytil dážďovkový čaj – odpadový produkt kompostovacieho procesu, ktorý je výsledkom prebytočnej vlhkosti.		

Post-editors made more grammatical errors compared to translators; however, most of them were related to the raw machine translations, which had to be edited to correct the grammatical errors.

5.3.3.2.4 Grammatical register

Table 16: ***Examples of errors in the grammatical register subcategory***

Group	**Translators**	**Segment**	**14**
Source	Drill at least 10 quarter-inch holes in the lid.		
Target	Do veka vyvŕtajte minimálne 10 otvorov s rozmermi 0,6 centimetra.		
Suggested target	Do veka vyvŕtajte minimálne 10 otvorov s priemerom 0,6 centimetra.		
Group	**Post-editors**	**Segment**	**14**
Source	Drill at least 10 quarter-inch holes in the lid.		
Target	Do veka vyvŕtajte najmenej 10 otvorov s rozmerom približne 6mm.		
Suggested target	Do veka vyvŕtajte najmenej 10 otvorov s priemerom približne 6 mm.		

In the grammatical register subcategory, translators and post-editors made the same error. Instead of the term *priemer* (diameter), they chose the term *rozmer* (dimension) to express the diameter of the openings; this could confuse the reader of the manual.

5.3.3.3 Terminology

The terminology category encompasses two types of errors: inconsistency with the termbase and an inconsistent use of terminology.

Table 17: ***The average error rate in the terminology category***

	Average error rate	
Typology of errors	**Translators**	**Post-editors**
Inconsistent with termbase	0	0.2
Inconsistent use of terminology	0.7	0.7

Based on the table of error types in the terminology category, it can be seen that in the inconsistent use of terminology, translators and post-editors performed similarly with an average error rate of 0.7. In terms of being inconsistent with the termbase, there were only errors in the group of post-editors, meaning that they did not sufficiently correct the machine translation to make the terminology match the provided termbase.

5.3.3.3.1 Inconsistencies with the termbase

Table 18: ***Examples of errors in the inconsistent with termbase subcategory***

Group	**Post-editors**	**Segment**	**24**
Source	Worms prefer temperatures between 55°F and 80°F (13°C and 27°C).		
Target	Červy dávajú prednosť teplotám medzi 13 °C a 27 °C.		
Suggested target	Dážďovky dávajú prednosť teplotám medzi 13 °C a 27 °C.		

Only the post-editors had inconsistencies with the termbase. The two errors noted concerned the term "worm," which was incorrectly left as *červ* when, according to the termbase, it should have been translated as *dážďovka*. Again, this was an error caused by machine translation which was not corrected by the post-editors, who either allowed themselves to be influenced by it or did not correct the error out of inattention or a lack of final proofing.

5.3.3.3.2 An inconsistent use of terminology

Table 19: ***Examples of errors in the inconsistent use of terminology subcategory***

Group	**Translators**	**Segment**	**23**
Source	After you make or buy a bin, store it in a good location for both you and the worms.		
Target	Keď už ste hrdým majiteľom zakúpeného alebo vyrobeného domáceho kompostéra, umiestnite ho na miesto, ktoré vyhovuje vám ale je vhodné aj pre dážďovky.		
Suggested target	Keď už ste hrdým majiteľom zakúpeného alebo vyrobeného domáceho vermikompostéra, umiestnite ho na miesto, ktoré vyhovuje vám, ale je vhodné aj pre dážďovky.		
Group	**Post-editors**	**Segment**	**20**
Source	Use non-toxic glue to attach fine screen over holes (optional).		
Target	Pomocou netoxického lepidla pripevnite jemné sito cez otvory (voliteľné).		
Suggested target	Pomocou netoxického lepidla pripevnite jemnú sieťku cez otvory (voliteľné).		

In the inconsistent use of terminology subcategory, the translators had a particular problem with the term *vermikompostér* (vermicomposter), where they incorrectly used the term *kompostér* (composter) instead. In addition to the same inconsistencies the translators made, the post-editors also had problems with the terms "screen" and "tray." In the example above, the post-editor translated "screen" as *sito* even though in previous segments it had been translated as *obrazovka*. This inconsistency was already present in the machine translation and was not corrected by the post-editor.

