

Overview of the Major Findings of the Szeged Post-Editing Research Project

Márta Lesznyák¹, Mária Bakti², Eszter Sermann³

¹. Szegedi Tudományegyetem, Bölcsész- és Társadalomtudományi Kar, Angoltanár-képző és Alkalmazott Nyelvészeti Tanszék, ². Szegedi Tudományegyetem, JGYPK, Modern Nyelvek és Kultúrák Tanszék, ³. Szegedi Tudományegyetem, Bölcsész- és Társadalomtudományi Kar, Olasz Nyelvi és Irodalmi Tanszék

Abstract: This paper reports on a multi-year research project carried out by the Translation and MT Post-editing Competence Research Centre of the University of Szeged, Hungary. The Centre conducted a complex project to investigate the role of different competence elements in traditional from-scratch human translation (FSHT) and in the post-editing of neural machine translation (MTPE). The project began in 2020 and relied primarily on the PACTE model of translation competence. Between 2020 and 2024, data were collected from 106 students, who translated or post-edited a research paper abstract or a legal text from English into Hungarian, completed language tests and background knowledge tests, filled in a form on their beliefs about translation and a follow-up questionnaire. Drawing on models of translation and post-editing competence, we analysed how competence elements, source-text characteristics and task type (FSHT vs. MTPE) interact. In this paper, we attempt to synthesise the main findings of the project, with a particular focus on comparing FSHT and MTPE with respect to the effect of source text (ST) genre and domain, and on feedback from students working with MTPE. Overall, our findings suggest that post-editing offers clear advantages only when the genre, the content and the register of the ST are familiar to the translators / post-editors. An implication of these findings is that the development of language competence, particularly reading competence, should receive more attention in translator training and should by no means be taken for granted.

Keywords: translation competence, post-editing competence, translator training, role of the source text, language competence

1. Introduction

The aim of this paper is to summarise the key findings of a multi-year research project carried out by the Translation and MT Post-editing Competence Research Centre of the University of Szeged, Hungary. First, we introduce four competence models;

Márta Lesznyák: mlesz@edpsy.u-szeged.hu • <https://orcid.org/0000-0002-9354-0633>

Mária Bakti: bakti.maria@szte.hu • <https://orcid.org/0000-0002-4304-6805>

Eszter Sermann: sermann.eszter@szte.hu • <https://orcid.org/0000-0001-8753-5312>

next, we describe the methodology used in our investigations; then, we present our results, which are followed by a discussion and a summary.

The aim of our investigations is to map the role different competence elements play in the post-editing of machine translation. The starting point of our project is PACTE's model of translation competence (PACTE 2003), according to which translation competence is made up of five inter-related sub-competences, namely bilingual sub-competence, extralinguistic sub-competence, instrumental sub-competence, knowledge about translation sub-competence and strategic sub-competence. In addition, their model includes psycho-physiological components (PACTE 2003). PACTE's construct of translation competence has four basic tenets:

- (1) translation competence is expert knowledge and is not possessed by all bilinguals
- (2) translation competence is mainly procedural knowledge
- (3) translation competence is made up of sub-competences
- (4) one element of translation competence is strategic competence, which plays an outstanding role.

The first model of post-editing (PE) competence, proposed by Nitzke et al. (2019), is based on PACTE's translation competence model (2003) and the revision competence model of Robert et al. (2017). The model is founded on the premise that post-editors must be able to make informed decisions about the risks associated with machine translation (MT), which in turn requires a distinct set of competences. The authors identify several risk-related factors that post-editors must consider, including data breaches, loss of control over the translation process, increased liability for clients, the shortage of qualified translators and post-editors, and the quality of post-edited texts. They argue that the post-editing competence model is basically a model of making informed decisions related to risk management. The criteria for making decisions include text type, MT output quality, turnaround time and the life span of translations, data security and controlled language.

The post-editing competence model of Nitzke et al. (2019) comprises four core competences and eight subsidiary competences. The core competences are risk assessment (to assess the risk of the SL text), strategic competence (based on the risk assessment, the post-editor decides whether to implement light or full post-editing, or to leave the MT output as it is), and consulting competence (which involves informing the customer or project manager about the potential risks and the problem-solving strategies). The fourth core competence is service competence, which includes a variety of skills, namely the ability to calculate prices, the ability to handle CAT and revision tools, and MT systems. In addition, the post-editor should be familiar with the translation market, should be able to negotiate with customers and prepare an appropriate offer that includes a realistic time frame for the job.

The post-editing competence model's subsidiary competences are partly based on PACTE's (2003) translation competence model and the revision competence model proposed by Robert et al. (2017). The subsidiary competences include bilingual competence, extralinguistic competence, instrumental competence, research competence, revision competence, translation competence, MT competence and PE competence. By PE competence, they mean identifying errors triggered by neural MT systems. In addition, the model enumerates surrounding factors the post-editing task is influenced by, namely the psycho-physiological factors, the post-editor's self-perception, the post-editing brief, and the post-editor's affinity for technology.

