Product and Process Analysis of Machine Translation into the Inflectional Language

Dasa Munkova1, Michal Munk1, Katarina Welnitzova1, and Johanna Jakabovicova2

Abstract
This study focuses on the influence of quality of Machine Translation (MT) output on a translator’s performance. We analyze the translator’s effort by product analysis and process analysis. The product analysis consists of MT quality evaluation according to the Dynamic Quality Framework; using error typology and the criteria such as fluency and adequacy. We examine translator’s effort from the point of view of typing time, in the context of MT quality—focusing on error rate in language, accuracy, terminology, and style, and also in fluency and adequacy to the source text. We have found that the translator’s performance is influenced by MT quality. The typing time is very closely related to errors in language, accuracy, terminology, and style as well as to fluency and adequacy. We used the Mann-Whitney test to compare the productivity of post-editing of MT with human translation. The results of the study have shown that post-editing—compared to human translation of journalistic text from English into the inflectional Slovak language is more effective.

Keywords
typing time, machine translation, translation quality, productivity, post-editing

Introduction
As the need for communication and transmission of information is globally increasing, the use of technological advances is more urgent than ever before. The tasks which were exclusively based on human thinking and intelligence are gradually starting to be operated by machines. As Absolon et al. (2018) remark, we can see a parallel in the process of translation. The evidence might be seen in Quah’s (2006) specification of applied translation studies map devised by Holmes (1975). Originally, the map was based on three areas—translator training, translation aids/devices and translation criticism. Lately, the map was extended with translation tools—automatic tools and computer-assisted translation tools. In translation, there exist several technological innovations using computer-aided translation tools and machine translation systems. These innovations have significantly influenced and changed the way people communicate (Gromová & Múglová, 2012). To use them as functional tools in today’s dynamically developing multilingual society, their products (outputs) must be systematically evaluated. Generally, machine translation output is not expected to be ‘flawless’. It aims to be comprehensible in the form of the information, text assimilation (understanding the content and the meaning of the translation); communication (translations of e-mails, text messages), discussion fora (chatting); internet markets; or with the specific intervention of a translator (a high-quality translation for publication). While computer-aided translation tools—due to their productivity and consistency of translation—are considered to be part of a translator’s work, lately, machine translation systems have also been accepted by professional translators and scholars (Bowker, 2019; Vieira & Alonso, 2020; Way, 2018). However, many translators are still adopting the changes brought by translation technologies to the translation industry, and to the translation process itself. The fact that machine translation systems are gradually improving (despite common arguments between translators about whether the use of machine translation is more helpful or harmful), is undeniable. As for machine translation, the man’s help—human intervention—inevitable so that the target text is comprehensive, fluent, and without grammatical or other mistakes. Generally, a translator’s intervention in the text within the translation process is primarily performed during the quality control of the target text in the form of post-editing of the target text. However, translator’s intervention may also be performed at the beginning of the translation process through pre-editing of the source text (by an improvement of text’s comprehension or by reduction of polysemantic words), but it has not been in practice (Hirooka & Yamada, 2019).

1Constantine the Philosopher University in Nitra, Slovakia
2Slovak University of Agriculture in Nitra, Slovakia

Corresponding Author:
Dasa Munkova, Constantine the Philosopher University in Nitra, Tr. A. Hlinku 1, Nitra 949 74, Slovakia.
Email: dmunkova@ukf.sk
According to the extent of intervention, Baker and Saldanha (2009) distinguish two types of machine translation: fully automated machine translation (without human intervention) and supported machine translation (with a degree of human intervention), including machine-assisted human translation and human-assisted machine translation (interactive machine translation).

Machine translation (MT) systems are essentially developed from human translations. Nowadays, the systems contain millions of sentences translated by humans, and the “machines” are trained on these translations (large data sets) and learnt the best match or pattern for machine translation. The MT systems are being constantly improved—increasing the quality of translation to make translator’s work more efficient. Their architecture is ceaselessly being further developed, moreover, they are equipped with an increasing number of high-quality translations.

The aim of our study is twofold. The first objective—based on editing time—aims to examine the quality of machine translation, that is, how the post-editor’s effort measured by typing time (pressing keys on a keyboard) is related to the quality of machine translation (MT quality was measured according to the TAUS Dynamic Quality Framework, DQF). According to Görög (2014), the objective of DQF is an increase in satisfied customers and a more credible translation quality assurance in the translation industry. DQF is underpinned by the recognition that quality is when the customer is satisfied (Görög, 2014, p. 157).

The second objective—based on the typing time—aims to find out whether the post-editing of machine translation (PE) is more effective than human translation “from scratch” (HT) and also to identify the errors of MT regarding language, accuracy, terminology, and style of journalistic texts. Nowadays, such a genre is one of the most common to be translated by MT.

We examine machine translation from English into Slovak—languages with different language typologies: English as an analytical type and Slovak as an inflectional type of language with a large number of morphological forms (including grammatical polysemy and homonymy), which often cause inadequate semantic and syntactic analyses and transfers in machine translation.

**Post-editing**

Post-editing (PE) could be defined as a task involving editing, modifying or correcting a text which has already been translated by a MT system (Munková, 2013).

Based on the purpose of MT output (for what purpose it will be used), it is essential to decide whether MT output requires PE (or not) and to what extent. Allen (2003) distinguishes three PE types: no PE, light or minimum PE, and full PE.

PE is a process which follows certain rules and procedures so that the result can reach the required quality. PE guidelines provided by the Translation Automation Users Society (TAUS) in 2016, include the instructions on how to post-edit effectively, productivity tests, and on how to analyze different kinds of PE (due to the price).