5.3.3.4 Style

The style category highlights stylistic problems. This category encompasses five types of errors: awkward, company style, inconsistent style, third-party style, and unidiomatic. It was found that the only error was in the inconsistent style subcategory and had been made by a translator.

Pavlíková, Diana. 2022. Comparing Human Translation and Machine Translation Post-Editing. In: L10N Journal. 1(1), pp. 65–100.

Table 20: ***The average error rate in the style category***

	Average error rate	
Typology of errors	**Translators**	**Post-editors**
Awkward	0	0
Company style	0	0
Inconsistent style	0.1	0
Third-party style	0	0
Unidiomatic	0	0

5.3.3.4.1 Inconsistent style

Table 21: ***Examples of errors in the inconsistent style subcategory***

Group	**Translators**	**Segment**	**20**
Source	Use non-toxic glue to attach fine screen over holes (optional).		
Target	Jemný triedič prelepíte cez diery netoxickým lepidlom (nepovinné).		
Suggested target	Jemnú sieťku prelepte cez diery netoxickým lepidlom (nepovinné).		

In this subcategory, there was one error. The translator had written the verb *prelepiť* in the indicative form and not in the imperative form as is used in the rest of the translation and as would be appropriate in a text with instructions.

5.3.3.5 Design

The design category highlights problems with formatting. This category encompasses five subcategories: length, local formatting, markup, missing text, and truncation/text expansion. Post-editors had a higher average error rate, but only made errors in the markup subcategory; translators again varied more in error types, and, in addition to the markup subcategory, they also made errors in the local formatting subcategory.

Pavlíková, Diana. 2022. Comparing Human Translation and Machine Translation Post-Editing. In: L10N Journal. 1(1), pp. 65–100.

Table 22: ***The average error rate in the design category***

	Average error rate	
Typology of errors	**Translators**	**Post-editors**
Length	0	0
Local formatting	0.1	0
Markup	0.1	0.9
Missing text	0	0
Truncation/text expansion	0	0

5.3.3.5.1 Local formatting

Table 23: ***Examples of errors in the local formatting subcategory***

Group	**Translators**	**Segment**	**1**
Source	{1>INDOOR COMPOSTING<1} with a worm bin		
Target	Interiérové kompostovanie s vermikompostérom		
Suggested target	{1>INTERIÉROVÉ KOMPOSTOVANIE<1} s vermikompostérom		

Within the local formatting subcategory, only the translators made errors. One translator did not follow the formatting of the source text and wrote the term *interiérové kompostovanie* in the lower case even though it was originally written in the upper case.

5.3.3.5.2 Markup

In the markup subcategory, there was only one error that the translators made; one translator omitted tags in the first segment altogether. Post-editors had a bigger problem with tags. In addition to omitting tags altogether, there was incorrect tag notation with redundant spaces. For the post-editors, these errors in tags were already present in the machine translation.

Table 24: ***Examples of errors in the markup subcategory***

Group	**Translators**	**Segment**	**1**
Source	{1>INDOOR COMPOSTING<1} with a worm bin		
Target	Interiérové kompostovanie s vermikompostérom		
Suggested target	{1>INTERIÉROVÉ KOMPOSTOVANIE<1} s vermikompostérom		
Group	**Post-editors**	**Segment**	**1**
Source	\{1\>INDOOR COMPOSTING\<1\} with a worm bin		
Target	\{1\> VNÚTORNÉ KOMPOSTOVANIE \<1\} s vermikompostérom		
Suggested target	\{1\>VNÚTORNÉ KOMPOSTOVANIE\<1\} s vermikompostérom		

5.3.3.6 Locale convention

The locale convention category highlights cases where the translation does not conform to country-specific conventions. This category encompasses six subcategories: address format, date format, currency format, measurement format, shortcut key, and telephone format.