A revised model of post-editing competence was published two years later (Nitzke and Hansen-Schirra 2021). It depicts these competences as a "house of post-editing competences". The foundation of the house (which resembles a Greek temple) is translation competence, involving bilingual, extralinguistic and research competences. The authors of the model stress that post-editors should be skilled translators, as both post-editors and translators "need the same basic skill set" (Nitzke and Hansen-Schirra 2021: 70).

The three pillars of the house of post-editing competences include the additional competences that post-editors need. The first one is error-handling, including error spotting, error classification and error correction. The second pillar is MT engineering, involving training and assessment of MT systems, in other words, what kind of pitfalls different MT systems have. The third pillar of the house is consulting competences, which are tailored to the needs of the post-editing job. The roof of the house is made up of post-editing soft skills, namely risk assessment competence and service competence. The surrounding factors influencing the post-editor's task are quite similar to those listed in the 2019 model: psycho-physiological components, affinity for technological developments, the post-editing brief and the post-editor's self-perception.

It should be noted that the two post-editing competence models are similar to a great extent in that they include the same competence elements, at the core of which are PACTE's competence elements, while the main difference lies in their visual presentation. The core competences of the 2019 model are the pillars and the roof of the house of post-editing competences, whereas the subsidiary competences of the 2019 model are the foundation competences of the 2021 model.

The EMT Competence Framework (2022) describes five main areas of competence: Language and Culture, Translation, Technology, Personal and Interpersonal, and Service Provision. In this framework, Translation competence encompasses strategic, methodological and thematic competences, and as part of this, students know how to "post-edit MT output using style guides and terminology glossaries to maintain quality standards in MT-enhanced translation projects" (2022: 8). The fact that post-editing has become part of translation competence can be seen as a major development in the field.

2. Methodology

The aim of the Translation and MT Post-editing Competence Research Centre's investigations was to map the roles the following competence elements and variables play in the quality of the translation or post-editing products: English language competence, text type and domain (licence to publish statement vs a research paper abstract), domain knowledge (legal and applied linguistics), knowledge or beliefs about translation, expertise, experience and task (from-scratch human translation or post-editing of machine translation).

Our investigations were carried out among MA students of translation attending the Translation and Interpreting MA programme of the University of Szeged, Hungary. This is a four-semester training programme, in which students choose a specialisation (translation or interpreting) after completing the first two semesters. In addition to their A language (Hungarian), students choose a B and a C language, one of which has to be English.

Data collection took place between May 2020 and May 2024, in eight data collection waves. In each academic year, MA students of translation were asked to translate or post-edit a legal or a humanities source language text from English into Hungarian. Students were randomly assigned to one of four groups: "translation: humanities text", "translation: legal text", "post-editing: humanities text" and "post-editing: legal text". Each student worked on one text only. At the September data collection events, participants included first-year MA students of translation, who were literally at the beginning of their studies. In contrast, in the May data collection round, second-year students participated, a few weeks before their final exam.

With the exception of the 2020 and 2021 data collection rounds, all data were collected in a classroom setting. Students were allowed to use their own laptops, although university laptops were also provided for those who did not have a device. Students were also permitted to use any internet sources, but students in the translation groups were instructed to refrain from using any MT engine. The time limit was 120 minutes, but none of the students actually required that much time.

Altogether, 106 students produced 170 Hungarian target language texts, 64 of which were produced by the same students at two different points in time. All the participants used pseudonyms. It should also be noted that at the time of the data collection, students had received no explicit training in post-editing. Participation was voluntary for students, although they received extra points added to their seminar scores for contributing to the research.

3. Material, methods and instruments

To test the translation performance of the participants, two English source language texts were used: a research paper abstract on bilingualism and a copyright agreement. The post-editing performance of the participants was tested using the machine-translated Hungarian target texts of these two SL texts. The first SL text was a 127-word abstract on bilingualism. We assumed that the domain and the register of this text were more familiar to the students than that of the second SL text, a 350-word copyright agreement. This text had longer, more complex sentences, and the domain and register were unfamiliar to the students. Our assumption on the difficulty of the texts was tested and verified with a survey conducted with translation trainers (see 4.2 below). The MT output of the academic abstract was produced by the free version of DeepL, a neural machine translation system, while eTranslation, the European Commission's secure neural machine translation service, was used to generate the MT version of the copyright agreement. The domain of eTranslation was set to General Language. Both outputs were generated in 2020.

