# MT and Legal Translation: applications in training

**Suzana Noronha Cunha**                               scunha@iscap.ipp.pt
CEOS.PP, ISCAP, Polytechnic of Porto, S. Mamede de Infesta, 4465-004, Portugal

**Abstract**

This paper investigates the introduction of machine translation (MT) in the legal translation class by means of a pilot study conducted with two groups of students. Both groups took courses in legal translation, but only one was familiarised with post-editing (PE). The groups post-edited an extract of a Portuguese company formation document, translated by an open-access neural machine translation (NMT) system and, subsequently, reflected on the assigned task. Although the scope of the study was limited, it was sufficient to confirm that prior exposure to machine translation post-editing (MTPE) did not significantly alter both groups' editing operations.

The pilot study is part of a broader investigation into how technology affects the decision-making process of trainee legal translators, and its results contributed to fine-tuning a methodological tool that aims to integrate MTPE procedures in an existing process-oriented legal translation approach developed by Prieto Ramos (2014). The study was repeated this year. This time both groups of trainees were introduced to and used the tool in class. A comparison of both studies' results is expected to provide insight onto the productive use of MTPE in other domain-specific texts.

## 1. Introduction

The advent of the World Wide Web and the emergence and evolution of computer-based tools has been changing the way translation is done for more than twenty years. More recently, machine translation (MT) and neural machine translation (NMT) brought about the second major technological shift in the translation industry (Doherty, 2016). These systems appeal to a wide range of users thanks to their ability to provide instant translation of large amounts of information with high-quality output in numerous language pairs. Furthermore, most major NMT providers now offer free versions of their systems that are accessed and used globally for translation of numerous types of texts, in numerous domains and for multiple purposes.

Translation trainees use open-access NMT systems, unaware of the systems' limitations and before acquiring competence in the particulars of in-domain translation. In legal translation, lack of familiarity with the subject matter, the textual genre, and intersystemic transfer barriers may cause them to miss errors or, conversely, overcorrect them. The increasing sophistication and "human-like qualities" of NMT make it even more difficult for trainees to flag errors (Yamada, 2019). In light of the above, the challenge for the translation trainer in the 2020s lies in providing guidance through the automated process of translating in-domain texts.

While the integration of technology in translation training is consensual, to our knowledge, not many studies focused on the integration of post-editing processes in domain-specific translation processes, which paved the way for investigating the specificities of implementing them in the legal translation class. The first step was to carry out a pilot study with 2 groups of participants, of which only one was formally trained in MTPE. The study and preliminary results are discussed in more detail in the present paper. In general, and although the

sample was not representative, editing operations were similar in both groups, both reveal special caution with terminology and no one is consistent when post-editing format, punctuation, or mechanic grammatical errors.

These results reinforced the conviction that a methodological tool could be useful to systematize the processes of post-editing the NMT output of legal documents. Prieto Ramos (2011, 2014) legal translation integrative methodology was the framework for the first proposal. This problem-solving model integrates the legal and linguistic dimensions in a 4-stage process comprising analysis of skopos and macro-contextualization; source text analysis; transfer and target text production; and revision (2014).

Our proposal reorganized the second and third stages of Prieto Ramos' model in an attempt to reflect a relevant shift in the translation workflow: that translators no longer work from decoding the source to transferring and recoding the same intent and information in the target text. Today, translators work with, at least, 3 texts simultaneously – the source, the MT output and the target - and, after considering the brief and the communicative situation, focus not on analyzing the source text but the NMT output, thus performing PE rather than producing a translation from scratch.

This shift needs to be addressed actively and openly in the specialised translation class, so as to make it very clear that, although MTPE apparently reduces the relevance of human intervention in the process by dislocating it to the end of the process, it entails as much competence in the legal domain and discourse as translation from scratch does. Post-editing skills are so critical in the process of comparing texts in different languages and assessing them for accuracy and naturalness, that it should be presented as a "validation" activity rather than a revision and proofreading procedure (Pym and Torres-Simon, 2021).

The following sections of the present paper discuss the relevant literature, describe and discuss the pilot study and preliminary results, and introduce future work on the development of the integrative methodological tool.

## 2.   Related Work

Ten years ago, Pym alerted for the need to define new skill-sets for the translator in the MT age (2013). Traditionally, translation had been a "generative" activity consisting mostly in identifying solutions to translation problems, whereas today it requires selecting the adequate solution to a specific communicative situation among numerous possible candidates. This, in his words, is "a very simple and quite profound shift", unsettling for both trainers accustomed to the traditional process of transfer between source and target texts, and trainees provided with many ready-made solutions and not enough insight on how to address them.