Table 25: ***The average error rate in the locale convention category***

Typology of errors	**Average error rate**	
	Translators	**Post-editors**
Address format	0	0
Date format	0	0
Currency format	0	0
Measurement format	0.6	0.3
Shortcut key	0	0
Telephone format	0	0

In the locale convention category, translators and post-editors only made errors in the subcategory of measurement format; translators showed a higher error rate.

5.3.3.6.1 Measurement format

The translators had more trouble with the localization of units than the post-editors. Sometimes they decided not to localize the units, which meant that readers of the

manual in the Slovak market would have to convert the units themselves in order to assemble the vermicomposter. In addition, the translators made errors in the actual conversion of the units. This was also the case in the given example, where the translator converted the data but, probably due to inattention, used the wrong unit prefix to the numeric value, thus using centimeters instead of millimeters.

Table 26: ***Examples of errors in the measurement format subcategory***

Group	**Translators**	**Segment**	**14**
Source	Drill at least 10 quarter-inch holes in the lid.		
Target	Do veka vyvŕtajte najmenej desať 6 centimetrových otvorov.		
Suggested target	Do veka vyvŕtajte najmenej desať 6-milimetrových otvorov.		
Group	**Post-editors**	**Segment**	**8**
Source	Dimensions should be approximately 12 x 12 x 12 inches (one cubic foot) but do not have to be precise.		
Target	\{i\>Rozmery by mali byť približne 12 x 12 x 12 palcov (jedna kubická stopa), ale nemusia byť presné.\<i\}		
Suggested target	\{i\>Rozmery by mali byť približne 30 x 30 x 30 cm (s objemom zhruba 28 litrov), ale nemusia byť presné.\<i\}		

Although post-editors had a lower average error rate in this subcategory, some in this group chose not to convert units at all.

To conclude the qualitative analysis, there are three research questions that need to be answered. Within the translation error rate, there was an average error rate of 10 for post-editors and 7.8 for translators. The second question asked which group produced more correct translations. Since the average error rate was directly related to penalty points, more correct translations were produced by the translators. On average, translators had 30.2 penalty points and post-editors had 38.8. The last research question asked what types of errors were made by the translators and what types of errors were made by the post-editors. After a closer analysis of each error category, the post-editors on average made more errors in the categories of accuracy, fluency, terminology, and design. Translators performed worse in the style and locale convention categories. When analyzing the subcategories, the translators also made errors in subcategories where the post-editors did not make any.

5.4 Temporal analysis

In addition to the number and types of errors, the temporal aspect of the translation and post-editing effort was looked at. This aspect can be easily measured and compared. Translators and post-editors were informed in advance that the time spent on the translation would be measured and that they should therefore only work on this activity and hand in the final document once they were satisfied with it.

Table 27: ***The average time spent on translation and post-editing***

Group	***Average time spent on translation/post-editing***
Translators	36.8 minutes
Post-editors	25.9 minutes

Looking at the average times of the translator and post-editors, it is clear that the post-editors generally performed the post-editing faster. Compared to the translators, they had raw machine translation; this probably sped up their work and made them more efficient with an average time of 25.9 minutes; however, it should be noted that the translators' speed for the length of the text was also good, given that they were able to translate it in 36.8 minutes on average. In conclusion, the given sample showed that the post-editors were faster. This finding also answers the last research question regarding the temporal analysis. By analyzing the measured times, it seems that the post-editors were more time-efficient compared to the translators.

5.5 Discussion

After presenting the results, it is important to consider why they occurred and whether they could have been predicted. A qualitative analysis first looked at the average error rate and translation correctness in both groups. Since the error rate was directly related to translation correctness through the penalty point system, it seems that in both cases the post-editors performed worse. To find out why this was the case, it was necessary to look at the error rate analysis within the typology of errors. This revealed which errors were made most often by post-editors and by translators and what might have been the reason for this. When analyzing the typology of errors, there were six main categories: accuracy, fluency, terminology, style, design, and locale convention. In four of the six categories, the post-editors had a higher average error rate; these were in the categories of accuracy, fluency, terminology, and design. This means that the translators performed worse only in the categories of style and locale convention. Each

category also had subcategories of error types, which further revealed that although the post-editors made errors more frequently, translators showed a greater variety of errors.