Students who agreed to participate in the study were asked to complete several questionnaires and tests before carrying out the translation task. These included a demographic background questionnaire, English language tests, a thematic knowledge test, and a questionnaire on beliefs about translation. All tests were administered online, in Google Forms. The instruments are described below.

To measure the source language (English) proficiency of the students, selected tasks from Cambridge Proficiency Tests were used. The Reading part of the test students had to complete consisted of three parts: two multiple choice tests and one task with missing paragraphs. This part of the test was worth 12 points. The second part, *Use of English*, also comprised three sections: two gap-filling tasks and one sentence transformation task. The *Use of English* section was worth 29 points.

The students' background (domain-specific) knowledge was assessed using two tests, each consisting of 10 multiple-choice questions. With the assistance of subject-matter experts, one test was developed to assess knowledge of copyright law and the other to assess knowledge of bilingualism and pragmatics.

To test the students' beliefs about translation, a specific questionnaire was compiled by our research group. The construction and validation of the questionnaire were described in detail in Lesznyák, Sermann and Bakti (2023, 2025), so only a brief summary of the process is provided here. As the final version of the questionnaire was completed by the summer of 2021, only data collected after that date could be included in the correlation analysis.

The questionnaire was based on PACTE's questionnaire on declarative knowledge about translation (PACTE 2008, 2011, 2014). Knowledge about translation is one of the five sub-competences of the PACTE model, and is described as "predominantly declarative knowledge, both implicit and explicit, about translation and

aspects of the profession” (PACTE 2008: 106). PACTE based their questionnaire about knowledge of translation on two paradigms of thinking about translation (dynamic vs static), with statements concerning the following eight factors: the concept of translation and translation competence, translation units, translation problems, the phases in the translation process, the methods required, the procedures used in translation, the role of the translation brief, and the role of the target language reader.

In our investigations, we compared the results of MA students in translation and interpreting and those of BA and MA students majoring in a foreign language and students majoring in foreign language teaching. However, we encountered problems, namely that of low reliability, and it also became clear that the questionnaire did not differentiate appropriately between the groups tested. We decided to introduce radical changes to the questionnaire. The new questionnaire focused on participants’ beliefs about translation, rather than knowledge of translation. Beliefs about translation are, in our view, “ideas, notions, conceptions and opinions that guide the process of translation on the one hand and translation competence acquisition on the other hand” (Lesznyák, Sermann and Bakti 2025: 80). Following piloting and restructuring, 18 items were included in the final version of the questionnaire, focusing on the following five factors: concepts of translation and translation competence, translation problems, methods and translation procedures, the translation context or setting, and post-editing.

Students were also invited to fill in a follow-up questionnaire after completing the translation or post-editing task. The first set of questions consisted of closed-ended items on the perceived difficulty of the task (human translation or post-editing), addressing both its overall difficulty and specific aspects. The open-ended questions invited students to identify terminological problems and to list the terms they had found most difficult to translate. These were followed by closed-ended questions on experiences related to post-editing. These focused on the perceived difficulty and speed of MTPE as compared to from-scratch human translation, together with the rating of factors that helped or hindered the MTPE process. The English language tests, the background knowledge test, the ‘Beliefs about Translation’ test, and the follow-up questionnaire were administered online as Google Forms, which students completed at home.

The quality of the translated or post-edited Hungarian texts was evaluated using the following MQM (ref.) error categories: accuracy, fluency, style, terminology¹. In addition, total error numbers were calculated. Two or three evaluators, who are themselves translator trainers, assessed and annotated each Hungarian target language text, and the mean of their scores was used for analysis.

In the analysis, together with the evaluation of the target texts, the scores from the English language tests, the domain-specific knowledge tests and the ‘Beliefs about Translation’ questionnaires were used. All statistical analyses were carried out

¹ <https://themqm.org/the-mqm-typology>

in SPSS v. 24. It should be noted that the results of the 2024 English test scores were excluded from the analysis, as surprisingly good results suggested potential use of AI-based tools or other forms of academic misconduct.

In spite of the relatively large number of students participating in the study, some data loss and corruption occurred because a few students did not complete all the online questionnaires or, surprisingly, could not handle digital technology appropriately. As a result, only a limited number of texts could be included in the analysis presented in this paper: bilingualism abstract = 13 FSHT and 21 MTPE; copyright agreement = 32 FSHT and 37 MTPE.

4. Results and discussion

This section will provide a summary of the main findings of the research project, with particular focus on performance indicators (i.e. target text quality and time on task), the role of competence elements assessed in the study and students' perception of post-editing. Some results have already been published in previous papers; these will be outlined and referenced here, but full statistics will not always be provided.