Two essential factors influence the PE process itself. The first factor includes the quality of MT output, while the second one concentrates on the quality of the target text to be achieved. Accordingly, one of the mentioned approaches—full or light PE—is chosen.

For the light PE, “acceptable” quality of the target text is sufficient. Even here, however, certain rules must be followed. These are mentioned by TAUS as follows:

- Translation should be semantically correct,
- No information should be unintentionally added or removed,
- Offensive, inappropriate or culturally unacceptable content should be altered,
- As much of the original machine translation as possible should be used,
- Basic grammar rules should be followed,
- Purely stylistic errors do not have to be corrected,
- It is not necessary to change sentences and word order solely to make them sound more natural (TAUS, 2016).

It is a type of PE in which post-editors carry out minimal and inevitable changes so that the translation is clear and expresses the same meaning as the original text and/or is congruent in meaning with the original text. However, minimum interference in MT output is relative and it depends on the post-editors. For these reasons, light PE can be extremely difficult (task) for professional translators. When translating “from scratch” (without the help of machines) translators try to avoid grammatical, stylistic or semantic mistakes, and provide a reader with a translation which can be quickly and easily comprehensible. They should only correct the most serious grammatical mistakes, typos, incorrect words, delete unnecessary or incorrect information, or the information accidentally added by the machine. The final text may seem artificial or less natural, but equaling the quality of human translation is not the aim of light PE. The text does not have to be perfect, being adequate is good enough (Densmer, 2014).

Full PE is a far more time-consuming process, and it aims to achieve the translation (quality) equal to human translation. Readers should not notice that they read a post-edited MT output. For readers, they cannot distinguish the translation from the original text. Within full PE, a translator has to correct MT output and to make it stylistically, grammatically, lexically, and semantically correct, and keeping the terminology consistent. MT should also correspond with the original text visually, with a correct formation and facts and cultural references (such as idioms and examples); in other words, it should be as close to human translation as possible in every single aspect (Densmer, 2014).

Although the PE process (either light or full) is reflected by each translator/post-editor differently and depends on...
him/her attitude, some of the basic rules should be followed. In full PE, a translator/post-editor can pay more attention to details which results in a better quality translation. On the contrary, light PE is far more time-saving and brings a clear text of sufficient quality. The decision-making is up to the client who decides which of the two processes to use; under the quality of the original MT output and with the expected quality of the target text.

Koponen (2012a, 2012b) points out that MT can produce texts solid enough to be used in the practice and therefore it can increase the productivity of translators in (at least) some language combinations. Koponen (2015) mentions the survey of Gaspari et al. (2015), which has shown that the absolute majority of 438 respondents working in the translation and localization field use or are likely to use machine translation in the future. About 38% of those who use machine translation (30% of all the respondents) claim that MT output always needs to be post-edited afterward. About 30% of all the respondents do not use PE at all while 32% of them use it to different extents. The survey has thus shown that the use of MT output with PE in the course of translation and localization is on the rise and that the demand for these services is constantly increasing. The authors of the survey expect this trend to carry on in the future. Mariniello and Steiert (2016) discuss the role of a translator in MT. They agree with other experts with the fact that machine translation is still being improved and translators do not need to worry about their jobs since the translation industry cannot completely rely on the machines themselves. Medical and legal MT outputs seem to be the most problematic as even a small mistake in the translation could easily result in the death of the patient or the wrong decision of the court. Since doctors and lawyers need to avoid mistakes in translations, they still need a human translation. Mariniello and Steiert (2016) also mention that the types of translations in which MT has been truly successful are technical texts. These are mostly written clearly and simply, with no ambiguous meanings and a large number of repetitions. This is exactly what suits for MT and it thus makes such texts easier to be post-edited. As a result, MT has become a valuable tool for many companies which need translations of technical documentation. They carry out productivity tests to find out whether PE or human translation is more productive. Plitt and Masselot (2010) performed a productivity test on 12 participants who translated texts from English into French, Italian, German, and Spanish. The test lasted for 2 days and was divided into two phases. In the first phase, the participants translated a text without the help of the MT engine. Plitt and Masselot (2010) obtained a reference value for each of them, and by these references, they were able to compare the results of the second phase dealing with post-editing of machine translation (PEMT). The authors used Moses for MT system. As for PE, they created own system which recorded decision-making and the modification time for each PE sentence. The participants were supposed to do full PE, so the quality of the PEMT had to be equal to the quality of a commonly translated text. The comparison of the quality was carried out by a quality assessment team, who did not know which texts were translated by a human translator and which were post-edited. The results of the productivity test have shown that MT helped to increase the productivity of all the subjects to a varying degree: from 20% up to 131%, giving an average of 74%. This means that machine translation saved 43% of their time. The test has also shown that slower translators benefited the most from MT output. As for modification time itself, the subjects spent 19% of the translation process writing and 10% PE. MT managed to reduce modification time by 70%, which means that if the MT output is appropriate quality, translators do not have to correct anything and they thus spend less time typing. The use of MT and its subsequent PE helps translators gradually increase their productivity. This was confirmed in case of all participants since their performance at PE was much higher than during traditional translating without the help of MT.

Sánchez-Gijón et al. (2019) carried out an experimental test with eight professional EN-ES translators. Participants were asked to answer a questionnaire about their perception of MT quality, to assess the translation quality (30 segments: 15 segments: fuzzy matches under 100% from translation memory [TM] and 15 from MT) and also to post-edit another 30 segments (15 TM and 15 MT). The results have shown that PE involves less editing than TM segments.