More recently, Rodríguez de Céspedes discussing the implications for training of the shift of translation from human to machine, also warned that translator intervention at later stages of the process added a new dimension to the human cognitive act of translating (2019). Furthermore, although it is getting harder to distinguish between human and machine translation, because MT systems are trained and fed with human translations (Doherty, 2016), human translation remains the standard for quality evaluation of MT output, and the activity of editing and correcting MT output is carried out by humans (ISO 18587:2017).

Assuming that post-editing has become central in translation practice[1] and that it requires both specific competence and translation competence (Yamada, 2019), to prepare trainees for performing the roles of post-editor and translator interchangeably (Vieira, 2019) it is imperative that MTPE is openly integrated in the specialized translation class.

---

[1] For an overview of post-editing as an increasingly central practice in the translation field and of research on MTPE, see Koponen (2016).

Legal translation, as any other specialized language, requires knowledge of the domain and familiarity with its terminology (Wilss, 1996). However, unlike medicine or engineering, legal terminology is system-bound (Cao, 2007), which means that the concepts of source and target equivalents do not fully coincide, but rather overlap partially, leading to incongruity between what can be called legal functional equivalents. Incongruity is compensated by established translation strategies – e.g. borrowing, paraphrase, literal and functional equivalence – and constitutes a real challenge in legal translation: balancing accurate transfer of legal concepts with naturalness and target reader expectations. It also helps explain why legal translations are often "hybrid texts" that read like and look like translations: their purpose is not to replace but to reveal the source legal entity "through target language knowledge systems" (Biel, 2009).

Prieto Ramos' holistic model for legal translation competence (2014) systematizes the translation process workflow under the overarching procedural/methodological competence, combining legal and linguistic subcompetences. Such a model is flexible enough to accommodate specificities of legal translation and post-editing procedures. ISO 18587 requires that post-editors possess general knowledge of MT and the typical errors[2] it makes (2017), much in the same way as the European Master's in Translation framework recommends that translators know the basics of MT systems (2022).

To carry out the "observable operations performed by the translator over pre-existing text" – the 4 "editing actions" of inserting, deleting, replacing and moving (Carmo, 2017) –, post-editing usually involves error typology quality evaluation. The error typology approach is useful in identifying and fixing errors and it should be flexible to allow for the addition or deletion of error categories and sub-categories, according to the features of each text or the requirements of a translation task. The Language Service Provider (LSP) industry provides various guidelines for error evaluation, such as those of TAUS, the Translation Automation User Society (2017). TAUS and the German Research Center for Artificial Intelligence (DFKI) have harmonized their respective DQF (Dynamic Quality Framework) and MQM (Multidimensional Quality Metrics) into one DQF-MQM framework (Valli, 2015). A simplified adaptation of this harmonized framework was used to analyse the extracts post-edited by the participants in the pilot study.

## 3.   The Pilot Study

As stated above, this paper investigates whether prior introduction to PE is enough for productive use of MTPE in the legal translation class. The pilot study is the initial stage of a larger project that aims to integrate adapted post-editing procedures and error categories in an existing problem-solving model for legal translation. The study tries to answer the following questions:

Q1. Do students with prior training in PE perform better in identifying or fixing MT errors?

Q2, Are there noticeable differences in students' lexical, grammatical or formatting edits?

Q3. Do students with PE training follow a different process or method?

The author evaluated the participants' post-edited texts using a simplified version of the Dynamic Quality Framework Knowledge Base (TAUS, 2017). First, the MT output was assessed and the author selected twenty-nine recommended edits covering 4 of the framework's high-level error types: Accuracy, Fluency, Style and Design. Each category was then divided into granular error types to facilitate tracking them in the edited texts of the

---

[2] Kenny (2022) enumerates 4 typical non-human errors: linguistic ambiguity, non-isomorphism, discontinuous dependencies and non-compositionality.

two groups. These were analysed to verify which errors were detected and fixed and whether there were relevant differences in each group's edits. Finally, the evaluated data were correlated with the answers to pre-questionnaire questions *7. Higher education studies in translation* and *8. Professional experience in translation*, and post-questionnaire questions *4. How do you rate the quality of the MT output?* and 5. *Your alterations to the MT text were…[few, some, many (in number and relevance)]* to compare participants' profiles with editing behavior. Questions 6 and 7 of the post-questionnaire, concerning error categories detected and MT output usefulness in fixing them, were also correlated with the evaluation results. The data was used to answer the first 2 questions,

Answers to question *9. Briefly describe the process you followed to carry out this post-editing task* (post-questionnaire) were analysed to check differences in the groups' PE procedures, and correlate participants' perceptions with the evaluator's analysis, to try and draw some insight for the third research question.

## 3.1.    Participants' profile

Twenty-one students attending courses in legal translation at ISCAP, the Accounting and Business School of the Polytechnic Institute of Porto, participated in the pilot study on two consecutive days, in May 2022. The post-editing task was carried out by fourteen students of the Master's in Specialised Translation and Interpreting (MSTI), on day one, and by seven students from the Post-graduation in Specialised Translation and Translation Tools (PGSTTT), on day two. ISCAP's MSTI is a member of the EMT network and more than half the students take the master's following completion of the degree in Management Assistance and Translation at the same institution. The PGSTTT, in turn, is a lifelong learning, one-year course preferred by graduates wanting to start a career in translation and professional translators in search of updating.