4.1. Performance indicators: target text quality

When students had to translate and post-edit the bilingualism abstract, post-editors' overall performance was slightly but statistically significantly better than that of FSHT students. The independent samples *t*-test showed that they committed fewer errors (FSHT $M = 5.67$, $SD = 0.98$, MTPE $M = 4.69$, $SD = 0.84$, $t(32) = 3.1$, $p = .004$). For the legal text, however, no significant differences were found between the two groups (FSHT $M = 7.00$, $SD = 1.57$, MTPE $M = 6.48$, $SD = 1.48$, $t(66) = 1.4$, $p = .167$). Findings from earlier stages of the study showing similar tendencies have been published previously (Lesznyák, Bakti and Sermann 2023; 2024). Overall, these results suggest that post-editing may enhance students' performance, but it is not always the case. The source text seems to be a factor confounding MTPE's effectiveness; therefore, we decided to examine whether the difficulty of the source text affected students' performance.

4.2. The effect of the source text on students' performance

It was assumed that both the genre and the content of the legal text were unfamiliar to the students, which might have led to poorer performance compared to the abstract on bilingualism. To confirm that the copyright agreement was more difficult to translate than the bilingualism abstract, 12 translator trainers were asked to judge the difficulty

of the two source texts on a five-point scale (1 = very easy, 5 = very difficult). The responses indicated that the trainers indeed found the copyright agreement more difficult ($M = 4.16$) than the bilingualism abstract ($M = 2.75$).

In the next step, student error numbers per sentence were compared in both the FSHT and the MTPE group. Findings in both groups, as illustrated in tables 1 and 2, suggested that, generally, students made significantly more errors in the copyright text than in the abstract on bilingualism. On the one hand, this does not come as a surprise, but on the other hand, it shows that MT cannot be used to compensate for deficiencies in legal language or legal background knowledge.

Table 1

Differences in error numbers between the bilingualism abstract and the copyright agreement in the FSHT group. Results of the independent samples t-tests

	Bilingualism ($n = 13$)		Copyright ($n = 32$)		<i>t</i> -test	<i>p</i>
	<i>M</i>	<i>SD</i>	<i>M</i>	<i>SD</i>		
<i>Accuracy errors*</i>	2.09	0.75	2.66	0.70	-2.41	.020
<i>Fluency errors</i>	1.19	0.75	1.22	0.59	-.125	.901
<i>Style errors*</i>	1.51	0.48	1.13	0.44	2.52	.016
<i>Terminology errors*</i>	0.92	0.28	1.98	0.60	-8.04	.000
<i>Total number of errors*</i>	5.67	0.98	7.00	1.57	-2.80	.008

Table 2

Differences in error numbers between the bilingualism abstract and the copyright agreement in the MTPE group. Results of the independent samples t-tests

	Bilingualism ($n = 21$)		Copyright ($n = 37$)		<i>t</i> -test	<i>p</i>
	<i>M</i>	<i>SD</i>	<i>M</i>	<i>SD</i>		
<i>Accuracy errors*</i>	1.82	0.61	2.45	0.73	-3.33	.002
<i>Fluency errors*</i>	0.65	0.39	1.12	0.53	-3.89	.000
<i>Style errors</i>	1.22	0.48	1.01	0.43	1.64	.106
<i>Terminology errors*</i>	1.02	0.20	1.89	0.44	-10.09	.000
<i>Total number of errors*</i>	4.69	0.85	6.48	1.48	-5.06	.009

Summarising the findings on target text quality, we can argue that several background factors influence whether working with post-editing leads to more success than working with traditional from-scratch human translation. One such factor is the source text. Post-editors produced higher quality target texts when they were familiar with the genre and the topic of the ST, and when it was easier to translate (as rated by experts). Nevertheless, they seemed to lose their advantage over from-scratch transla-

tors when the text belonged to a domain unfamiliar to them. Poor quality MT output, as will be detailed below, may have exacerbated the problem. What seems certain, however, is that MT is only as effective as the person using it and cannot compensate for deficiencies in post-editing or translation competence.

4.3. Performance indicators: time on task

A series of independent samples *t*-tests showed that there was a tendency for MTPE students to work somewhat faster than FSHT students on both the bilingualism abstract and the legal text, particularly at the end of the training. Nevertheless, the difference between the two groups never reached statistical significance (Lesznyák, Bakti and Sermann 2023; 2024). The lack of statistical significance may be attributed to small sample size, but another plausible explanation is that students lacked appropriate training in post-editing, which led to an inefficient use of time.

The findings seem to somewhat contradict Nitzke and Oster's (2016) results, which indicated that domain-specific texts require more effort in FSHT and less effort in MTPE than non-domain-specific texts. They also found post-editing to be faster for both domain-specific and general texts. Our results, however, suggest that temporal effort was similar in the FSHT and the MTPE group. There are several explanations for this. First, without going into much detail, the two studies had very different designs, in as much as different types of instruments were used and the sample size and characteristics also differed. But most importantly, Nitzke and Oster (2016) most likely used SMT in 2012, whereas we used NMT in our study. In addition, in the present study, both texts were domain-specific, although one was more familiar to the students. All these minor factors may have contributed to the different findings.