Läubli et al. (2019) empirically tested, with four professional translators, how the inclusion of NMT in translation impacts speed and quality of process and product in the professional translation of financial texts in the banking and finance domain. They found out that PE of neural machine translation provides substantial time savings and leads to equal or slightly better quality compare to human translation in language pairs with small amounts of in-domain data for MT system adaptation.

Federico et al. (2012) examined the productivity of 12 translators during PE (with English being the source language and German and Italian being the target languages). The texts for translation were from the legal area and information technology. Under common conditions, translators were allowed to use translation memories during the productivity test. The authors paid their attention to speed and “effort during PE,” that is, the average percentage of the number of words corrected by subjects. This research differed from previous tests in the fact that the translators worked in a “real” environment; they were allowed to use CAT tools which they tend to use in their actual work as well. The results of the research have shown another increase in the performance of each participant during PE. For 10 of the 12 participants, this increase was statistically significant.

O’Brien (2002), also gives an interesting view on the topic. It deals with issues concerning translators, as well as the need for different PE courses. The main question is why translators should learn PE is a demand. Since translation
demand keeps growing every day, a large number of translators use various technical tools including translation memories, terminology management tools or MT technologies and thus they increase their productivity and meet higher demands. If using MT in their work, translators should have a certain amount of knowledge and either full or light PE skills. To put it another way, not every translator is necessarily a good post-editor as well. Some translators have a problem with the fact that their translation is not perfect after light PE. According to O’Brien (2002), some translators have a problem with correcting multiple mistakes which one would normally never make and they are worried about their language skills which may get worse and worse because of permanent contact with weak MT output. Or, they feel unwilling to accept the fact that during PE they are not as free as during translating. This is why novice translators are recommended to take part in PE courses, which would also resolve their aversion toward this matter.

### The Impact of the Source and the Target Language on the Post-editing Process

Differences between the source language and target language in translation represent important factors in MT and its subsequent PE. MT is the most effective in the languages which are close and belong to the same language family (which makes them at least a little similar). However, this is not our case since we have focused on English as the source language and Slovak as the target language. The fact that both of these languages are Indo-European is irrelevant since this is the most widespread language family in the world. Such languages are also divided into further subgroups and there, Slovak and English belong to different categories. English thus belongs to the Germanic languages while Slovak belongs to the Slavic ones.

English and Slovak mainly differ in their prevailing grammatical properties. English is an analytical type of language (including certain synthetic elements) while Slovak is a synthetic one. And this is exactly what makes them different and causes the greatest problems in translation. In the analytical languages, words are mostly created by one morpheme while in the case of the synthetic languages, the words are made up of several morphemes. Therefore, a distinctive feature of synthetic languages is declension. Declension is not typical for English; synthetic characteristics are used only in the several cases, for example, forming the plural of nouns (suffix –s: door-doors, bed-beds) or in the comparison of adjectives (suffixes –er or –est: small-smaller-the smallest, tall-taller-the tallest). On the other hand, declension is typical for Slovak. While Slovak uses case suffixes, in English they are expressed by prepositions, for example,

- A: knihu—a book
- L: (o) knihe—about a book
- I: knihou—with a book

There is also a different word order in English and Slovak. In English, the word order and it uses SVO word order (subject-verb-object); any other word order is incorrect as it would change the meaning of the sentence. In Slovak, there are more possibilities of word order to express the same idea “otec vari večeru,” “večeru vari otec,” or even “otec večeru vari.” Each of the sentences has the same meaning even though SVO word order sounds most natural and clear. However, if we change the word order in English sentence “dad cooks dinner” (“otec vari večeru”) we would get “dinner cooks dad,” which would mean “večera vari otec” in Slovak. By changing the word order, the meaning of the English sentence changes as well, so it is necessary to follow the SVO word order. The model of a simple sentence can be developed by other elements: S (subject) V (verb) O (object) M (manner) P (place) T (time); for example, “I would like to go to Rome next summer”—“Budúce leto by som rád šiel do Ríma.”

Unlike Slovak, English does not distinguish the grammatical gender of nouns and adjectives. Although it recognizes masculine, feminine, and neuter gender, just as Slovak does, they are natural and do not affect grammar. This is mostly reflected in singular verbs in the third person. In Slovak, one could say “muž sedel,” “žena sedela,” or “dievča sedelo” while in English, we use the same verb form for all three cases, that is, “a man was sitting,” “a woman was sitting,” and “a girl was sitting.” This may cause problems especially if the text refers to the people’s names whose sex is unknown or have not been mentioned. As for the sentence “Alex Smith is my best friend,” one would have to find out first whether Alex is a man or a woman and then to translate or post-edit “best friend” correctly into Slovak (“nejlepší priateľ” or “nejlepsia priateľka”).

Problems in English-Slovak translation can be also caused by the understanding of the active and passive voice as the latter is typical for English. In Slovak, however, the active voice is a more natural form, which may cause problems in machine translation as it usually sticks to the original language. “I was confused by his reaction.”—“Jeho reakcia ma zmiala.” instead “Bol som zmätený jeho reakciou.”

Other differences between English and Slovak include vocabulary; in English, it is far wider than in Slovak and rich for synonymy. It is thus very important to be familiar with synonyms in both languages and pay attention to homonyms and words with more than one meaning. This also includes a large number of English tenses, which are replaced by perfective and imperfective verbs in Slovak. Additionally, the conversion is also problematic, which means that the word form “work” in English could be translated as “pracovat” (verb), “práca” (noun), or even “pracovný” (adjective) in Slovak. This may be tricky for MT since it is important to
understand the context and other parts of speech surrounding a particular word.