The two groups (henceforward, the Master's – G1 and the Post-graduation – G2 groups) were not homogenous: G1 students were generally younger and less experienced. In the master's group, there were 5 males and 9 females, and all 7 G2 participants were females. English language proficiency was evenly distributed between C1 and C2 in G1; 70% of G2 participants indicated C2 as their level. As stated before, only G2 participants attended a module in PE.

## 3.2.    Materials and procedures

The task consisted in post-editing a 330-word extract of the Articles of Association (AoA) of EDP, a Portuguese public limited company (*sociedade anónima*) from the energy sector. The extract consisting of the initial three articles had been previously translated into British English in the free version of DeepL[3], a German-based NMT system that used the existing dataset of the translation search engine Linguee[4]. Two reasons motivated the selection of the document: translation of company formation documents is frequent in Portugal and this type of text that can be easily accessed online, in both Portuguese and English. Participants were informed of and agreed with the content and purpose of the experiment.

British English was the preferred variant and participants were instructed to post-edit accordingly. Translation into L2 has been used for pedagogic purposes in countries with languages of limited diffusion (T. Pavlovic 2013; Fonseca, 2015) or where the hegemony of English creates a context where all other languages are *minority languages*. In these countries, such as Portugal, Germany, Spain and Brazil, professional translation into

---

[3] https://www.deepl.com/translator
[4] https://www.linguee.com/

English (L2) is an established reality that must somehow be dealt with in training (Kiraly, 2000; N. Pavlovic, 2007; Vigier, 2016; Ferreira et al., 2018). Post-editing in L2 is one of the exercises to address this issue[5].

The translation brief required post-editing for human translation quality (TAUS, 2016), since the purpose of the target text was the internationalisation of the company. In order to address the research questions, the brief did not include post-editing guidelines. In both groups, students were familiar with the extract's branch of law and text genre, and had translated examples of company AoA. They were also aware that documentary translation (Nord, 2016) is advisable when legal documents are translated for information of the target audience. At the time of the experiment, differences between the USA and the UK legal systems and how these reflect in legal discourse had been discussed in class.

On the days of the experiment, each group of participants filled in the pre-task questionnaire (9 questions) to collect data on educational and professional backgrounds and their perceptions of the benefits and limitations of MT. They then carried out the post-editing task followed by a post-task questionnaire, comprising 10 questions, in which they stated familiarity with NMT systems, perceptions on the quality and usefulness of the NMT output, degree of difficulty and satisfaction with the task and provided a brief description of how they carried out the post-editing task. There were no time constraints, and both groups took approximately the same time to complete the three tasks.

## 3.3.    Analysis

On each day of the experiment, after completion of the post-questionnaire, participants emailed the post-edited texts to the author. The anonymised documents were downloaded and analysed using the MS Word Track Changes feature. Assessment of variation in the number and type of errors detected and fixed by each group[6] was supported in the Error Typology Best Practice Guidelines (TAUS, 2017).

TAUS recommends a limited number of error categories for quality evaluation and describes the four most commonly used: Language, Terminology, Accuracy and Style (2017). For evaluations that "seek to understand in detail the nature or cause of errors" (2017) a more detailed and flexible typology is advisable and the TAUS Dynamic Quality Framework Knowledge Base is referenced[7]. The quality error typology evaluation template sets 7 high-level error types, each divided into granular error types. From those, the author selected 4 high-level error types – Accuracy, Fluency, Design and Style – divided them in granular error types and used them to label the 29 recommended edits in the MT output. These were, then, tracked in G1 and G2 post-edited texts.

| High-level error type | Granular error type | No. of edits |
|---|---|---|
| Accuracy | Mistranslation | 5 |
|  | Under-translation | 3 |
|  | Untranslated text | 1 |
|  | Inconsistency | 3 |
| Fluency | Syntax | 2 |

---

[5] Compiling small comparable corpora to compensate for lesser fluency, grammatical accuracy and phraseology, while providing context for analysis of incongruity in intersystemic translation into L2 (Scott, 2012; Vigier, 2016) are other exercises carried out in the legal translation class.

[6] In G2, there was 1 invalid contribution. Only 6 post-edited extracts were evaluated, although 7 answers to pre and post-questionnaires were validated.

[7] TAUS launched the first attempt at an industry-developed standard for translation quality evaluation in 2011, with a dynamic quality framework that lived up to today's translation quality requirements that change depending on content type, purpose and audience (Gorog, 2014).

| | | |
|---|---|---|
| | Grammar | 2 |
| Design | Formatting | 5 |
| | Conventions | 1 |
| Style | Awkward | 3 |
| | Unidiomatic | 4 |
| TOTAL | | 29 |

Table 1. TAUS Error Categories (adapted).

