Eglė Konstantinavičienė

MA in Translation,
Faculty of Social Sciences, Arts and Humanities
Kaunas University of Technology, Lithuania,
https://orcid.org/0009-0009-7788-5840
e-mail: konstantinaviciene@yahoo.com

Ramunė Kasperė

Dr., Professor
Faculty of Social Sciences, Arts and Humanities
Kaunas University of Technology, Lithuania,
https://orcid.org/0000-0003-0782-3758
e-mail: ramune.kaspere@ktu.lt

Svitlana Matvieieva

Dr., Professor
Faculty of Foreign Philology
Mykhailo Dragomanov State University of Ukraine,
Faculty of Social Sciences, Arts and Humanities
Kaunas University of Technology, Lithuania,
https://orcid.org/0000-0002-8357-9366
e-mail: s.a.matvyeyeva@udu.edu.ua

POST-EDITING DECISIONS
IN MACHINE TRANSLATION OF LEGAL CONTRACTS

Abstract

Legal translation requires special language use. It has therefore been many a time credited to be one of the most difficult genres to translate. Post-editing of machine translation is becoming more and widely employed in the translation industry. Even in languages that cannot boast huge amounts of resources, translating texts of certain genres is rarely conducted without resorting to machine translation, as the resulting quality seems to be approaching parity with human translation. This study attempts to analyze and compare post-editing decisions in machine translated legal contracts from English to Lithuanian and from English to Ukrainian. Legal contracts are part of the legal discourse with language characterized as being archaic, complex, culturally bound and typically stocked with formulaic phrases, specialized terms, and Latinisms. Being a technical type of discourse, legal translation may benefit from machine translation. The conclusions of the study are based on the findings of assessment of machine translation post-editing decisions made by professional translators. The findings demonstrate and support the results of previous research with other languages that the main trends are as follows: changes in noun and verb phrases in both languages, co-reference and reordering changes. Further study of the stated problem will allow expanding (targeting and detailing) the typology of post-editing actions, considering the specifics of Baltic (Lithuanian) and Slavic (Ukrainian) languages.

Keywords: machine translation, post-editing, post-editing actions, legal contracts.

1. Introduction.

Legal translation is among the most complex types of technical translation. It is also by far the most demanding from the translator's point of view. Professional translators working with legal texts have to be well versed in legal language which is typically archaic, sophisticated, culturally bound and typically stocked with formulaic phrases, specialized terms, and Latinisms. Chéragui (2012) suggests that although technical translation work is challenging and strenuous, most of it is dull and repetitive, at the same time necessitating diligence and precision. While computer-assisted tools are now a standard aspect of a translator's routine, the adoption of machine translation becomes an alternative solution where human resources are insufficient. Having this in mind, legal translators and/or clients may benefit considerably from machine translation. Post-editing (PE) is becoming more and more widely employed and relied on in technical and, therefore, legal translation.

This study aims to analyze post-editing decisions in machine translated legal contracts, which are a typical manifestation of legal texts. To equip legal linguists, translators, and post-editors with the necessary tools of their trade, legal translation has become the subject of "linguistic research within the framework of comparative legal linguistic and legal translation and interpreting theories and practices" (Gortych-Michalak, 2017). Therefore, in addition, a comparison between post-editing decisions in English to Lithuanian and English to Ukrainian is carried out in order to determine if there are any cross-linguistic idiosyncracies and parallels, which under further research may allow developing universal generalized systemic approaches to the post-editing of legal texts for the studies target languages.

The present study is expected to contribute to our understanding of cognitive efforts that post-editors make and the reasons behind their choices in two different yet distantly related languages. This article is organized as follows: literature review overviews existing studies on machine translation post-editing of legal content in a variety of languages, followed by aims, objectives and methodology adopted in the study; the section on the results is supplemented with examples of the major study outcomes, which are further

highlighted and compared in the discussion section; the conclusions summarize the results, highlight the implications, acknowledge the limitations and outline future research directions.

2. Literature review.

Many languages of the world are under-resourced in terms of parallel corpora that are important for precision in machine translation. Lithuanian and Ukrainian are among such languages (Utka et al., 2016; Grabar et al., 2019). Efficient post-editing that also requires less temporal and cognitive effort largely depends on the quality and precision of machine translation output. Below we overview recent publications reporting the results of research on machine translation and post-editing in various language pairs, mostly focusing on under-resourced languages.

In a study by Killman and Rodriguez-Castro (2022) where an experimental research design was carried out involving 26 translators who translated and post-edited a legal text in the English to Spanish pair, the general implication focused on the benefits of post-editing in terms of quality and time gains compared to translation. The output of machine translation was found to be especially useful at the terminological and phraseological levels.

A usability study of machine translated segments from English to Croatian was conducted by Omazić and Šoštarić (2023). Text versions processed by five machine translation tools, including generative artificial intelligence based, were found to be highly usable with the majority requiring only minor or major revisions, which allowed the authors of the study to conclude that machine translation output requires post-editing as a necessary step in translation of legal texts (Omazić & Šoštarić, 2023).