English and Slovak are two different languages which have just a few features in common. This can result in MT problems, and thus a post-editor has to pay attention to the text so it must be properly translated and understood in every aspect.

The Impact of the Translator/Post-editor Competence on the Post-editing Process

Post-editor’s characteristics and skills are very important for the process of PE. The excellent translators, do not have to be excellent post-editors and vice versa. Many translators do not fit for this job (task) because they do not have the skills necessary for making their PE (and also the use of MT) effective. Good post-editors must be able to evaluate the quality of the MT output, decide very quickly (whether the MT output can be used or not) and also should not be perfectionists in light post-editing strategy. De Almeida and O’Brien (2010) summarized the skills required for a good post-editor as follows:

- He/she should be able to identify the mistakes of machine translation output that need to be dealt with and replace them accordingly (essential changes),
- He/she should be able to do the post-editing fast enough so he/she can meet the deadline that is set for this activity (approximately 5,000 words a day on average),
- He/she should be able to follow the procedure and rules of post-editing to minimize the number of preferential changes—changes that are not necessarily needed and their replacement depends only on the preferences of the translator himself.

The biggest issue with PEMT can be found in an excessive amount of preferential changes. It is not always necessary to apply them and if post-editors decide to, they lose their time and productivity. Post-editors must be able to select and decide if it is necessary to make certain changes. They must be able to use the maximum of raw MT to shorten the time for changes. If they delete the MT output (that was partly usable) and rewrite a whole sentence, they would again lose the time and productivity would decline as well.

PE should be only done by professionals who are familiar with the method and know exactly what, how and how much to edit (need to be edited in the text). The idea of PE loses its importance as the work of post-editors would not be more effective as translators’ who translate traditionally “from scratch.” Also, when evaluating the quality of MT, we cannot take into consideration the personality of translators. The machine does not have human personal attributes and therefore the evaluation of MT quality is specific.

Nowadays, other basic translation skills include technical skills. It means that good translators/post-editors should be able to use various software that help them while translating (e.g., CAT Tools, Microsoft Office), they should know how to use the internet to find or verify information needed in translation. Another advantage is the ability to type fast. All these general IT skills include using a keyboard and mouse, with the most effective way of using them. The more the translators are skilled, the faster they can translate the text they work with or to post-edit the MT output. When referring to PE, it means the more skilled the PC user, the less time for edition is required and the higher productivity of the translator is. The idea of PE loses its effectiveness as translators who translate traditionally “from scratch.” Also, when evaluating the quality of MT, we cannot take into consideration the personality of translators. The machine does not have human personal attributes and therefore the evaluation of MT quality is specific.

Typing Time

The efficiency of MT depends on several factors. For effective PE, it must be carried out in a shorter time than the time required for human translation of the same source text. This also increases the productivity of a translator, who can translate a larger amount of text at the same time. One of the main factors of the translation process is the time that a translator spends “behind the keyboard.” It is the time that involves active writing in the target language (production units), that is, translating and rewriting text from the source language into the target language (it refers to production units). In PE, however, the translators do not rewrite the whole text, they make only the inevitable corrections in the MT output. According to Daems et al. (2017), the PE effort can be assessed through product or process analysis, that is, by comparing MT output to a reference translation, or by observing aspects of the PE process, for example, time or pauses. Krings (2001, pp. 178–182) distinguishes three aspects of PE effort: temporal effort (time to turn the MT output into a high-quality translation, the easiest one to define and measure), technical effort (physical actions in PE including deletions, insertions, substitutions, reordering), and cognitive effort (mental processes and cognitive load in mind during PE).
The temporal effort measured by edit time is defined as a combination of thinking time (cognitive effort) and typing time (technical effort). They are of different length as post-editors usually need more time for thinking than for typing or vice versa. Typing time during the post-editing is shorter when the source language has a fixed structure (i.e., sentence word order), whereas it is longer when a source language has a less-strict structure since the occurred MT errors are more serious (Absolon et al., 2018). In a study by Plitt and Masselot (2010), the typing time is defined as the sum of time intervals between keystrokes under 1 second. It is the time when the translator writes fluently without the pauses. This time differs from the thinking time or pauses by counting time intervals longer than 1 second into the edit (decision-making) time.

Assumptions

The study is focused on the impact of the quality of MT on the number of its corrections. We observe the quality of MT output in terms of the error rate (language, accuracy, terminology, and style), fluency, and adequacy to the source (original) text. We deal with the temporal effort (measured by time) by which we observe the MT efficiency and translators’ productivity. Typing time is the time from the first click/change in segment until the last click/change (within the edit operations: insert, delete, substitute, and reordering).

We stated the following assumptions:

- We assume that the translation productivity given by typing time is not related to the method of translation (PE vs. HT).
- We assume that typing time in the PE process does not depend on the error rate of the MT output in language, accuracy, terminology, and style (as defined by TAUS, 2013).
- We also assume that typing time in the PE process is not dependent on the adequacy and fluency of the MT output (as defined by TAUS, 2013).