The 4 categories were selected taking into consideration the small size of the extract and the relevance of accurate representation of meaning in the translation of legal texts (Sarcevic, 2000; Prieto Ramos, 2014). Because "Legal terminology is the most visible, […/…] and it is also one of the major sources of difficulty in translating legal documents" (Cao, 2007) all the examples included in the Accuracy category are terms. This is the most detailed category to which a granular error subcategory was added for Inconsistency in term usage. Fewer examples and sub-categories for Fluency are also due to the small size of the extract. Formatting illustrates the category's errors in NMT output. The Style category aims at highlighting content that, although grammatical, does not reflect the legal style of the target system. The category of Style is not analysed here.

| Source text | NMT output | Recom. edit | PE text G1 | PE text G2 |
|---|---|---|---|---|
| 1. Contrato de Sociedade | Memorandum and Articles of Association | Articles of Association | Error undetected:7<br>Error fixed:**2**<br>Error introduced:**5** | Error undetected:**4**<br>Error fixed:**2**<br>Error introduced:**0** |
| 2. simples deliberação | simple resolution | ordinary resolution | Error undetected:**12**<br>Error fixed: **0**<br>Error introduced:**2** | Error undetected:**1**<br>Error fixed:**5**<br>Error introduced:**0** |
| 3. agências | agencies | branches | Error undetected:**12**<br>Error fixed:**0**<br>Error introduced:**2** | Error undetected:**4**<br>Error fixed:**0**<br>Error introduced:**2** |
| 4. proceder | proceed | define/ formulate | Error undetected:**13**<br>Error fixed:**1**<br>Error introduced: **0** | Error undetected:**6**<br>Error fixed **0**<br>Error introduced:**0** |
| 5.c) assegurar…<br>   d) assegurar… | c) ensure…<br>d) ensure… | c) ensure…<br>d) undertake… | Error undetected:**14**<br>Error fixed:**0**<br>Error introduced:**0** | Error undetected:**5**<br>Error fixed:**0**<br>Error introduced:**1** |

Table 2. Error Type evaluation of Accuracy (mistranslation).

In all five mistranslations over 70% of G1 and G2 participants fail to identify the error in the NMT output, probably because they lack domain competence (1 and 2) and efficient PE strategies (3 and 4). Apart from the slightly higher number of 'error introduced' by G1, there are no significant differences in PE procedures. In *deliberação simples*, an ellipsis, typical of expert communication, occurs. Undetected by the untrained NMT system, it is only identified and fixed by G2 participants. However, voting procedures in general meetings and the title of the incorporation document had been discussed in both groups' classes.

Literal translation of terms is another typical issue of online NMT output (3 to 5). A second ellipsis occurs in *agências*, short for *agências bancárias* (branches of banks), mistranslated by the cognate **agencies**. The term is followed by a synonym — *delegações* (**branches**, in the NMT output) which may explain why the 4 participants who identified the error failed in properly editing it, either replacing **branches** by **delegations** (literal mistranslation in the context) or by **subsidiaries** (a collocate of **branches**, though differing in meaning). The verbs *proceder* and *assegurar* (4 and 5) are also translated by cognates that do not represent the same

16

meaning in the context of Article 3 of the AoA. The fact that the same verb is repeated at the beginning of consecutive sentences (5) may also contribute to failure in error identification.

| Source text | NMT output | Recomm. edit | PE text G1 | PE text G2 |
|---|---|---|---|---|
| 6. sede | head office | registered office | Error undetected:**10** Error fixed:**2** Error introduced:**2** | Error undetected:**6** Error fixed:**0** Error introduced:**0** |
| 7. participações | participations | participating interests | Error undetected:**12** Error fixed:**1** Error introduced:**1** | Error undetected:**4** Error fixed:**1** Error introduced:**1** |
| 8. suprimentos | loans | shareholders' loans | Error undetected:**9** Error fixed:**0** Error introduced: **5** | Error undetected:**4** Error fixed **0** Error introduced:**2** |

Table 3. Error Type evaluation of Accuracy (under-translation).

As in the previous subcategory, between 70% to 80% of participants fail to spot errors. Differences in PE texts in G1 and G2 are limited to a slightly higher number of errors identified by G1. Consequently, G1 also introduce a higher number of errors. Procedures adopted to edit *suprimentos* – **loans** illustrate this: when the term **loans** is repeated in the following sentence in the NMT output: "4. The company may provide services and grant **loans** and other forms of **loans**…", participants would likely have consulted the ST to check the 2 different terms in Portuguese: *suprimentos* and *empréstimos*. When in doubt, G1 tend to try and fix the error, whereas most G2 participants prefer to accept the NMT suggestion.

| Source text | NMT output | Recomm. edit | PE text G1 | PE text G2 |
|---|---|---|---|---|
| 9. sociedade anónima | sociedade anónima | (public) limited liability company | Error undetected:**0** Error fixed:**12** Error introduced:**2** | Error undetected:**0** Error fixed:**5** Error introduced:**1** |

Table 4. Error Type evaluation of Accuracy (untranslated text).