Of mention here is a study on machine translation for post-editing practices in the English to Lithuanian language pair, which is also one of the objects of the current study. Although Povilaitienė and Kasperė (2022) employed a news text, the findings may be relevant here to note since three well known machine translation engines were deployed and demonstrated mainly minor errors in the machine translated output. In addition, accuracy errors were in all three cases the most abounding.

The three above mentioned studies, two in the direction from the language of rich resources, i.e., English, to the languages deemed under-resourced (namely, Croatian, Lithuania) as well as the study describing findings in Spanish, reported on the machine translation quality to a lesser of greater extent. Human interference can be emphasized to be relevant.

On the other hand, a look at studies analyzing machine translation and/or post-editing quality in a reverse direction, i.e. from a low-resourced language to English as the language with the most resources, reveals similar results. Here we overview the results of machine translation and/or post-editing quality assessment from Greek, Korean and Hungarian, which are considered as either under-resourced or medium resourced languages (see Sosoni et al., 2022; Park et al., 2020; Yankovskaya, 2023).

A study where machine translation from Greek to English, the language pair deemed under-resourced, was assessed involved a comparison of post-edited and human translated legal texts (Sosoni et al., 2022). The authors reported no productivity gains in the post-edited version; they did not either observe major differences in accuracy or fluency between the post-edited and human translated versions. The human translated texts exhibited errors of accuracy, while the post-edited text contained style and verity errors (Sosoni et al., 2022).

Kovács (2022) compared the quality of machine translated legal texts in both directions: from English to Hungarian vice versa. One tendency that was equally noted in both translation directions was inconsistency of adverbs that was found to be more common

compared to nouns and verbs. However, in terms of the other tested parameters – source-language faithfulness and target-language well-formedness – machine translation from English to Hungarian resulted in a lower quality compared to human translation. On the other hand, inconsistency of special legal terminology use was more pronounced in Hungarian to English machine translation (Kovács, 2022). Therefore, post-editors' experience and competence are crucial in completing high quality requiring post-editing tasks.

The study of the English-language texts of official OSCE reports translated into Ukrainian using machine translation systems (Koroliova et al., 2020) demonstrated that the most difficult task for machine translation is the matching of language units by person, numbers and cases, resulting in incorrect selection of correspondences. These issues make approximately 25% of all identified errors. Furthermore, about 13–14% of errors stemmed from literal translation, causing distortions in the original message. The authors of the study emphasize that machine-translated texts, while providing a general understanding, require refinement through post-editing by a human to ensure accuracy.

T. Arterchuk (2014) proves the effectiveness of employing machine translation followed by post-editing not only for direct translation, but also for training future lawyers. The researcher concludes that the teacher should emphasize to students the critical nature of precise specialized text translation, clarify the key lexical, grammatical and stylistic features of legal language, highlight the main shortcomings of machine-translating legal texts, and pay attention to the human factor, i.e. teach fundamental techniques for post-editing machine-translated results.

Some other studies also report evidence of research that machine translation post-editing might not be helpful in case of legal translation. In a study where English to Chinese human translation and post-editing scenarios were compared, Cui et al. (2023) found that semi-professional translators rated legal texts as requiring more cognitive effort than other types of texts. The authors argue that legal jargon and complex syntax make it difficult for translators and machine translation seems not to be beneficial. An example of an incorrectly translated legal term is provided as evidence that machine translation may still struggle with legal terminology where precise rather than straightforward, word-for-word translation of terms is required (Cui et al., 2023). Such cases might be a confirmation of lower efficiency of post-editing, which may also turn to be demanding from the perspective of time.

Although legal jargon has been disclosed to be an issue for machine translation, Mileto (2019) proposes a solution that may provide needed quality in machine translation post-editing of legal documents. The author maintains that custom-designed termbases used along with machine translation may help achieve consistency and lead to reduced temporal efforts while retaining undiminished quality (Mileto, 2019).

Given all the pros and cons reported in the literature on the use of machine translation for post-editing purposes and the peculiarities of machine translation in low-resourced languages, we hope that post-editors may benefit from an overview of the tendencies of post-editing actions reported in this study.

3. Aim and Objectives.

This study attempts to analyze and compare post-editing actions made by professional translators in machine translated legal contracts from English to Lithuanian and from English to Ukrainian. The objectives are as follows: (1) to compare post-editing actions of machine translated output of legal contracts in two languages; (2) to evaluate the quality of machine translated legal text before and after editing; (3) to analyze and determine the tendencies in post-editing actions (PEA) in a machine translated legal contract from English to Lithuanian and from English to Ukrainian.