Method

To collect data for the analysis, we conducted an experiment with seven students of translation studies at the master’s degree level with no experience with MT systems and with no explicit PE training but with similar translation competencies and practice (They were of the same year and the same language combination; they were selected due to the same evaluation [marks] from all study subjects.). The participants carried out two different tasks: post-editing of MT output from English into Slovak and making human translation in the same direction—from EN into SK (their mother tongue). The post-editing process was carried out in a special virtual environment recording keystroke logs and measuring thinking time, typing time, productivity (production segment), and the number of edit operations done by post-editors.

To be able to identify the relationship between specific MT issues and post-editing effort (typing time), all translators annotated MT output for translation quality according to the TAUS Dynamic Quality Framework (DQF)—Error Typology (Language, Accuracy, Terminology, and Style) and Fluency/Adequacy were used as the measures for MT evaluation (TAUS, 2013).

Participants

Participants were seven students from second year of the Master’s Translation program (novice translators), all female. Their median age was 24 years (range 24–25). All participants have the same typing skills (attended 6 weeks of typing course at university) and their median typing speed was 37 words/minute (range 35–39).

Materials

Two articles with a comparable complexity level were selected (according to participants’ opinions). The first article—a source text for post-editing—was a journalistic text published in Human Rights Magazine, consisting of 1,834 words and 79 sentences (6.5 SP). The translation direction was from English to Slovak (mother tongue). The Slovak MT output was gained from Google Translate. The second article—a source text for human translation, published in The Observer Magazine consisting of 1,820 words, 83 sentences, totally 6.5 standard pages—was translated by the same students (human translators). It was translated from a foreign language into the mother tongue, that is, from English into Slovak “from scratch,” without the MT or CAT support, exclusively using print or online dictionaries.

Methods

Error typology is a standard approach to translation quality evaluation, comprising four language areas, in our case: MT issues, that is (1) Language—grammar and syntactic errors, punctuation, (2) Accuracy—inaccuracy in meaning, omitted or added information, (3) Terminology—unacceptable terms for the concept, general terms, or glossary, and (4) Style—sometimes judged subjectively but needed to be evaluated according to the style standards in a given language. The error rate was ranked on a three-point scale according to the sensitivity of evaluator, that is, 1—representing a minor error, 2—major error and 3—critical error.

Evaluation scales—Adequacy and Fluency. We used a definition of the Linguistic Data Consortium (LDC), in which Adequacy is defined as a measure of meaning convey between the source and target text, that is, “how much of the meaning expressed in the gold-standard translation or the source is also expressed in the target translation,” and
Fluency “how fluent translation is,” that is, to what extent the translation is “one that is well-formed grammatically, with correct spellings, adheres to the common use of terms, titles and names, is intuitively acceptable and can be sensibly interpreted by a native speaker.” Both measures were ranked on five-point scales, where 5—means everything, 4—most, 3—much, 2—little, and 1—none for measure adequacy. For fluency, 5—means flawless, 4—good, 3—non-native, 2—disfluent, and 1—incomprehensible.

**Results**

The final dataset comprised 553 post-edited MT sentences and 581 HT sentences. For each MT sentence, the average error weight (separately for language/accuracy/terminology/style) was calculated by summing up the particular error weight of all annotators and divided by the number of annotators/evaluators. The same was calculated for fluency and adequacy.

The objective of the analysis is to identify the relationship between PE effort (typing time) and error typology, as well as between PE effort (typing time) and fluency or adequacy of MT output. We also assumed that typing time in the PE process does not depend on the error rate of MT output in language, accuracy, terminology, and style (as defined by TAUS). The following null hypotheses follow from our assumptions:

H01: Typing time in the PE process does not depend on the error rate of the MT output in language.
H02: Typing time in the PE process does not depend on the error rate of the MT output in accuracy.
H03: Typing time in the PE process does not depend on the error rate of the MT output in terminology.
H04: Typing time in the PE process does not depend on the error rate of the MT output in style.
H05: Typing time in the PE process is not dependent on the adequacy of the MT output.
H06: Typing time in the PE process is not dependent on the fluency of the MT output.

To not decrease the power of statistical tests, we used non-parametric procedures concerning deviations from normality. To test H0s, we used a non-parametric correlation, namely Kendall’s Tau correlation significance test. In the case of our hypotheses, we do not distinguish between dependent and independent variables, given that, it is a symmetric degree of dependence. However, for our research, we consider the variable time to be dependent and errors to be independent.

After the rejection of null hypothesis at the .001 significant level (typing time & language: $Z = 4.800307$, $p = .000002$; typing time & accuracy: $Z = 6.366129$, $p = .000000$; typing time & terminology: $Z = 4.798103$, $p = .000002$; and typing time & style: $Z = 6.601318$, $p = .000000$) we can claim, that the typing time is significantly affected by language errors (grammatical errors, incorrect, or omitted punctuation), accuracy (misunderstanding of meaning, additions, omission of essential information, incorrect translation within the context), terminology (using incorrect terms), and style (incorrect order of phrases).

Despite the different influence of these language areas, errors, or mistakes (Figure 1), the typing time is affected by error rate and similarly, the efficiency of the MT and the productivity of translation itself. From the graph (Figure 1), where the $x$-axis depicts errors and $y$-axis time, we can see a statistically significant directly proportional dependency between time and errors.

Fluency and adequacy were analyzed similarly. We also assumed that typing time in PE process is not dependent on adequacy and fluency of MT output (as defined by TAUS). Based on the analysis results, we reject the null hypothesis at the .001 significant level (typing_time & fluency: $Z = 6.455408$, $p = .000000$; typing_time & adequacy: $Z = 5.54195$, $p = .000000$). We can claim that there is a significant dependency between typing time and adequacy and also between typing time and fluency of MT output. Adequacy reflects whether the sentence conveys information from the source text; fluency refers to comprehensibility. In PE, the meaning of the sentence is prior. In the case of meaning discrepancies, typing time increases. Within fluency, we observe the correctness of sentence structure. Moreover, the degree of correctness of the transformed segment (the segment adequate to its equivalent), influences typing time in a significant way.