Of the five analyzed subcategories concerning accuracy, the post-editors on average were more likely to make errors only in the subcategory of mistranslation. The post-editors made errors in one subcategory, and so they made the same errors which were influenced by the machine translation. Many of these were because the post-editors did not correct the errors created by the machine translation. Either they relied on the machine translation more than they should have, or they simply did not notice them there. The biggest problem they had was with mistranslated words, which, according to the context, did not fit the text. In terms of under-translation or untranslated text, only the translators made errors because the machine translation had all the words translated and omitted no information in the output. In the first category alone, the machine translation had a clear impact on the types of errors that the post-editors made; in some ways, it may even have confused them or they may have trusted it too much and thus not delivered a translation of the highest quality.

The same analysis of the results could be applied to the fluency category, where the post-editors also showed a higher average error rate; they appear to have been influenced by the machine translation. The biggest difference can be seen between the spelling and grammar subcategories. In the spelling subcategory, there was a higher error rate for translators because machine translation does not produce typos in the text. This is a purely human error which occurred for translators as well as for post-editors who decided to overwrite a certain segment in the raw output of the machine translation; however, the post-editors showed a higher average error rate in the grammar subcategory and there was the largest difference between the groups. The raw machine translation had a few grammatical errors which several post-editors did not identify. They either did not notice them or trusted the machine translation more than they should have; importantly, grammatical correctness is one of the most important aspects of a good translation.

Within the terminology category, the largest difference in error rates was in the inconsistent with termbase subcategory. Only the post-editors made any errors. This was influenced by the machine translation, which translated the terms incorrectly or differently from the terms in the glossary; the post-editors had to unify these mistranslations with the termbase. In the final translation, there was an inconsistent use of terminology, which the post-editors did not correct in the output of the machine translation, as

well as inconsistencies with the termbase, which were not present at all in the translators' texts. There was only one error in the style category, and this was in the translators' group. This was probably an error that the translator did not check. The post-editors had no problem with the style; apparently the output of the machine translation was stylistically acceptable.

Within the design subcategory, translators again showed a more varied error rate; however, the post-editors had a higher overall error rate, especially for the tags (markup) subcategory. The post-editors had more trouble with the tags because the machine translation caused the output to have multiple gaps in the tags and even omitted tags altogether in some segments. Several post-editors did not notice these changes and did not correct them. They were again affected by the output of the machine translation, and so the average error rate increased. In the locale convention category, there was only one subcategory (measurement format) where translators and post-editors both made errors. Some translators chose not to localize the units of measurement and used them in foreign values; by contrast, this happened only once in the post-editors' group. The machine translation in one segment chose not to give temperatures in both Fahrenheit and Celsius, only using the Celsius units. This segment may have prompted the post-editors to localize the units in other segments as well, thus unifying the final translation. The translators had a choice whether to localize the units or not, and this choice resulted in cases where the translators did not localize them.

To summarize the qualitative analysis, it seems that the post-editors had a higher average error rate and thus less accurate translations; however, based on the specific examples of errors, they were significantly influenced by the machine translation. It had certain types of errors, which the post-editors had to watch out for, but it did not have other types of errors. After all, machine translation rarely omits parts of a text, adds information to a translation, or makes typos. It can be said that post-editors should mainly focus on correcting tags, inconsistencies in the text (especially the terms and words used), grammatical errors, and mistranslations of words and terms produced by the machine translation.

The temporal analysis aimed to determine which group produced the translation faster. With an average time of 25.9 minutes, the post-editors were faster. Translators had an average time of 36.8 minutes. This result was predictable, and several articles (Krings 2001; Tatsumi 2010) have reported that machine translation post-editing is faster and thus more efficient than human translation. This experiment confirmed this statement.

It should be added, however, that this research has some limitations, especially regarding the evaluation of the quality of the translation, which was carried out only by the present author. Despite efforts at objectivity, subjectivity must have been reflected in the results to some extent for obvious reasons. To achieve greater objectivity in future research, it would be advisable to use more than one person to evaluate the quality of the translation, or at least a person who is not directly involved in the research, has not prepared the text in advance, and who is unaware of the knowledge of the translators and post-editors involved.