4. Methodology.

Various taxonomies of machine translation errors have been designed exploiting examples from a number of languages (Flanagan, 1994; Vilar et al., 2006; Dugast et al., 2007; Lommel et al., 2014, among others). Studies in numerous languages and a variety of text types and genres have been conducted. However, focusing on the errors is one step only in the search for the best translation possible. Identifying the actions post-editors make to correct errors may help in understanding more about the nature of post-editing and in providing guidelines for translation students, novice translators and post-editors. In this study, we adopt Blain et al.'s (2011) approach who proposed a typology of post-editing actions which are the smallest logical edits that linguistically make sense made by the post-editor in a machine translated text. These edits are opposed to mechanical edits – insertion, deletion, substitution and move (Blain et al., 2011). The typology was based and tested on translation from English to French. The edits fall in 8 categories, namely related to (1) noun phrases in terms of lexical changes (including determined change, noun meaning change, noun stylistic change, noun number change, case change, adjective change, multiword change, noun phrase structure change); (2) verb phrases in terms of grammatical changes (including verb agreement correction, verb phrase structure change, verb meaning change, verb stylistic change); (3) preposition change; (4) co-reference change; (5) reordering; (6) post-editing error; (7) miscellaneous (unnecessary) stylistic change; (8) other unclassified miscellaneous changes. The original post-editing action analysis was conducted with rule-based machine translation and statistical machine translation. In Blain et al.'s (2011) study, the vast majority of changes were found to fall within the category of noun phrase changes (90% in case of rule-base machine translation and 92% with statistical machine translation).

In our study, we analyze a text of legal contract translated from English to Lithuanian and Ukrainian by a freely available neural machine translation system. The data analyzed consists of the extract from a sample partnership agreement. The extract from the agreement consisted of 1075 words. The machine translated texts were post-edited by six translators/post-editors who were then given a short survey regarding their perceptions on the machine translation and post-editing of the text. The statements (adapted from Castilho, 2016; Kasperavičienė et al., 2020) were as follows: (1) the text of the contract is easy to understand and clear; (2) the language of the contract text (grammar, vocabulary) is easy to understand; (3) the sentences in the text of the contract sound natural; (4) the provisions contained in the text of the contract are easy to understand and clear; (5) the parties can use the text of the contract and assume the rights and obligations provided for in it; (6) it is not necessary to read the text in the original language in order to understand it; (7) I could recommend the text of the contract to a person looking for such a contract. The post-editors had to respond to the statements by selecting out of 5 options: disagree, partially disagree, neither agree nor disagree, partially agree, or agree. Additionally, the post-editors were interviewed by asking them whether it was difficult to post-edit such a text.

5. Results.

As expected, noun phrase changes were most common among all other 8 categories. In this category, noun replacement with another noun by changing the meaning was most frequent. All other subcategories within the noun phrase change category were found, except for the noun phrase structure change. In the category of verb phrase changes, cases of verb meaning change and verb stylistic change were found, but no verb agreement correction or verb structure change were found. Cases of co-reference changes, i.e. introduction or removal of a pronoun, and reordering when an adjective or an adverb is repositioned to a better location were also present. No post-editor introduced errors nor unnecessary stylistic

changes, i.e. miscellaneous style changes, were found. On the other hand, other miscellaneous changes that could not be classified with Blain et al.'s typology were observed. This may require further investigation as corrections of technical errors – unnecessary punctuation marks, extra spaces, repetitive phrases – were more common than accidental. Examples of post-editing actions made by Lithuanian and Ukrainian translators are provided below for illustration purposes. Also, given that the analyzed text was of a relatively small size, some post-editing actions may not have occurred, thus requiring more extensive research.

<div align="center">

Example 1

</div>

Source: *In order to accomplish their aforesaid <u>desires</u>, the parties hereto <u>desire</u> to join together in a <u>general partnership</u> under and pursuant to the <u>Uniform Partnership Act</u>, amended from time to time (the "Act").*

MT output in LT: *Siekdamos įgyvendinti savo pirmiau minėtus <u>norus</u>, šalys <u>nori</u> prisijungti prie <u>tikrosios ūkinės bendrijos</u> pagal ir pagal <u>Vienodos ūkinės bendrijos įstatymą</u>, retkarčiais keičiamą (toliau – Aktas).*

Human post-editing LT1: *Siekdamos įgyvendinti savo pirmiau minėtus <u>pageidavimus</u>, šalys <u>siekia</u> prisijungti prie <u>viešosios bendrijos</u> pagal <u>Vienodos ūkinės bendrijos įstatymą</u>, retkarčiais keičiamą (toliau – Aktas).*

Human post-editing LT2: *Siekdamos įgyvendinti savo pirmiau minėtus <u>norus</u>, šios sutarties šalys <u>pageidauja</u> prisijungti prie <u>tikrosios ūkinės bendrijos</u> pagal <u>Tipinės partnerystės įstatymą</u> su visais pakeitimais (toliau – Įstatymas).*

Human post-editing LT3: *Šalys, siekdamos įgyvendinti prieš tai nurodytus savo <u>pageidavimus</u>, <u>pageidauja</u> tapti <u>bendros jungtinės veiklos partnerėmis</u> pagal <u>sudarytą bendrovės aktą</u> su numatomais vėlesniais pakeitimais (toliau – aktas).*