All participants identified the untranslated term, although both in G1 and G2 some chose to keep it in the SL between inverted commas and 3 fail to identify the type of company correctly.

| Source text | NMT output | Recomm. edit | PE text G1 | PE text G2 |
|---|---|---|---|---|
| 10. denominação denominação | name shall be known | name the Company's name is | Error undetected:**9** Error fixed:**4** Error introduced:**1** | Error undetected:**4** Error fixed:**2** Error introduced:**0** |
| (6). sede 11.sede social | head office registered office | registered office registered office | Error undetected:**10** Error fixed:**0** Error introduced:**4** | Error undetected:**6** Error fixed:**0** Error introduced:**0** |
| 12. objecto objecto objecto social objecto social | object object corporate object corporate object | object object (Company's) object (Company's) object | Error undetected:**0** Error fixed:**0** Error introduced: **3** | Error undetected:**0** Error fixed **0** Error introduced:**2** |

Table 5. Error Type evaluation of Accuracy (inconsistency).

Table 5 illustrates variation in the translation of the terms *denominação*, *sede* and *objecto*, occurring twice: in the Section's Title and in the text of the articles. Although there is no obvious reason not to repeat the term the second time it occurs, the majority in both groups fails to detect the inconsistency in the NMT text. The ellipsis referred to above – the omission

of the adjective *social* in 11 and 12 – may account for the inconsistency in the NMT text, but does not explain why **name** is occasionally replaced by **denomination**, **registered office** by **head office** or **object** by **aim** or **purpose** in the PE texts. As with previous examples, these terms had been dealt with in class in the context of company law. A possible explanation is that some participants introduced error in the terms' second occurrences because they were concerned with terminological consistency, and thus repeated the mistranslated term of the first occurrences[8].

As regards the category of Accuracy as a whole, we observed that typical NMT errors, with the exception of inconsistency, are not detected by the majority of G1 and G2 participants. While G1 carry out more PE operations, G2 participants seem slightly better at handling terminology, thus not making it possible to state whether it is domain competence or prior training in PE to make a difference.

In the category of Fluency, two examples of syntax and two of grammar editing are analysed to illustrate how the groups handled post-editing long/complex sentences when compared with mechanical editing procedures. Numbers 1 and 2 of Article 3 were chosen as examples of a long sentence (n.1) and a syntactically complex one (n.2).

| Article 3 |
| --- |
| 1. The object of EDP is to promote, stimulate and manage, directly or indirectly, undertakings and activities in the energy sector, both at national and international level, with a view to increasing and improving the performance of all the companies of its group. |
| 2. EDP, in the development of its corporate object, must, in relation to the companies of its group* |
| a) proceed to the definition of the joint global strategy of those companies; |
| b) coordinate their action, in order to guarantee compliance with the duties which at each moment are assigned to them; |
| c) **to** ensure joint representation of the interests common to all of them* |
| d) **to** ensure, globally, the functions common to all of them, namely in the financial area, with a view to obtaining group synergies. |

Table 6. Error Type evaluation of Fluency (syntax).

In both groups and in the two sentences, PE operations were limited to minor changes, mainly, in conjunctions, prepositions and articles. In G1, some unnecessary alterations were made, and one error was introduced in n.1. Conversely, in n.2, alterations were more relevant, as both groups detected the awkward phrasing. Again, G1 participants made more alterations than G2, but were not more successful in improving fluency.

Fluency is also achieved through more mechanical operations, such as checking punctuation, spelling or enumerative structures[9]. A minority of participants in G1 and G2 fail to fix the two punctuation signs (see table 6*) missing in n. 2, Article 3. As for inconsistent use of preposition **to** in c) and d) of n.2, the majority in both groups does not restore the syntactic parallelism at the start of each item of the enumerative structure. Differences in each group PE behaviour are not noticeable.