Conclusion

This article has discussed the comparison between human translation and machine translation post-editing. The objective of the research was to compare human translation with machine translation post-editing on a given sample of students and to find out which process was better in terms of speed of work and quality of the final translation. To meet this goal, a qualitative and temporal analysis was conducted. Based on the results of both analyses, it can be evaluated that machine translation post-editing has its advantages. It has been shown to be more time-efficient and useful in eliminating some types of errors such as typos, omissions, and additions of information. On the other hand, the machine translation still contains errors and so post-editing is essential. Errors are often repeated, and a trained post-editor should be able to correct errors such as inconsistent terminology.

Students without in-depth training in machine translation post-editing produce lower quality post-edited translations than students who have had several years of experience with human translation during their studies. The error rate is directly linked to the unfamiliarity of working with machine translation. Teaching post-editing could eliminate unnecessary errors in the categories of grammatical correctness, inconsistency, and the translation of terms. Students could also learn how to approach the output of the machine translation correctly and how to perform post-editing more efficiently, so that it is not only useful in terms of time but also in terms of quality.

Bibliography

Absolon, Jakub. 2018. Strojový preklad a posteditovanie. Nitra: Univerzita Konštantína Filozofa.

Allen, Jeffrey. 2003 Post-editing. In: Computers and Translation: A translator's guide. Amsterdam: John Benjamins Publishing Company, pp. 297–318.

Pavlíková, Diana. 2022. Comparing Human Translation and Machine Translation Post-Editing. In: L10N Journal. 1(1), pp. 65–100.

Bubnic, Mauro. 2022. Meeting the Needs of End Users with AI-based NLP Tools. In: Translating Europe Forum 2022. https://www.youtube.com/watch?v=CXOSCjkcCNc&list=PLLqI-RaiVCGCSrlTMx1iG1pRlyYR1ZKSyi&index=15. Accessed on: 11 December 2022.

Common Sense Advisory. 2016. Global Market Research Firm Common Sense Advisory Finds Post-edited Machine Translation (PEMT) Among Fastest-growing Segments of the Language Industry. https://csa-research.com/Blogs-Events/CSA-in-the-Media/Press-Releases/PEMT-Among-Fastest-growing-Segments-of-the-Language-Industry. Accessed on: 3 December 2022.

Dengová, Stanislava. 2020. 3 most popular CAT tools compared. https://www.lexika-translations.com/blog/comparison-of-cat-tools/. Accessed on: 17 January 2022.

Doherty, Stephen a Gaspari, Federico. 2013. Effective Post-Editing in Human & Machine Translation Workflows: Critical Knowledge & Techniques. Dublin: Dublin City University.

Finnegan, Noel. 2018. Linguistic Quality Assessment: How Do You Measure Translation Quality? https://www.argosmultilingual.com/blog/linguistic-quality-assessment-how-do-you-measure-translation-quality. Accessed on: 23 January 2022.

Frívaldský, Dalibor. 2020. Memsource Translate: Interview with Dalibor Frívaldský Memsource Chief Technology Officer. https://www.memsource.com/blog/memsource-translate-interview-with-chief-technology-officer/. Accessed on: 17 January 2022.

Gene, Viveta. 2019. Machine Translation: From Translation to Post-editing. https://www.gala-global.org/events/events-calendar/machine-translation-translation-post-editing. Accessed on: 16 January 2022.

Green, Spence; Heer, Jeffrey a Manning, Christopher D. 2013. The efficacy of human post-Editing for language translation. In: Proceedings of the SIGCHI Conference on Human Factors in Computing Systems (ACM). Association for Computing Machinery, pp. 439–448.

Guerberof Arenas, Ana. 2010. Project management and machine translation. In: Multilingual: 21(1) pp. 1–4.

Hu, Ke a Cadwell, Patrick. 2016. A Comparative Study of Post-editing Guidelines. In: Baltic J. Modern Computing: 4(2) pp. 346–353.