MT output in UA: *Щоб виконати свої вищезазначені <u>бажання</u>, сторони <u>бажають</u> об'єднатися в <u>загальне товариство</u> відповідно до <u>Уніфікованого закону про партнерство</u>, до якого час від часу вносяться зміни ("Закон").*

Human post-editing UA1: *Для досягнення вищезазначеного сторони цього Договору <u>бажають об'єднатися</u> в <u>повне товариство</u> відповідно до <u>Єдиного закону про партнерство</u> зі змінами та доповненнями ("Закон").*

Human post-editing UA2: *З метою реалізації вищезазначених <u>намірів</u> сторони цього Договору <u>бажають</u> вступити в <u>повне товариство</u> відповідно до <u>Єдиного закону про товариства</u> з періодичними змінами й доповненнями ("Закон").*

Human post-editing UA3: *Для виконання вищезазначеного <u>бажання</u> сторони Договору <u>об'єднуються</u> в <u>повне товариство</u> відповідно до <u>Єдиного закону про товариства</u> з подальшими змінами й доповненнями ("Закон").*

In this example, it is important to pay attention to the correction of the translation of the term *Uniform Partnership Act*. Lithuanian translators chose the options *tipinės partnerystės įstatymą* and *bendrovės aktą*. Considering the fact that the legal basis in Lithuania differs from the legal basis of the source text ("The Uniform Partnership Act provides governance for business partnerships in several U.S. states" (Kenton, 2022)), there is no equivalent in Lithuanian. In this case, the post-editor who chose the option – *tipinės partnerystės įstatymą* performed a multi-word change – multiword expression change (meaning change). The post-editors who chose the option *bendrovės aktą* made an edit, classified as PE Error, i.e., the post-editor made a mistake in Blain et al.'s taxonomy, whereas the meaning of the edited term refers to a document prepared by a company and not to a law.

There are also no direct equivalents in the Ukrainian language. As mentioned above, the cause of such a situation lies in the differences in the structure and principles of legal

systems resulting in lacunae in terminological systems (Matvieieva, 2020, p. 187), which can be filled or corrected by a human translator or a post-editor. In this sample, the post-editors used two variants for the construction *Uniform Partnership Act* – *Єдиний закон про партнерство* and *Єдиний закон про товариства*. These equivalents are offered instead of a literal machine translation – *Уніфікований закон про партнерство*. This edit can be classified as a change in adjective choice for better fit with the modified noun.

Another term that was changed during post-editing is the term *General Partnership*. Machine translation resulted in the Lithuanian term *tikrosios ūkinės bendrijos*. The post-editors made changes to rended this term as *viešosios bendrijos pageidauja tapti bendros jungtinės veiklos partnerėmis*. The selected term *viešosios bendrijos* may be classified as a PE Error, as the term is not used in Lithuanian. Such a form of legal entity does not exist in Lithuania, and the same time such rendering has no meaning and cannot be clearly understood. The changes made by the post-editor who went for the option *pageidauja tapti bendros jungtinės veiklos partnerėmis* may be classified as multi-word change – multiword expression change (meaning change).

Meanwhile, all three Ukrainian post-editors changed the option suggested by the machine (*загальне товариство*) by the construction *повне товариство*. Such an equivalent is found in the English-Ukrainian legal dictionary (Karaban, 2003, p. 461). The construction *загальне товариство* is also used in Ukrainian, but it is not registered in legal dictionaries and official documents. It is used in analytical and popular texts to describe this type of companies in other countries, not Ukraine. The change for *повне товариство* is identified as a change in adjective choice as well.

The Lithuanian post-editors made changes to the machine translation variant of source words *desires / desire*; instead of *norus / nori*, Lithuanian post-edited text contained *pageidavimus / pageidauja, siekia*. Following Blain et al.'s typology, such changes may be classified as Misc style, which may be regarded as an unnecessary stylistic change. Both Lithuanian words offered by machine translation are used and correctly translated, so the change was not necessary. On the other hand, the change resulted in a higher formality level, typical of such texts.

For the Ukrainian language in machine translation, there are equivalents *бажання / бажають,* for which each post-editor suggests changes: (1) Ø / *бажають (об'єднатися)*, (2) *наміри / бажають (об'єднатися)*, (3) *бажання / об'єднуються*. The reason for such changes is to prevent tautology, which, although often inherent in legal texts, is recommended to be avoided. The noun *намір* used in the second variant, is not fixed in the dictionary as a synonym for the noun *бажання*, but has such a seme in its meaning: "намір – задум, бажання зробити що-небудь" / transl.: intention – a plan or desire to do something (Busel, 2005, p. 724). So, in the first case, the construction 'adjective + noun' (*aforesaid desires*) is replaced with an adjectival form serving as a noun; in the second and third cases, the same parts of speech are used. This PEA can be classified as stylistic change where the words (nouns and adjectives) are replaced by the synonyms with minimum meaning change to prevent stylistic errors and inaccuracies.