Furthermore, it must be emphasised that large individual differences were found in both the FSHT and MTPE groups (Bilingualism abstract: FSHTmin = 28.67 min, FSHTmax = 55.62 min, MTPEmin = 25.08 min, MTPEmax = 59.90 min; Copyright text: FSHTmin = 54.66 min, FSHTmax = 83.52 min, MTPEmin = 31.40 min, MTPEmax = 73.22 min). This suggests that temporal efficiency may also be related to personality (see Rojo López et al. 2021) and not only to work mode.

One of our most surprising findings is that correlations between time on task and the quality of the target texts were mostly not present or very weak (Lesznyák, Bakti and Sermann 2023; 2024). A possible explanation for this is that speed could be a highly individual variable that is more related to the translator's/post-editor's personality (e.g. conscientiousness) than to the underlying translation or post-editing competence (Rojo López et al. 2021). In addition, lack of training in post-editing may have contributed to low correlations between time on task and performance, as students could have spent a considerable amount of time on unsuccessful strategies.

4.4. The role of competence elements – correlation patterns behind performance

The role of competence elements in performance will be discussed separately for the abstract on bilingualism and the legal text, as the role of these elements seems to be partly (con)text-specific.

4.4.1. *Translating and post-editing the bilingualism abstract*

In the FSHT group, no significant correlations were found between language competence (reading and grammar) test scores and error numbers. A moderate negative correlation emerged between thematic knowledge and total error numbers, although this only trended towards statistical significance ($r = -.50$, $p = .063$). Scores on beliefs about translation showed strong, significant negative correlation with accuracy error counts ($r = -.85$, $p = .032$). Negative correlation means that higher test scores were associated with fewer errors, suggesting that greater knowledge of bilingualism and a professional mindset concerning translation are related to higher quality translations.

In the MTPE group, grammar test scores showed a significant moderate negative correlation with total number of errors ($r = -0.58$, $p = .023$), but reading scores were again not related to post-editing performance. Significant moderate correlations were observed between thematic knowledge and several aspects of performance: that is, fluency ($r = -.47$, $p = .036$), style ($r = -.46$, $p = .041$) and total error count ($r = -.46$, $p = .041$). Finally, strong correlations were found between ‘Beliefs about Translation’ test scores and total error count ($r = -.76$, $p = .017$). Moreover, there were moderate correlations between ‘Beliefs about Translation’ test results and accuracy ($r = -.61$, $p = .079$) and style ($r = -.64$, $p = .066$) respectively, but values were only approaching statistical significance.

Overall, these results indicate that beliefs about translation and thematic knowledge may play a more important role than source language competence in performance when the content and the genre of the text are more familiar to the students. This holds true for both from-scratch human translation and post-editing.

4.4.2. *Translating and post-editing the copyright agreement*

For the legal text, somewhat different correlation patterns were found to those for bilingualism abstract. Reading scores showed moderate to strong significant correlations with some performance variables, including accuracy ($r = -.57$, $p = .004$), fluency ($r = -.64$, $p = .001$), and total error numbers ($r = -.48$, $p = .018$). The statistical analysis revealed similar associations between grammar test scores and performance

as measured by accuracy ($r = -.57, p = .003$), fluency ($r = -0.45, p = .027$) and total error count ($r = -0.42, p = .044$).

‘Beliefs about Translation’ test scores were significantly correlated with multiple performance measures, notably style ($r = -.61, p = .009$), terminology ($r = -.62, p = .008$) and total error number ($r = -.54, p = .024$). The correlations were moderate and negative, indicating that more professional views on translation are associated with better performance on the variables mentioned above.

In the MTPE group, significant moderate relationships were observed between reading competence and three performance variables: accuracy ($r = -.57, p = .001$), terminology ($r = -.43, p = .016$) and total error numbers ($r = -.54, p = .002$). Grammar test scores were significantly associated with the same indices of performance: accuracy ($r = -0.43, p = .019$), terminology ($r = -0.36, p = .049$) and total error count ($r = -.40, p = .027$). Additionally, there was a marginally significant moderate correlation between ‘Beliefs about Translation’ test scores and terminology error count ($r = -.44, p = .053$). All correlations described above are negative, indicating that higher test scores are accompanied by better performance. The most surprising finding of the research was that neither the FSHT nor the MTPE group’s performance was associated with thematic (background) knowledge for the legal text.