Hudecová, Elena et al. 2021. Porovnanie strojového, posteditovaného a ľudského prekladu technickej dokumentácie zo slovenčiny do nemčiny. In: Slovenská Reč: 86(2) pp. 192–207.

Hutchins, W. John a Somers, Harold L. 1992. An Introduction to Machine Translation. San Diego: Academic Press.

Indoor Composting with a Worn Bin. New York: NYC Department of Sanitation. https://www1.nyc.gov/assets/dsny/downloads/pdf/promotional-materials/indoor-worm-bin-composting-brochure-06340-f.pdf. Accessed on: 24 January 2022.

Kabát, Marián. 2022. Všeobecná štylistická príručka pre lokalizáciu softvérových produktov. Bratislava: Stimul.

Korkas, Vassilis. What is linguistic quality assurance and how it works in localization. https://lexiqa.net/what-is-linguistic-quality-assurance-and-how-it-works-in-localization/. Accessed on: 23 January 2022.

Krings, P. Hans. 2001. Repairing Texts. Empirical Investigations of Machine Translation Post-Editing Processes. Kent: The Kent State University Press.

Läubli, Samuel et al. 2019. Post-editing Productivity with Neural Machine Translation: An Empirical Assessment of Speed and Quality in the Banking and Finance Domain. In: Proceedings of Machine Translation Summit XVII Volume 1: Research Track. Dublin: European Association for Machine Translation, pp. 267–272.

Pavlíková, Diana. 2022. Comparing Human Translation and Machine Translation Post-Editing. In: L10N Journal. 1(1), pp. 65–100.

Mesa-Lao, Bartolomé. 2013. Introduction to post-editing. In: SEECAT – Speech & Eye-Tracking Enabled CAT. Copenhagen: Copenhagen Business School.

Munková, Daša. 2013. Evalvácia strojového prekladu. Nitra: Univerzita Konštantína Filozofa.

O'Brien, Sharon. 2002. Teaching Post-editing: A Proposal for Course Content. In: European Association for Machine Translation Conferences/Workshops. Dublin: Dublin City University.

Pavlíková, Diana. 2021. Porovnanie strojového a humánneho prekladu terminológie. In: Prekladateľské listy 10. Bratislava: Univerzita Komenského, pp. 64–77

Plitt, Mirko a Masselot, François. 2010. A productivity test of statistical machine translation postediting in a typical localization context. In: The Prague Bulletin of Mathematical Linguistics, pp. 7–16.

Robert, Anne-Marie. 2013. Vous avez dit post-éditrice? Quelques éléments d'un parcours personnel. In: The Journal of Specialised Translation: 19(1) pp. 29–40.

Sánchez-Gijón, Pilar; Moorkens, Joss a Way, Andy. 2019. Post-editing neural machine translation versus translation memory segments. In: Machine Translation: 33(1), pp. 31–59.

Tatsumi, Midori. 2009. Correlation between automatic evaluation metric scores, post-editing speed and some other factors. In: MT Summit XII – The twelfth Machine Translation Summit International Association for Machine Translation hosted by the Association for Machine Translation in the Americas. Association for Machine Translation in the Americas.

Tatsumi, Midori. 2010. Post-Editing Machine Translated Text in A Commercial Setting: Observation and Statistical Analysis. Dublin: Dublin City University.

TAUS. 2010. Post-editing in practice. https://www.taus.net/think-tank/reports/postedit-reports/postediting-in-practice. Accessed on: 23 January 2022.

TAUS. 2016. Machine Translation Post-editing Guidelines. https://info.taus.net/mt-post-editing-guidelines. Accessed on: 23 January 2022.

TAUS. Start tracking errors with DQF-MQM. https://info.taus.net/dqf-mqm-error-typology-templ. Accessed on: 23 January 2022.

Torrejón, Enrique a Rico, Celia. 2012. Skills and Profile of the New Role of the Translatoras MT Post-editor. In: Revista Tradumatica: 10(1), pp. 166–178.

Wu, Yonghui et al. 2016. Google's Neural Machine Translation System: Bridging the Gap between Human and Machine Translation. Ithaca: Cornell University.