To mention another common mistake in machine translations – a needless repetition – an unnecessary word repetition was made in Lithuanian machine translation output for the English *under and pursuant*, i.e. *pagal ir pagal* (lit. en. *according to and according to*) in the middle of the sentence, which occurs as an obvious error. Doublets are two synonyms used together, a problematic issue in translation (Heller & Zoyirova, 2022). They are "obsolete, redundant, unnecessary, pleonastic and ambiguous" and, what is more important, non-existent in many languages (Buşila, 2017, p. 123). In English legal language, the doublet *under and pursuant* is typical and common. In Lithuanian, no such doublet is established.

Machine translation renders both synonymous words literally. Post-editors therefore deleted one occurrence of *pagal*, thus making a change attributed to the class of miscellaneous changes.

As for the Ukrainian translation, the named construction (*under and pursuant*) was machine translated correctly (*відповідно до*) and didn't require any post-editing. As for other transformations, the pronoun *свої* was removed in post-editing as redundant and stylistically incorrect – a co-reference change. The punctuation of the translation was also corrected, as the post-editors used a complete syntactic restructuring of the sentence (a complex sentence has been replaced with a simple one).

Example 2

Source: *Property Legal title to the property of the Partnership shall be held in the name of or in such other name or manner as the Partners shall determine to be in the best interest of the Partnership.*

MT output in LT: *Teisinė nuosavybės teisė į Partnerystės turtą įgyjama Partnerystės vardu arba kitu vardu ar būdu, kurį Partneriai nustato kaip geriausiai atitinkantį Partnerystės interesus.*

Human post-editing LT1: *Nuosavybės teisė į Partnerystės turtą įgyjama Partnerystės vardu arba kitu vardu ar būdu, kurį Partneriai nustato kaip geriausiai atitinkantį Partnerystės interesus.*

MT output in UA: *Власність Правове право власності на майно Товариства належить від імені або іншим ім'ям чи способом, який Партнери визначатимуть як найкращий в інтересах Товариства.*

Human post-editing UA1: *Майно. Право власності на майно Товариства набувається на ім'я Товариства або в інший спосіб, визначений Партнерами як такий, що якнайкраще відповідає інтересам Товариства.*

In Lithuanian, the phrase *Property Legal title to the property* was machine translated as *Teisinė nuosavybės teisė*. In this case, the English phrase *Property Legal title to the property* can be rendered to Lithuanian as either *Nuosavybės teisė* or *Teisinė nuosavybė*. The machine translation output shows redundant words with the same meaning: adjective *teisinė* (en. *legal*) and noun *teisė* (en. *right, title*); therefore, the first word – adjective *teisinė* – was removed in the post-edited text. This change is classified as a lexical change – multiword expression change without any change in the meaning).

In Ukrainian, the construction *Property Legal title to the property* was machine translated as *Власність Правове право власності*, which is incorrect both linguistically and logically (unreasonable repetition of the same or same-root or similar words). The post-editor used the construction *Майно. Право власності*, applying the noun *майно* ("речі, які комусь належать за правом власності" / transl.: items owned by someone under the right of ownership (Busel, 2005, p. 637)) as a synonym for the noun *власність* ("майно, належне кому-, чому-небудь" / transl.: property belonging to someone or something (Busel, 2005, p. 194)). Such transformation can be classified as a stylistic change with no meaning change.

Example 3

Source: *It is the intent of the majority partner to begin making this transfer after ___ years.*
MT output in LT: *Daugumos partneris ketina šį pervedimą pradėti vykdyti po ___ metų.*
Human post-editing LT1: *Pagrindinis partneris ketina šį pervedimą pradėti vykdyti po ___ metų.*
Human post-editing LT2: *Didžiąją ūkinės bendrijos kapitalo dalį turintis partneris ketina šį pervedimą pradėti vykdyti po ___ metų.*

MT output in UA: *Мажоритарний партнер має намір почати здійснювати цю передачу через ___ років.*
Human post-editing UA1: *Мажоритарний партнер має намір почати цю передачу через ___ років.*
Human post-editing UA2: *Партнер, який володіє контрольним пакетом акцій, має намір розпочати цю передачу через ___ років.*

In the Lithuanian machine translated version, the noun phrase *the majority partner* was rendered as *daugumos partneris*. Such rendering is not correct in Lithuanian. Therefore, post-editors made changes in this instance and rendered it as *pagrindinis partneris* (en. *main partner*) and *didžiąją ūkinės bendrijos kapitalo dalį turintis partneris* (literal en. *large partnership capital part owning partner*). On the one hand, an adjective change occurred, and on the other, a multiword expression change was introduced, replacing the literal machine translated version with a descriptive phrase.

The Ukrainian, the post-editors made different decisions about correcting the machine translation. In the first case the terminological structure *the majority partner* was not changed, while in the second case the post-editor used descriptive translation, as construction *мажоритарний партнер* is not common in Ukrainian legal discourse. This case is classified as a noun-phrase structure change with the sense preserved, together with reordering for a better location of the parts of the utterance.