In summary, results for the legal text show that source language competence, particularly reading, had primary importance in determining performance (see also Lesznyák, Bakti and Sermann 2024), followed by beliefs about translation. Interestingly, legal background knowledge test scores were not related to performance. It should be noted here that the participants were humanities students with very little legal background knowledge. Their level of thematic knowledge may not have reached a level where it could positively contribute to performance. In such a setting, language competence, particularly reading, is used to compensate for deficiencies in background knowledge (Lesznyák, Bakti és Sermann 2024).

4.5. Students’ perceptions of post-editing

The follow-up questionnaire generated a wealth of data, part of which is still being analysed, while a separate study based on the remaining data is planned for publication elsewhere. Consequently, only a summary of the most essential findings will be presented here.

The follow-up questionnaire included several statements focusing on the specific features of the post-editing process. An open-ended question invited students to share any thoughts they had about the task. A recurring topic for the legal text was the quality of the MT output. Students repeatedly expressed their dissatisfaction with the translation produced by eTranslation, and some of them suggested that they basically had to retranslate the text.

Post-editor ‘Kilenc9’:

(1) *Több időt vett igénybe az utószerkesztés, mintha fordítottam volna a szöveget.*

(1b) *It took longer to post-edit the text than it would have taken to translate it.*

It should be noted here that the MT output was generated in 2020 and because of the longitudinal aspect of the project, it could not be changed in later years. Obviously, the quality of MT- and AI-generated translations into Hungarian has improved considerably since 2020. Nevertheless, it is not uncommon to meet machine translated texts, particularly in the legal domain, which are of hardly any use.

In relation to the legal text, some further interesting findings emerged. Both FSHT and MTPE students found their task of translating and post-editing generally difficult. The mean value of perceived difficulty was 3.74 ($SD = 0.62$) in the FSHT group and 4.08 ($SD = 0.64$) in the MTPE group. However, the independent samples t -test suggested that the difference between the two groups was only approaching significance ($t(46) = -1.87, p = .068$), which means that the difference between the two groups could be a measurement error. Nevertheless, post-editing in this setting was by no means felt to be easier than traditional from-scratch translation.

In summary, students' perception of the difficulty of the task, coupled with the observed low quality of the MT output, indicates that post-editing may not always be easier than from-scratch translation, and the simple fact that there is MT output will not make the job of inexperienced translators easier. As performance data above has shown, NMT or AI will not even automatically help students produce better target texts. The domain of the source text and the quality of the MT output play a key role in determining how useful MT can be in the translation process and whether it actually facilitates the work of inexperienced translators.

Finally, in the MTPE group working on the copyright agreement, a few significant correlations were observed. Notably, overall performance was significantly associated with both perceived overall difficulty ($r = -.40, p \leq 0.05$) and perceived syntactic difficulty ($r = -.53, p < 0.01$). Correlations were negative, indicating that the higher the perceived difficulty, the lower the number of errors in the final target text. This pattern may provide some tentative evidence for the assumption that realistic task assessment is particularly important in MT- and, probably, in AI-driven translation. Without such awareness, student or novice translators may easily fall into the trap of assuming that the job is easy because MT or AI has produced a solid text to rely on. Our findings suggest that those students who were more aware that the post-editing task was difficult in spite of having MT output at hand actually performed better. An implication of this finding is that students should be cautioned against taking the MT output for granted and assuming that the post-editing task is easy.

Some statements in the questionnaire asked students to assess whether certain features of MTPE made their work easier or harder. Students' responses showed a great variety and appeared to be strongly affected by text type (legal or academic)

and previous experiences (first- or second-year students). There were only two variables where responses seemed to be consistent across the different sub-samples. First, contrary to what has frequently been reported in the literature (Cadwell et al. 2018; Läubli and Orrego-Carmona 2017; Sakamoto 2019; Alvarez-Vidal et al. 2021), students did not indicate any issues with the emotional aspect of post-editing. In other words, they did not suggest that it was more menial or less creative than traditional from-scratch translation (Lesznyák 2026). Second, students consistently reported that the greatest advantage of post-editing was that it saves typing time (Lesznyák 2026). Responses concerning other potential advantages like having a solid starting point and having terminology work already done were not so clear-cut.

These results support the idea that the perceived benefits and limitations of post-editing are highly dependent on the individual characteristics of the translation setting itself, including the features of the source text (content and genre), the quality of the MT output and the translator's level of expertise. Nevertheless, our research provided evidence that the younger generation may not feel any inherent lack of creativity in post-editing. Somewhat ironically, however, the advantage that students most consistently emphasised was simply that post-editing saves typing time.

5. Conclusions

The comprehensive investigation outlined in this paper has generated a large volume of data, only a part of which could be presented and analysed here. In this final section, we would like to highlight the most important findings and a number of implications concerning translator training.