We can see (Figure 2) a statistically significant directly proportional dependency between time and adequacy/fluency, where the $x$-axis depicts adequacy/fluency and $y$-axis time.

Besides the product analysis of MT quality from the aspect of typing time, we also focused on the process analysis in terms of translator productivity or MT system efficiency, that is, we tried to identify the relationship between the translation as a process (machine vs. human translation) and PE effort (typing time). We assumed that the translation productivity given by typing time is not related to the method of translation (PE vs. HT). From the assumption (the translation productivity given by typing time is not related to the method of translation (PE vs. HT), the null hypothesis is followed:

H07: There is no statistically significant difference between PE and HT in term of time.

We tested the differences between the way of translation (PE vs. MT) and typing time using the Mann-Whitney $U$ test. Based on the results ($Z = -2.66109; p = .007789$) we can reject a null hypothesis at the .01 significant level. We can claim that there is a statistically significant difference between human translation and post-editing of MT output from a typing time perspective in favor of PE (Figure 3).
Median for typing time of human translation, consisting of 87 sentences (1,820 words) (Figure 3) is 44 seconds, a lower quartile (25%) is 25 seconds, and the upper quartile (75%) is maximum 75 seconds, that is, 25% of the values of typing time lie below 25 seconds and 75% of the values of typing time below 75 seconds. The total typing time (Figure 3) is 267 seconds and the interquartile range is 46 seconds in the case of human translation.

The second method of translation—post-editing—the PEMT output consists of 79 sentences (1,834 words). Median for typing time (Figure 3) is 37 seconds. About 25% of the sentences (a lower quartile) have a typing time lower than 8 seconds, and also 75% of the sentences (upper quartile) have a typing time larger than 8 seconds, but lower than 63 seconds. The total range for PE (Figure 3) is 220 seconds, and the interquartile range is 55 seconds for post-editing.

**Discussion**

The results have shown that the typing time is very closely related to errors in language, accuracy, terminology, and style as well as to fluency and adequacy. The highest dependency is between the typing time and the style of the given text. Mistakes in the style mostly consist of incorrectly arranged phrases or sentence elements within sentences. The typing time is longer because post-editors need mostly to post-edit and sometimes rewrite the whole sentence or compound/complex sentence to give meaning to the reader. Slovak—for its structure and feature—is demanding for MT systems and post-editors. To correct MT sentence precisely, they need to consider all aspects of the sentence—lexical, morphological, syntactic, and semantic.

Another very significant category of errors that affect typing time is language (grammar). The similarities and differences between translation languages have impacts on MT quality and PE. In the case of our study, the source language is English and the target language is Slovak. There are many (above-mentioned) differences between them which may result in problems in MT and PE processes. The key differences lay in the origin of the languages—English is a Germanic language, Slovak belongs to the Slavonic group of languages. This fact indicates that the first significant differences would result from their origin. Problems in PE also occur in the cases in which post-editors need to spend more
Figure 2. Dependency between typing time and fluency or adequacy in PEMT.

Figure 3. Differences between HT and PEMT from the typing time perspective.
typing time to correct grammar form and to keep the congruence at a syntactic level in the category of gender, number, and case, mainly in noun phrases. Although there are a lot of language errors in the MT, the typing time may be shorter than in the area of style or accuracy because post-editors do not need to rewrite whole sentences but fixe or concentrate only on the words with a mistake in them.

The factor that affects the typing time the least (but still statistically significant) is terminology. Similarly to the language issue, post-editors have to manually rewrite only the incorrect term.

As already mentioned, the efficiency of PE in machine translation depends on the error rate of machine translation. In our research, we have found out that editing time is (very) closely related to errors in language, accuracy, terminology, and style, as well as fluency and adequacy (see Results).

The minimum—the shortest PE time—was 0 second. The time of 0 seconds during the PE means that the post-editors did not edit or correct any MT segment; they only confirmed its correctness. This could have occurred for several reasons.

The first reason is that a segment contains only proper names which are unnecessary to translate. Names are not translated even in human translation as they are unchangeable. Exceptions may be found in belles-lettres style: the names of fictional characters are translated if they carry, for example, a feature that would be lost by keeping the original name. However, this is not the case of the examined text as it contains the names of real people which are the same in both languages. The editing time of such segments (sentences) was 0 second as the MT had almost no point to make a mistake, as the text of the original and the translation were the same.

Another case of unnecessary editing of MT output are numbers. Translating of numbers is a simpler procedure than translating of names as the numbers have the same form in all languages using Arabic numerals. However, discrepancies can occur in writing decimals: in Slovak, decimals are marked by a comma, in English by a decimal point/dot. Thousands are also marked in different ways and attention should be paid to the symbols of currency which are placed after the number in Slovak but before the number in English.

Some common phrases are also not necessary to be post-edited as MT can recognize them easily and translates them in the same and correct way. In our case, MT correctly translated “available at” as “k dispozícii na.” This phrase is used for example, with hypertext links (links to web sites with further information). Hypertext links represent another case of unnecessary post-editing. The link has the same form in both languages as it links to the same site.