Concerning the survey results of post-editors about the machine-translated text they post-edited, none of them selected the option "agree" to any of the 7 provided statements. Most of the answers were within the range of partially disagree and partially agree. For statements 1, 2 and 3, the post-editors mostly chose the option "partially disagree". For statements 4, 5 and 6, the answers ranged between "neither agree nor disagree" and "disagree". For statement 7, the only answer option selected by the post-editors was "disagree".

When asked whether it was difficult to post-edit the legal text, post-editors answered that it was quite a difficult task, especially when they felt that they would rather have to translate the text from scratch rather than post-edit it. Post-editors shared that some segments and/or sentences appeared to be correct at first glance, both grammatically and semantically. However, inconsistency of the terms used was among the issues they struggled with. In their opinion, post-editing took a relatively long time. They indicated that it took them from one to three hours to do the post-editing task.

6. Discussion.

As mentioned in the literature review, post-editing quality depends on the quality and precision of machine translation output. However, the post-editors' competence is of great importance. Studies with various language pairs generally report the usefulness of using machine translation and post-editing of legal texts, making note of reduced temporal and cognitive efforts. However, in certain cases, especially involving under-resourced languages or language pairs or unrelated languages, research provides evidence that post-editing legal texts actually requires more cognitive effort than other types of texts. The issues to be taken into consideration while post-editing machine translated legal texts are many, including consistency in the use of terminology, non-matching legal jargon, complex syntax, to mention but a few. Because of the lower quality of machine translation output, translators may be reluctant to do post-editing of legal texts and may prioritize translation from scratch (Levanaitė, 2021).

In this study, an initial objective was to analyze and compare post-editing actions made by professional translators in machine translated legal contracts in two languages pairs involving under-resourced languages, namely Lithuanian and Ukrainian. It was hypothesized that the main post-editing actions will involve issues related to noun phrases, or in other words, terms and legal jargon. However, we also aimed at sketching the tendencies in the variety of post-editing actions in general as well as the post-editors' opinions on the difficulty of the task of legal text post-editing.

Regarding the main issues to be addressed in a machine translated legal text, our study findings, as expected, are consistent with those obtained in other studies where terms and legal jargon are reported to be most problematic. Inconsistency of specialized legal terminology was found to be pronounced in machine translation from Hungarian to English (Kovács, 2022) as well as from English to Chinese (Cui et al., 2023). On the other hand, a study by Killman and Rodriguez-Castro (2022) in the English to Spanish language pair reported that machine translation was especially useful at the terminological and phraseological levels.

Our data also reflect the findings of Blain et al. (2011) that it is the changes in the noun phrases that are most prominent in post-editors' actions taken. In addition, we also observed some other types of post-editing actions, including verb phrase, co-reference and reordering changes. No post-editor introduced errors were observed in our sample, which may point to the fact that either the post-editors were experienced and skilled and/or the machine translation tool used was relatively high quality requiring less cognitive effort on the part of post-editors. This interpretation needs to be further tested with a larger sample and a different research design.

What concerns the post-editor's opinions about the difficulty of the post-editing task, the findings of our study may be seen to be in line with those obtained in a study by Cui et al. (2023) who found that legal texts were rated by semi-professional translators as requiring more cognitive effort than other types of texts. Although different in research design and aims of the study, our research results may also substantiate this fact since the post-editing task was reported to be a difficult task where human translation would have been prioritized. This preference may have been likely related to inconsistency of the terms provided by machine translation, which resulted in spending more time looking for the needed translation and exerting more cognitive effort in an attempt to use the terms in the consistent manner.

Post-editor's skill and effort are the main factors affecting post-editing speed (Koehn & Germann, 2014). In a study on post-editing a news text machine translated using 4 different systems by 4 post-editors, a great variance in post-editing speed was observed. The difficulty of the text to be post-edited could be among the reasons for slower productivity. In our case, the 6 Lithuanian post-editors indicated that it took them between 1 and 3 hours to complete the post-editing task, which shows huge difference. Although we did not measure the productivity in words per second (as in a study by Koehn & Germann, 2014), our results of great differences in the post-editing time may be an indirect indication of an association between the difficulty of the text and the time required. However, the finding should be tested in further studies.

7. Conclusions.

This study presents findings of a comparative analysis of post-editing actions in a text of a legal contract machine translated from English to Lithuanian and from English to Ukrainian. The study found a tendency of noun phrase changes in terms of lexical units in both languages. This finding correlates well with the post-editors' expressed opinions that they were mostly struggling with terms and with the fact that the post-edited task was a legal

text, which is by and large loaded with legal terms and semi-fixed phrases. General purpose machine translation, which we used in this study, may not be accurate enough for legal purposes. Using and/or training domain specific machine translation engines with legal domain data would result in higher accuracy outputs. The results also point to the fact that having sufficient resources, namely parallel corpora, in a particular language pair is of crucial importance for higher quality machine translation output facilitating the post-editor's work.