One of the objectives of the project was to identify how work mode, that is FSHT or MTPE, affects performance measures. Our findings indicate that MTPE may boost performance if the content, the genre and the register are familiar to the translator. However, the advantages of MTPE seem to disappear when the text is too demanding in terms of content or form. This was observed with the legal text in our study. It must be emphasised here that MT could not compensate for deficiencies in translation competence, most probably due to lack of legal background knowledge and deficiencies in legal language skills, both in Hungarian and English. The fact that MT cannot compensate for deficiencies in competence may sound self-evident, but it is well documented in the literature that non-professionals and novices can be characterised by overtrust and overreliance on MT and AI (e.g. Belhassen and Hamda 2025; Quinci 2023). In other words, they may actually hope that MT will be able to solve problems they cannot solve. Therefore, it is vital that students are made aware of the fact that MT is only as good as they are. In addition, MT- and AI-supported translation should not be introduced too early in the training, as this may lead to a fossilisation of post-editing strategies that tend to overrely on MT/AI. Fossilisation

is a term borrowed from foreign language learning and refers to the “cessation of [...] development despite input, instruction or exposure” (Selinker 1972). Zhang and Doherty (2025) suggest that particularly translators’ creative and critical thinking skills may be hindered or stalled in development by overreliance. Overreliance on AI and its effects may be a promising new research field in itself.

If we now turn to competence elements and their role in performance, we find very different correlation patterns depending on what type of text the translator is working on. Language competence appears to be a decisive factor when the text is highly demanding, as was the case with the legal text in our study. The leading role of language competence can be observed in both FSHT and MTPE. In contrast, language competence looks relatively insignificant when the genre and the content of the text are reasonably familiar to the translator. Translating or post-editing the bilingualism abstract represented such a scenario in our study. In these cases, thematic background knowledge shows stronger associations with performance. Beliefs about translation, that is, a professional view of and attitude towards translation, are always important, but their relative weight may vary with context. An implication of these findings is that the development of language competence, particularly reading competence, should receive more attention in translator training and by no means should be taken for granted. In addition, students should be made aware of which elements of their competence they can rely on when dealing with different text types.

A remarkable result from the questionnaire on students’ perceptions of post-editing was the significant association between more realistic task assessment and better performance for the legal text. This suggests that students should be made aware that post-editing is not easier but simply different from traditional from-scratch translation, and not putting sufficient effort into the post-editing process may lead to unsatisfactory results.

Taken together, our research findings support the view that translation and post-editing are at once similar and distinct processes. Both are very likely to remain essential activities of translator training in the future, although the balance between them may shift. Finally, raising students’ awareness of the specific demands and pitfalls of post-editing will be crucial in preparing them for a rapidly changing industry.

References

- Alvarez-Vidal, S., Oliver, A., Badia, T. 2021. What do post-editors correct? A fine-grained analysis of SMT and NMT errors. *Tradumàtica: Tecnologies de La Traducció* Vol. 19. 131–147. <https://doi.org/10.5565/rev/tradumatica.286>
- Belhassen, S., Hamda, A. 2025. Translation Students' Reliance on and Trust in Artificial Intelligence for Successful Translation Projects: Opportunities, Challenges, and Implications. *Arab World English Journal for Translation & Literary Studies* Vol. 9. No. 2. 106–119. <http://dx.doi.org/10.24093/awejtls/vol9no2.7>