In the design category, the aim was to verify editing behaviour in operations that did not require cognitive effort, nor any special domain or revision competences.

| Source text | PE text G1 | PE text G2 |
| --- | --- | --- |
| Title formatting | Error fixed:**13** | Error fixed:**3** |
| Title bold | Error fixed:**2** | Error fixed:**1** |
| Capital letters | Error fixed:**2** | Error fixed:**0** |

[8] This explanation is corroborated by answers to question 9 (post-questionnaire), in which more than half the participants refer terminology as a major preoccupation when post-editing.
[9] According to Ho-Dac et al. (2012), enumerative structures are characterised by an internal organisation and involve several sub-segments: a trigger (optional), segments composing the enumeration, a closure (optional).

| Typo error | Error fixed:**2** | Error fixed:**0** |
|---|---|---|
| Typo repetition | Error fixed:**9** | Error fixed:**3** |
| Address convention | Error fixed:**2** | Error fixed:**0** |

Table 7. Error Type evaluation of Design

Interestingly, both groups do not carry out what can be considered simple revision operations. This behaviour is more noticeable in G2, who scored slightly better in the category of accuracy. The fact that both groups fail to carry out what can be considered simpler, more mechanical revision operations may indicate that explicit instructions on how to carry out MTPE of legal documents are needed.

## 4. Results and Discussion

The present study aimed at investigating whether prior introduction to MTPE reflected in the edited products of 2 groups of participants. Analysis of the edited extracts does not reveal significant differences in each group's results: the single noticeable evidence is that the majority of participants in both groups fail to identify many of the errors in the 3 categories analysed.

In pre and post-questionnaires, participants reveal a positive attitude towards MT recognizing time saving and cost-effective benefits of technology, stating its limits in translating of culturally-marked texts and emphasizing the need for human intervention. Most of them claim to be familiar with and use MT systems – especially DeepL and Google Translate –, they rate the quality of the MT output as generally good and around three-quarters are quite satisfied with their edited texts. This positive attitude seems to be in line with the way both groups approach the editing task: confidence in the MT output quality and affirmed usefulness of the MT output, especially for the fluency category, may help explain the few edits participants make. Accordingly, the only granular category that participants state requires relevant edits is terminology. Operations for the high-level Accuracy category are the most numerous, with G2 performing slightly better than G1 in fixing these errors. At the same time, G1 tend to make more edits that do not fix but introduce a new error. In the category of Fluency, both groups results are also quite similar. Correlating these results with participant profile – G2 are, on average, older and have professional experience in translation – may indicate that familiarity with the domain is responsible for G2 more efficient MTPE.

Conversely, when expected to carry out editing operations that did not require specific knowledge of domain or discourse, both groups failed to identify grammar, punctuation, spelling and formatting errors. In the post-questionnaire questions 7 and 8, the majority of participants perceived the MT output good linguistic quality as quite useful for the PE process, stating that it made the editing task easier and faster, an attitude that may have played a role in the seemingly careless way with which tasks requiring less cognitive effort were approached.

Answers to open question 9 of the post-questionnaire, in which participants described the process followed to PE the extract, further corroborate these results: over 80% of G1 mention verifying every term by resorting to dictionaries, glossaries and databases. About half mention using other tools – such as Grammarly – to revise language, while others refer that not translating from scratch allows them to focus on lexical errors and saves time because there is no need to formulate sentences. Most G2 participants mention the same preoccupation with accurate rendering of terminology. Sources used include EDP's site, glossaries and parallel texts. Only one participant refers the need to correct literal syntax.

As far as it is possible to answer the questions in 3., no noticeable differences in both groups' lexical, grammatical or formatting edits (Q. 2) were detected and G2 slightly better performance in fixing errors cannot be traced to prior PE training, due to participant age group and professional experience. Younger and less experienced G1 participants make a higher number of edits and mention using other tools to PE. Other than this, there are no significant

differences in both groups' PE process description (Q. 3): more than half the participants explicitly say they compare source and MT texts, resort to external sources to check terminology, check language and do not mention design or style.

## 5. Conclusions and Future Work

As stated earlier, this was a small-scale, preliminary study, for which results have to be cautiously approached. The small number of participants, the reduced size of the edited extract, the heterogeneous profile of participants, and the experiment conditions (participants carried out the task as one more class exercise) do not favour the quantification of findings. From the trainer's point of view, though, variability is the rule: trainees change every six months and successful teaching strategies depend on factors that the trainer is not always able to control.

The decision not to control every variable in the study was due to its broader aim: to prepare the introduction of MTPE in the legal translation class. Preference was, therefore, given to detailed analysis of how each group edited each error (sub-)category and the categories selected covered the domain, discursive and stylistic components that need to be accounted for when training legal translators. It was possible to determine that, in general, participants focused their PE efforts on the category of accuracy, complying with the precision required to translate legal concepts. However, they relied too much on the perceived "fluency" of the MT output. The fast-growing linguistic quality of NMT output and a preference for literal solutions in legal translation, may account for participants' attitude, together with the fact that they post-edit into their L2. What these facts do not explain is why both groups failed in identifying and fixing error categories that are not cognitively demanding.

Legal translation is demanding for trainees and thus promoting the development of the methodological competence by proposing translation workflows to guide the translation process may reflect on more informed decision-making (Prieto Ramos, 2011). Although the addition of technological resources and tools to the process may add to the cognitive load, it is inevitable and, therefore, justifies further investigation on how to accommodate the shift from the traditional translation process to a process that is initiated with multiple texts, sources and tools from which the most adequate option has to be selected.