The limitation of this study is the scope of the study limited by the type and domain of the text chosen for the analysis. Future studies may be focused on replication of the analysis with different types and amounts of the source text and with different source or target languages in a pair with Lithuanian and Ukrainian or between Lithuanian and Ukrainian. Comparative analyses may be performed to analyse correlations between post-editing actions and other variables like post-editing time, quality, etc.

R e f e r e n c e s

Arterchuk, T. (2014). Mashynnyi pereklad v opratsiuvanni studentamy-pravnykamy nimetskomovnyh yurydychnyh tekstiv [Machine Translation of German-Language Legal Texts Processed by Law Students]. *Pedagogical sciences: theory, history, innovative technologies*, 3(37), 264–271. [in Ukrainian].

Blain, F., Senellart, J., Schwenk, H., Plitt, M., & Roturier, J. (2011). Qualitative analysis of post-editing for high quality machine translation. In Proceedings of Machine Translation Summit XIII: Papers, Xiamen, China. Retrieved October 25, 2023, from the website https://aclanthology.org/2011.mtsummit-papers.17

Busel, V. (ed.). (2005). Velykyi tlumachnyi slovnyk suchasnoii ukrainskoi movy [Great Explanatory Dictionary of Modern Ukrainian Language]. Kyiv: Perun. [in Ukrainian].

Buşila, A. (2017). The issue of translating legal doublets in notarial acts from English into Romanian. *Lingua Legis, 25*, 123-136. Retrieved October 4, 2023, from the website http://lingualegis.ils.uw.edu.pl/index.php/lingualegis/article/view/26

Castilho, S. (2016). Measuring acceptability of machine translated enterprise content. PhD Thesis, Dublin: Dublin City University. Retrieved from March 5, 2023, from the website https://doras.dcu.ie/21342/1/Castilho(2016)_Acceptability_of_Machine_Translated_Enterprise_Content_PhD Thesis.pdf

Chéragui, M. A. (2012). Theoretical overview of machine translation. Proceedings of the 4th International Conference on Web and Information Technologies, Sidi Bel Abbes, Algeria, 160–169.

Cui, Y., Liu, X., & Cheng, Y. (2023). A comparative study on the effort of human translation and post-editing in relation to text types: an eye-tracking and key-logging experiment. *SAGE Open, 13*(1). doi: https://10.1177/21582440231155849

Dugast, L., Senellart, J., & Koehn, P. (2007). Statistical post-editing on SYSTRAN's rule-based translation system. In Proceedings of the Second Workshop on Statistical Machine Translation, Prague, Czech Republic. Association for Computational Linguistics, 220–223. Retrieved May 22, 2023, from the website https://aclanthology.org/W07-0732/

Flanagan, M. (1994). Error classification for MT evaluation. In Proceedings of the First Conference of the Association for Machine Translation in the Americas, Columbia, Maryland, USA.

Gortych-Michalak, K. (2017). In search of equivalents in legal translation: a parametric approach to the comparison of legal terminology in Polish and Greek. Poznań: Wydawnictwo Naukowe Contact.

Grabar, N., Kanishcheva, O., & Hamon, T. (2018). Multilingual aligned corpus with Ukrainian as the target language. SLAVICORP, Prague, Czech Republic. Retrieved November 11, 2023, from the website https://shs.hal.science/halshs-01968343/document

Heller, T., & Zoyirova, D. (2022). English is difficult: modest proposals that can drastically improve the quality of legal English composition. *Pravnik: Revija za Pravno Teorijo in Prakso, 77*(7–8), 319–346. Retrieved November 12, 2023, from the website https://plus.cobiss.net/cobiss/si/sl/bib/157650179

Karaban, V. (2003). Anhlo-Ukraiinskyi yurydychnyi slovnyk [English-Ukrainian legal dictionary]. Vinnytsia: Nova Knyha. [in Ukrainian].

Kasperavičienė, R., Motiejūnienė, J., & Patašienė, I. (2020). Quality assessment of machine translation output: cognitive evaluation approach in an eye tracking experiment. *Texto livre: linguagem e tecnologia, 13* (2), 1–16. doi: https://10.35699/1983-3652.2020.24399

Kenton, W. (2022). *Investing Law*. Retrieved November 10, 2023, from the website https://www.investopedia.com/terms/u/uniform-partnership-act-upa.asp

Killman, J., & Rodríguez-Castro, M. (2022). Post-editing vs. Translating in the legal context: quality and time effects from English to Spanish. *Revista de Llengua i Dret, Journal of Language and Law, 78*, 56–72. doi: https://10.2436/rld.i78.2022.3831

Koehn, P., & Germann, U. (2014). The impact of machine translation quality on human post-editing. In Proceedings of the EACL 2014 Workshop on Humans and Computer-assisted Translation (pp. 38-46). Retrieved November 12, 2023, from the website https://aclanthology.org/W14-0307/

Kovács, T. (2022). Human and machine translation: a comparative analysis of neural machine- and human-translated EN-HU and HU-EN legal texts. *Porta Lingua. 2022*(1). 49–57. doi: https://10.48040/PL.2022.1.5

Koroliova, T., Zhmaieva, N., Kolchah, Yu. (2020). Postredahuvannia pry mashynnomu perekladi [Machine Translation Post-editing]. *Scientific Bulletin of PNPU named after K.D. Ushynskyi*, 30, 102–119. doi: https://doi.org/10.24195/2616-5317-2020-30-7. [in Ukrainian].