The last type of segments (sentences) that do not need to be post-edited are simple sentences. In the case of simple sentences and unambiguous words (words with one meaning, no polysemy or homonymy), MT can translate the segments correctly. We can demonstrate it on an English sentence from our text: “It seems as if we aren’t even on the radar screen.” It was translated into Slovak as: “Zdá sa, že nie sme ani na radare.” Although human translators would translate the sentence (without MT) differently, this sentence is from a lexical and terminological point of view translated correctly and no further grammatical corrections are needed.

When a source text mainly consists of the sentences just mentioned, it is highly suitable for translation by MT and subsequent post-editing. A text with a high number of names, hypertext links, numbers, terms, simple sentences, and unambiguous words recognized by MT system, the MT output is not needed to be post-edited (to exceptions). By this, the use of MT with PE becomes effective and the productivity and performance of translators/post-editors increases.

We examined the efficiency of PE compared to human translation “from scratch.” We compared the post-editing and human translation based on the typing time. Despite the relatively small dataset (553 post-edited MT sentences and 581 HT sentences), typing time during the human translation was longer (almost 18%). It has been shown that shorter PE time is related to corrections in morphology, word forms, and simple substitutions of words of the same word class; on contrary, longer PE time is related to corrections in word order issues, in the change of word classes and correction of mistranslated idioms (Koponen et al., 2012).

More segments were easier to post-edit and the typing time was short or even equals zero. There were also the segments with longer typing time which of course increased the average time of the total typing time of post-editing. However, such cases were not so often and the total typing time was shorter than the total typing time of human translation.

**Conclusion**

Undoubtedly, translation is one of the key tools of human communication. Currently, as translation technologies are almost fully integrated into the translation process, it is essential to know how to utilize them effectively. To achieve the desired efficiency, it is vital to approach the core of understanding critically and informedly to understand what the tools can provide and what they are not.

Machine translation systems are essentially developed from human translations. Nowadays the systems contain millions of sentences translated by humans, from which the “machines” are trained. Machine translation systems are being constantly improved, increasing the quality of translation to make translators’ work more efficient. Their architecture is ceaselessly being further developed, moreover, they are equipped with an increasing number of high-quality translations. Doherty (2016) claims that current and future problems of machine translation can be found in two...
areas: (1) quality of data from which systems learn; (2) balanced compromise between the time and the amount of data used for teaching MT systems to translate.

On the other side, there is a direct proportion between “good” machine translation (i.e., translation of higher quality) and “bad” machine translation (i.e., translation of lower quality). “Bad” machine translation is still unproductive, on the contrary, “good” machine translation can make post-editors’ work easier and thus increase their productivity. Researchers examining post-editing in MT agree that PE can increase productivity in terms of speed and sometimes in quality of the translation (Bangalore et al., 2015; Nitzke, 2019; O’Brien et al., 2014). Daems et al. (2017) carried out a test with eight professional translators and 10 students of translation studies, they found that post-editing of newspaper texts was significantly faster than translation from scratch. Lacruz et al. (2016) note that post-editing increases the translators’ productivity compared to the translation from scratch. Carl et al. (2015) conducted an experiment with 12 translators, in which 10 out of 12 translators were faster when post-editing, compared to their translation baseline. Screen (2017) showed (in his research with eight translators) that post-editing of MT could significantly speed up translation and also the increase in productivity does not lead to a decrease in overall quality.

In terms of time editing, PE is much more efficient than human translation as in the latter one translators need to write every single word they translate. In post-editing, however, all the words, sentences, and segments are pre-translated by MT. When they are translated correctly, translators do not have to deal with them anymore and the editing time is much shorter than typing (each word using the keyboard).

Texts containing numbers, names, hyperlinks, terms and phrases, or simple constructions without polysemy and homonymy are for MT easier to translate correctly and require no post-editing, except small exceptions. This makes the use of MT and its PE effective and increases the translators or post-editors productivity.

Therefore we can state that post-editing of machine translation has a significant advantage in comparison with human translation. PE of long and complicated sentences, sentences with incorrect syntax and those “misunderstood” by MT took us the most of post-editing time. Errors even occurred in the lexis. However, similar types of sentences cause problems even in human translation: the longer and the more demanding the sentence, the longer it takes the translator to write the translation.

If a human translator can efficiently use translation technologies such as MT systems, CAT tools with translation memory, or terminology databases, they can substantially increase their performance and productivity at work. Brunette and O’Brien (2011) call for increased research into people’s efficiency in their working environment, particularly on cognitive aspects of the interaction between human translators and technology (CAT tools such as TM or MT).

The machine translation market is growing at an incredibly fast and it is expected to reach US $ 980 million by 2022 (up from US $ 78.9 million in 2012 and US $ 89.4 million the following year) (Grand View Research, 2018). Nowadays, post-editing constitutes a significant element in the translation industry (De Palma et al., 2016). Its situation and quality can be improved with the improvement of technologies and MT systems. This modern and fast period requires translation of high quality and asap (as soon as possible) delivery.

Authors’ Note
Each of the authors confirms that this manuscript has not been previously published and is not currently under consideration by any other journal. Additionally, all of the authors have approved the contents of this paper and have agreed to the journals submission policies.

Each named author has substantially contributed to conducting the underlying research and drafting this manuscript.

Katarina Welnitzova is now affiliated to University of Ss. Cyril and Methodius in Trnava, Slovakia.

Declaration of Conflicting Interests
The author(s) declared no potential conflicts of interest with respect to the research, authorship, and/or publication of this article.