The study's preliminary results have led to fine-tuning the initial integrative methodology, specifically in the following: providing trainees with a brief introduction to MT's limits and typical error types and a simplified framework for error analysis, both adapted to the translation of legal texts; raising awareness to the abundance of incorrect literal options typical of legal translation and of automated systems output; redesigning steps 2 and 3 of the methodological tool by introducing a pre-documentation stage to edit the design and mechanic fluency categories and flag potential syntactic and terminological errors; promote guided documentation based on textual resources that efficiently provide insight on the legal and linguistic features that need to be catered for in PE; adding any sub-categories needed.

The fine-tuned model was presented in two legal translation classes that participated in the second version of the pilot study, in June of 2023. Comparison of both studies' results is expected to shed further light on the productive use of MTPE in domain-specific translation training. Professional post-editors of today may have been trained as translators, but present-day translation trainees will inevitably take up both activities simultaneously. It is not unlikely that PE is what trainees are now doing in translation classes, even if they and their trainers do not address it openly. It becomes imperative, therefore, to integrate MTPE in the specialised translation class.

## 6. References

Albi, A. B. (2005): Es posible traducir realidades jurídicas? Restricciones y prioridades en la traducción de documentos de sucesiones británicos al español. In E. M. Nebot, A. B. Albi (Eds.), *La traducción y la interpretación en las relaciones jurídicas internacionales* (pp. 65-89). Publicacions de la Universitat Jaume I.

Alcaraz, E. & Hughes, B. (2002) *Legal Translation Explained*. St. Jerome.

Bestué, C. (2016) Translating law in the digital age. Translation problems or matters of legal interpretation?, *Perspectives, 24*:4, 576-590, DOI: 10.1080/0907676X.2015.1070884

Biel, Ł. (2008). Legal terminology in translation practice: Dictionaries, googling or discussion forums? *SKASE – Journal of Translation and Interpretation, 3*:1, 22-38. http://www.skase.sk/Volumes/JTI03/pdf.doc/3.pdf

Biel, Ł. (2009). Organization of background knowledge structures in legal language and related translation problems. *Comparative Legilinguistics, 1*, 176–189. https://doi.org/10.14746/cl.2009.01.13.

Cao, D. (2007). *Translating Law*. Multilingual Matters.

Carmo, F. E. M. (2017). *Post-editing: A Theoretical and Practical Challenge for Translation Studies and Machine Learning*. [Ph.D. thesis]. Universidade do Porto.

Directorate-General for Translation (2022). *EMT Competence Framework*. European Master's in Translation. https://commission.europa.eu/system/files/2022-11/emt_competence_fwk_2022_en.pdf.

Doherty, S. (2016). The impact of translation technologies on the process and product of translation. *International Journal of Communication, 10*, 947–969. https://ijoc.org/index.php/ijoc/article/viewFile/3499/1573

Ferreira, A., Gottardo, A. & Schwieter, J. (2018). Decision-making processes in direct and inverse translation through retrospective protocols. *Translation, Cognition & Behavior, 1*, 98-118. DOI:10.1075/tcb.00005.fer.

Fonseca, N. (2015). Directionality in translation: Investigating prototypical patterns in editing procedures. *Translation & Interpreting, 7*(1), 111-125 . DOI: ti.106201.2015.a08

Gorog, A. (2014). Quantifying and benchmarking quality: the TAUS Dynamic Quality Framework. *Revista Tradumàtica: tecnologies de la traducció, 12*, 443-454. DOI:10.5565/rev/tradumatica.66

Harvey, M. (2002). What's so Special about Legal Translation? *Meta, 47*(2), 177–185. https://doi.org/10.7202/008007ar

Ho-Dac, L., Fabre, C., Péry-Woodley, M., Rebeyrolle, J. & Tanguy, L. (2012). An Empirical Approach to the Signalling of Enumerative Structures, *Discours*, *10*. http://journals.openedition.org/discours/8611. DOI: 10.4000/discours.8611.

Hu, K. & Cadwell, P. (2016). A comparative study of post-editing guidelines. *Baltic J. Modern Computing, 4*(2), 346-353. DOI: 10.13140/RG.2.1.2253.1446.

International Organization for Standardization (2017). *Translation services ‒ Post-editing of machine translation output ‒ Requirements* (ISO Standard No. 18587:2017).

Kenny, D. & Doherty, S. (2014) Statistical machine translation in the translation curriculum: overcoming obstacles and empowering translators, *The Interpreter and Translator Trainer, 8*(2), 276-294. DOI: 10.1080/1750399X.2014.936112.

Kenny, D. (2020). Technology in Translator Training. In M. O'Hagan (Ed.), *The Routledge Handbook of Translation and Technology* (pp. 498-515). Routledge.