Levanaitė, K. (2021). Lietuvos vertimo rinkos dalyvių požiūris į mašininį vertimą ir postredagavimą. Vertimo studijos, 14, pp. 22–39. doi: https://10.15388/VertStud.2021.2

Matvieieva, S. (2020). Linguistic Reconstruction of Cognitive Terminological Structure REFUGEE / БІЖЕНЕЦЬ in the English-Ukrainian legal corpus. Kyiv: NPU im. M.P. Dragomanova.

Mileto, F. (2019). Post-editing and legal translation. *Digital Humanities Journal, 1*(1). doi: https://10.21814/h2d.237

Omazić, M., & Šoštarić, B. (2023). New resources and methods in translating legal texts: machine translation and post-editing of machine-translated legal texts. *Language(s) and Law*. Osijek: Pravni fakultet Sveučilišta Josipa Jurja Strossmayera u Osijeku, 2023. str. 71–84.

Park, C., Yang, Y., Park, K., & Lim, H. (2020). Decoding strategies for improving low-resource machine translation. *Electronics, 9*(10), 1562. MDPI AG. doi: https://10.3390/electronics9101562

Povilaitienė, M., & Kasperė, R. (2022). Machine translation for post-editing practices. *Науковий Часопис Національного педагогічного університету імені М. П. Драгоманова. Серія 9. Сучасні тенденції розвитку мов = Scientific journal of national pedagogical Dragomanov University. Series 9. Current trends in language development*. Kyiv: National Pedagogical Dragomanov University. doi: https://10.31392/NPUnc-series9.2022.24.04

Sosoni, V., O'Shea, J., & Stasimioti, M. (2022). Translating law: A comparison of human and post-edited translations from Greek to English. *Revista de Llengua i Dret, Journal of Language and Law, 78*, 92–120. doi: https://10.2436/rld.i78.2022.3704

Utka, A., Amilevičius, D., Krilavičius, T., & Vitkutė-Adžgauskienė, D. (2016). Overview of the development of language resources and technologies in Lithuania (2012–2015). In *Human Language Technologies–The Baltic Perspective* (pp. 12–19). IOS Press. doi: https://10.3233/978-1-61499-701-6-12

Vilar, D., Xu J., D'Haro L. F., & Ney H. (2006). Error analysis of machine translation output. In Proceedings of the 5th International Conference on Language Resources and Evaluation (LREC'06), Genoa. Retrieved November 5, 2023, from the website https://aclanthology.org/L06-1244/

Yankovskaya, L., Tars, M., Tättar, A., & Fishel, M. (2023, March). Machine Translation for Low-resource Finno-Ugric Languages. In The 24rd Nordic Conference on Computational Linguistics, Tórshavn, Faroe Islands. Retrieved November 5, 2023, from the website https://aclanthology.org/2023.nodalida-1.77/

Бібліографічний опис:

Анотація

Юридичний переклад вимагає спеціального використання мови. Тому його часто вважають одним із найважчих різновидів перекладу. Постредагування машинного перекладу набуває все більшого поширення. Навіть у мовах, які не можуть похвалитися великою кількістю ресурсів, переклад текстів певних жанрів рідко проводиться без використання машинного перекладу, оскільки якість результатів, здається, наближається до якості перекладу, виконаного людиною. Ця розвідка є спробою аналізу та порівняння рішень постредагування юридичних договорів, перекладених з англійської на литовську та українську мови. Юридичні договори є частиною юридичного дискурсу.

Мова таких текстів характеризується архаїчністю, складністю, культурною залежністю, є зазвичай насиченою формульними виразами, спеціалізованими термінами та латинізмами. Як для технічного типу мовлення для юридичного перекладу може використовуватися машинний переклад. Висновки дослідження ґрунтуються на результаті оцінки рішень постредагування машинного перекладу, прийнятих професійними перекладачами. Отримані дані демонструють і підтверджують результати попередніх досліджень інших мов, згідно з якими основними тенденціями є такі: зміни в іменникових та дієслівних конструкціях в обох мовах, кореференція та зміна порядку частин. Подальше вивчення визначеної проблеми дозволить доповнити (таргетувати й деталізувати) типологію дій із постредагування з урахуванням специфіки балтійських (литовської) та слов'янських (української) мов.

Ключові слова: машинний переклад, постредагування, дії з постредагування, юридичні договори.