- Cadwell, P., O'Brien, S., Teixeira, C. S. C. 2018. Resistance and accommodation: Factors for the (non-) adoption of machine translation among professional translators. *Perspectives* Vol. 26. No. 3. 301–321. <https://doi.org/10.1080/0907676X.2017.1337210>
- EMT Board and Competence Task Force. 2022. *European Masters in Translation Competence Framework*. https://commission.europa.eu/system/files/2022-11/emt_competence_fw_k_2022_en.pdf
- Läubli, S., Orrego-Carmona, D. 2017. When Google Translate is better than some human colleagues, those people are no longer colleagues. In: Esteves-Ferreira, J., Macan, J., Mitkov, R., Stefanov, O. (eds) *Proceedings of the 39th Conference on Translating and Computer; London, November 16-17*. Geneva: Tradulex. 59–69. <https://www.asling.org/te39/wp-content/uploads/TC39-proceedings-final-1Nov-4.20pm.pdf>
- Lesznyák M. 2026. *A longitudinal investigation of translation trainees' translation and post-editing competence development I. Translating an Academic Abstract*. Budapest: Akadémiai Kiadó. https://mersz.hu/dokumentum/m1364alitt_1/
- Lesznyák M., Bakti M., Sermann E. 2023. Translation and post-editing performance of translation students – a cross-sectional analysis. *L10N Journal* Vol. 2. No. 2. 7–23. <https://l10njournal.net/index.php/home/issue/view/5>
- Lesznyák M., Bakti M., Sermann E. 2024. Reading takes it all? – The role of language competence and subject knowledge in legal translation. *The Interpreter and Translator Trainer* Vol. 18. No. 2. 192–211. <https://doi.org/10.1080/1750399X.2024.2344177>
- Lesznyák, M., Sermann, E., Bakti, M. 2023. A fordításra vonatkozó nézetek szerepe a fordítási teljesítményben. *Fordítástudomány*. 25. évf. 2. szám. 5–19. <https://doi.org/10.35924/fordtud.25.2.1>
- Lesznyák M., Sermann E., Bakti M. 2025. The role of beliefs about translation in translation performance. In: Schmidhofer, A., Angeles, M., Ariza, R. (eds) *Zukunftsperspektiven in der Translationswissenschaft/Future Perspectives in Translation and Interpreting Studies.: Ausgewählte Beiträge der Translata IV / Selected Papers of Translata I*. Innsbruck: Innsbruck University Press. 71–92. <https://doi.org/10.15203/99106-165-6-06>
- Nitzke, J., Oster, K. 2016. Comparing translation and post-editing: An annotation schema for activity units. In: Carl, M., Bangalore, S., Schaeffer, M. (eds) *New directions in empirical translation process research: Exploring the CRITT TPR-DB*. Springer. 293–308. https://link.springer.com/chapter/10.1007/978-3-319-20358-4_14
- Nitzke, J., Hansen-Schirra, S. 2021. *A short guide to post-editing. (Translation and Multilingual Language Processing 16)*. Berlin: Language Science Press. <https://zenodo.org/records/5646896>
- Nitzke, J., Hansen-Schirra, S., Canfora, C. 2019. Risk management and post-editing competence. *The Journal of Specialised Translation* Vol. 31. 239–259. <https://doi.org/10.26034/cm.jostrans.2019.185>
- PACTE. 2003. Building a Translation Competence Model. In: Alves, F. (ed.) *Triangulating Translation: Perspectives in Process Oriented Research*. Amsterdam: John Benjamins. 43–66. <https://doi.org/10.1075/btl.45>
- PACTE. 2008. First results of a translation competence experiment: 'Knowledge of translation' and 'Efficacy of the translation process'. In: Kearns, J. (ed.) *Translator and interpreter training. Issues, methods and debates*. London: Bloomsbury. 104–126. <https://scispace.com/pdf/first-results-of-a-translation-competence-experiment-577jtj7d7p.pdf>

- PACTE. 2011. Results of the Validation of the PACTE Translation Competence Model: Translation Problems and Translation Competence. In: Alvstad, C., Hild, A., Tiselius, E. (eds) *Methods and Strategies of Process Research: Integrative Approaches in Translation Studies*. Amsterdam: John Benjamins. 317–343. <https://doi.org/10.1075/btl.94.22pac>
- PACTE. 2014. First Results of PACTE Group’s Experimental Research on Translation Competence Acquisition: The Acquisition of Declarative Knowledge of Translation. *MonTI Special Issue – Minding Translation*. 85–115. <https://doi.org/10.6035/MonTI.2014.ne1.2>
- Quinci, C. 2023. “To Use or Not to Use MT?”: Some Insights into Trust and Reliance From the LeMaTTT Project. *Rivista Internazionale di Tecnica della Traduzione* Vol. 25. 147–165. <https://doi.org/10.13137/2421-6763/35502>
- Robert, I. S., Remael, A., Ureel, J. J. J. 2017. Towards a model of translation revision competence. *The Interpreter and Translator Trainer* Vol. 11. No. 1. 1–19. <https://doi.org/10.1080/1750399X.2016.1198183>
- Rojo López, A. M., Cifuentes Férez, P., Espín López, L. 2021. The influence of time pressure on translation trainees’ performance: Testing the relationship between self-esteem, salivary cortisol and subjective stress response. *PLoS ONE* Vol. 16. No. 9. <https://doi.org/10.1371/journal.pone.0257727>
- Sakamoto, A. 2019. Why do many translators resist post-editing? A sociological analysis using Bourdieu’s concepts. *The Journal of Specialised Translation* Vol. 31. 201–216. <https://doi.org/10.26034/cm.jostrans.2019.183>
- Selinker, L. 1972. Interlanguage. *IRAL: International Review of Applied Linguistics in Language Teaching* Vol. 10. No. 1–4. 209–231. <https://doi.org/10.1515/iral.1972.10.1-4.209>
- Zhang, J., Doherty, S. 2025. Investigating novice translation students’ AI literacy in translation education. *The Interpreter and Translator Trainer* Vol. 19. No. 3–4. 1–20. <https://doi.org/10.1080/1750399X.2025.2541478>