Kenny, D. (2022). Human and machine translation. In D. Kenny (Ed.), *Machine translation for everyone: Empowering users in the age of artificial intelligence* (pp. 23-49). Language Science Press. DOI: 10.5281.

Kiraly, D. (2000). Translation into a non-mother tongue: From collaboration to competence. In M. Grosman, M. Kadric, I. Kovacic & M. Snell-Hornby (Eds.), *Translation into Non-mother Tongues* (pp.117-123). Stauffenburg Verlag.

Koponen, M. (2016). Is machine translation post-editing worth the effort? A survey of research into post-editing and effort. *The Journal of Specialised Translation, 25*, 131–148. http://www.jostrans.org/issue25/art_koponen.pdf.

Nord, C. (2016) Skopos and (Un)certainty: How Functional Translators Deal with Doubt. *Meta, 61*(1), 29-41. https://doi.org/10.7202/1036981ar.

O'Brien, S. (2022). How to deal with errors in machine translation: Postediting. In D. Kenny (Ed.), *Machine translation for everyone: Empowering users in the age of artificial intelligence* (pp. 105-120). Language Science Press. DOI: 10.5281.

Pavlović, N. (2007). Directionality in Translation and Interpreting Practice: Report on a questionnaire survey in Croatia. *FORUM Revue internationale d'interprétation et de traduction / International Journal of Interpretation and Translation, 5*(2), 79-99. DOI: 10.1075/forum.5.2.05pav.

Pavlovic, T. (2013). Exploring Directionality in Translation Studies. *ExELL (Explorations in English Language and Linguistics), 1*(2), 149-165. http://www.exell.ff.untz.ba/wp-content/uploads/2014/04/ExELL.1.149.Pavlovic.pdf.

Prieto Ramos, F. (2011). Developing Legal Translation Competence: An Integrative Process-Oriented Approach. *Comparative Legilinguistics, 5*, 7-22. https://doi.org/10.14746/cl.2011.5.01.

Prieto Ramos, F. (2014). Parameters for Problem-Solving in Legal Translation: Implications for Legal Lexicography and Institutional Terminology Management. In *The Ashgate Handbook of Legal Translation* (pp. 121-134). Routledge.

Pym, A. (2013). Translation Skill-Sets in a Machine-Translation Age. *Meta*, 58(3), 487–503. https://doi.org/10.7202/1025047ar

Pym, A., & Torres-Simon, E. (2021). Efectos de la automatización en las competencias básicas del traductor: la traducción automática neuronal. In A. V. Suñe, & A. A. Alarcón (Dirs.), *Ocupaciones y Lenguaje: Indicadores y Análisis De Competencias lingüísticas En El Ámbito Laboral.* Universitat Rovira i Virgili.

Rodríguez de Céspedes, B. (2019). Translator education at a crossroads: the impact of automation. *Lebende Sprachen, 64*(1). https://doi.org/10.1515/les-2019-0005.

Šarčević, S. (2000), Legal Translation and Translation Theory: a Receiver- oriented Approach, In *Legal translation: history, theory/ies and practice.* [Proceedings of the International colloquium University of Geneva], February 17-19, 2000 - Berne: ASTTI, cop. 2000. http://tradulex.org/Actes2000/sarcevic.pdf

Scott, J. (2012). Can Genre-Specific DIY Corpora, Compiled by Legal Translators Themselves, Assist Them in 'Learning the Lingo' of Legal Subgenres?. *Comparative Legilinguistics, 12*, 87–100. https://doi.org/10.14746/cl.2012.12.05

TAUS. (2016). *Post Editing Guidelines*. Isabella Massardo, Jaap van der Meer, Sharon O'Brien, Fred Hollowood, Nora Aranberri, Katrin Drescher. https://www.taus.net/resources/reports/mt-post-editing-guidelines

TAUS, (2017). *Quality Evaluation using an Error Typology Approach*. TAUS BV, De Rijp. https://cdn2.hubspot.net/hubfs/2734675/Reports,%20ebooks/Error%20Typology%20Best%20Practices%20Guidelines.pdf.

Valli, P. (2015). The TAUS Quality Dashboard (Review of The TAUS Quality Dashboard). *Proceedings of the 37th Conference Translating and the Computer* (pp. 127–136). AsLing. https://aclanthology.org/2015.tc-1.17.pdf.

Vieira, L.N. (2019). Post-Editing of Machine Translation. In M. O'Hagan (Ed.), *The Routledge Handbook of Translation and Technology* (pp. 319-335). Routledge.

Wilss, W. (1996). *Knowledge and skills in translator behaviour*. John Benjamins. DOI: https://doi.org/10.1075/btl.15.

Yamada, M. (2019). The impact of Google Neural Machine Translation on Post-editing by student translators. *The Journal of Specialised Translation*, 31. 87-106. https://jostrans.org/issue31/art_yamada.php.