Funding
The author(s) disclosed receipt of the following financial support for the research, authorship, and/or publication of this article: This work was supported by the Scientific Grant Agency of the Ministry of Education of the Slovak Republic and of Slovak Academy of Sciences under contract VEGA-1/0809/18, and the Slovak Research and Development Agency under the contract APVV-18-0473.

ORCID iD
Dasa Munkova https://orcid.org/0000-0002-1003-7929

References
Absolon, J., Munková, D., & Welnitzová, K. (2018). Machine translation: Translation of the future? Machine translation in the context of the Slovak language. VERBUM.
Allen, J. (2003). Post-editing. In H. Somers (Ed.), Computers and translation: A translator’s guide (pp. 297–317). Benjamins.
Baker, M., & Saldanha, G. (Eds.). (2009). Routledge encyclopedia of Translation studies. Routledge.
Bangalore, S., Behrens, B., Carl, M., Gankhot, M., Heilmann, A., Nitzke, J., Schaeffer, M., & Sturm, A. (2015). The role of syntactic variation in translation and post-editing. Translation Spaces, 4(1), 119–144.
Bowker, L. (2019). Fit-for-purpose translation. In M. O’Hagan (Ed.), The Routledge Handbook of translation technology (pp. 453–468). Routledge.
Pre-editing plus neural machine translation for subtitling: Effective pre-editing rules for subtitling of TED talks [Conference session]. Machine Translation Summit XVII. Dublin, Ireland (pp. 64–72). https://www.aclweb.org/anthology/W19-6710.pdf

Holmes, J. S. (1975). The name and nature of translation studies. Translation Studies Section, Department of General Literary Studies, University of Amsterdam.

Koponen, M. (2012a). Is machine translation post-editing worth the effort? A survey of research into post-editing and effort. The Journal of Specialised Translation, 9(25), 131–148.

Koponen, M. (2012b, June 7–8). Comparing human perceptions of post-editing effort with post-editing operations [Conference session]. The NAACL 2012 Workshop on Statistical Machine Translation, Montreal, Canada (pp. 181–190).

Koponen, M. (2015, November 3). How to teach machine translation post-editing? Experiences from a post-editing course [Conference session]. MT Summit Workshop on Post-Editing Technology and Practice. Miami, United States (pp. 2–15). http://www.mt-archive.info/15/MTS-2015-W1-Koponen.pdf

Koponen, M., Aziz, W., Ramos, R., & Specia, L. (2012, October 28). Post-editing time as a measure of cognitive effort [Conference session]. Conference of the Association for Machine Translation of the Americas. Workshop on Post-Editing Technology and Practice, San Diego, United States (pp. 11–20).

Kring, H. P. (2001). Repairing texts: Empirical investigations of machine translation post-editing process. Kent State University Press.

Lacruz, I., Carl, M., Yamada, M., & Aizawa, A. (2016, March 7–11). Pause metrics and machine translation utility [Conference session]. The 22nd Annual Meeting of the Association for Natural Language Processing (NLP2016), Sendai, Japan (pp. 1–4).

Läubli, S., Amrhein, C., Düggelin, P., Gonzalez, B., Zwahlen, A., & Volk, M. (2019). Post-editing productivity with neural machine translation: An empirical assessment of speed and quality in the banking and finance domain. ArXiv, abs/1906.01685.

Marriniello, E., & Steiert, A. (2016). The human role in a machine-translation world. TCodeWorld, 32–35. https://afaftranslations.com/wp-content/uploads/2016/07/human_machine-translation%20world.pdf

Munková, D. (2013). Pristupy k strojovému prekladu (modely, metódy a problémy strojového prekladu). Univerzita Konštántina Filozofa.

Nitzke, J. (2019). Problem solving activities in post-editing and translation from scratch: A multi-method study. Berlin: Language Science Press.

O’Brien, S. (2002, November 14–15). Teaching post-editing: a proposal for course content [Conference session]. 6th EAMT Workshop Teaching Machine Translation, Manchester, UK (pp. 99–106).

O’Brien, S., Balling, L. W., Carl, M., Simard, M., & Specia, L. (Eds.) (2014). Post-editing of machine translation: Processes and applications. Cambridge Scholars Publishing.

Plitt, M., & Masselot, F. (2010). A productivity test of statistical machine translation post-editing in a typical localisation context. Prague Bulletin of Mathematical Linguistics, 93, 7–16.

Quah, C. H. K. (2006). Translation and technology. Macmillan.
Sánchez-Gijón, P., Moorkens, J., & Way, A. (2019). Post-editing neural machine translation versus translation memory segments. *Machine Translation, 33*, 31–59.

Screen, B. (2017). Productivity and quality when editing machine translation and translation memory outputs: An empirical analysis of English to Welsh translation. *Studia Celtica Posnaniensia, 2*(1), 1–24.

TAUS. (2013). *Dynamic quality framework (DQF) error typology*. https://www.taus.net/academy/best-practices/evaluate-best-practices/error-typology-guidelines

Vieira, L. N., & Alonso, E. (2020). Translating perceptions and managing expectations: An analysis of management and production perspectives on machine translation. *Perspectives, 28*(2), 163–184.

Way, A. (2018). Quality expectations of machine translation. In J. Moorkens, S. H. Castilho, F. Gaspari, & S. Doherty (Eds.), *Translation quality assessment: From principles to practice* (pp. 159–178). Springer.

TAUS. (2016). *TAUS post-editing guidelines*. https://www.taus.net/think-tank/articles/postedit-articles/taus-post-editing-guidelines