

Translators' Resource Dominance and the Success of Finding the Target Terms in Human Translation and Post-editing of Machine Translation

Vertėjų išteklių dominavimas ir sėkminga vertimo kalbos terminų paieška verčiant žmogui arba postredaguojant mašininį vertimą

ESZTER SERMANN

Department of Italian Language and Literature, University of Szeged, Hungary

ORCID id: <https://orcid.org/0000-0001-8753-5312>

MÁRTA LESZNYÁK

Institute of English and American Studies, University of Szeged, Hungary

ORCID id: <https://orcid.org/0000-0002-9354-0633>

ABSTRACT

This study presents the first results of process-oriented research on the types of on-line translation resources used by first- and second-year translation trainees when translating and post-editing a legal text from English into Hungarian. Based on the screen recordings of the students' workflow, the possible relations between resource dominance (termino-lexicographic or text-based), time on task and the success of finding the correct target terms were analysed. Our results indicate that students generally prefer termino-lexicographic sources to text-based sources. Interestingly, in most cases, the success of finding the correct target terms showed no significant correlations either with time on task or with resource dominance. The only exception was the post-editor group, where there was a significant correlation between the frequency of using text-based sources and the success of finding correct terms. In addition, evidence was found that post-editors worked more efficiently than from-scratch human translators in terms of time and search effort. The paper ends with possible explanations of the findings and suggestions for translator training.

KEYWORDS: translation, post-editing, translation process research, translation resources, adequacy of terminology.

ANOTACIJA

Šiame tyrime pateikiami pirmieji į vertimo procesą orientuoto tyrimo apie internetinių vertimo išteklių tipus, kuriuos naudoja pirmo ir antro kurso vertėjai praktikantai, versdami teisinį tekstą iš anglų kalbos į vengrų kalbą ir jį post-redaguodami, rezultatai. Remiantis studentų darbo eigos ekrano įrašais, buvo analizuojami galimi ryšiai tarp išteklių (terminografinių ir leksikografinių (angl. *termino-lexicographic*) ar tekstinių) dominavimo, užduočiai skirto laiko ir sėkmingo teisingų vertimo kalbos terminų radimo. Mūsų rezultatai rodo, kad studentai dažniausiai teikia pirmenybę terminografiniams ir leksikografiniams, o ne tekstiniams šaltiniams. Įdomu tai, kad daugeliu atvejų teisingų vertimo kalbos terminų sėkmingas suradimas neparodė reikšmingos koreliacijos nei su užduočiai skirtu laiku, nei su išteklių dominavimu. Vienintelė išimtis buvo postredaktorių grupė – nustatytas reikšmingas ryšys tarp tekstinių šaltinių naudojimo dažnumo ir sėkmingo teisingų vertimo kalbos terminų radimo. Be to, nustatyta, kad postredaktoriai dirbo efektyviau nei vertėjai praktikantai tiek laiko, tiek ir vertimo kalbos terminų paieškos pastangų atžvilgiu. Straipsnis baigiamas galimais tyrimo rezultatų paaiškinimais ir pasiūlymais dėl vertėjų mokymo.

ESMINIAI ŽODŽIAI: vertimas, postredagavimas, vertimo proceso tyrimas, vertimo ištekliai, terminijos tinkamumas.

1. INTRODUCTION

Translation tools have changed radically over the last two decades, with online digital resources almost completely replacing paper-based tools. In today's technological environment, many online digital resources are available to translators, but choosing reliable and effective tools requires care and routine. Research has confirmed that the use of online translation resources constitutes a significant proportion of the overall time spent on translation (Hvelplund 2017); moreover, constant technological changes have led to an increased interest in the types of translation resources and their use in the process of human translation and post-editing of machine translation (e.g. Gough 2019; Hvelplund 2016, 2017; Prieto Ramos 2021). In recent years, post-editing of machine translation has become a common practice for translators, but one of the major shortcomings of machine translation engines is that they produce a high number of terminological errors and inconsistencies, so post-editors must check the correctness and consistency of target language equivalents, which is also a time-consuming process requiring a high degree of precision.

The research presented here focused on determining what resources translation trainees prefer when translating or post-editing a legal text, and

how resource use dominance is linked to performance indices like time on task and correct terminology use. First, a brief literature review is provided focusing on translation process research, online translation resources and translation-oriented research activity. Next, the Szeged Translation Competence research project is described, and the methodology of the present investigation is detailed. This is followed by the presentation and discussion of the results and a conclusion.

2. LITERATURE REVIEW

2.1. Translation Process Research (TPR)

Translation process research began to proliferate in the 1980s (Károly 2022; Lesznyák 2024) when researchers turned their attention to the cognitive, psychological and behavioural aspects of translation. Early process-oriented research focused on translation competence and the development of research methods (e.g. Wilss 1988; Tirkkonen-Condit, Jääskeläinen 2000). In the 21st century, technological advances have brought radical changes not only to translation practice, but also to translation research (see Klaudy 2022 for more details), including new research tools and methods on, among other things, the psychological factors of translation (Angelone 2010), translation strategies, translator behaviour (Dragsted 2005; Lesznyák 2008), pedagogical issues in translation (Shreve 2006; Ericsson 2010), the process of simultaneous translation (Seeber 2015), and the physical environment of translators (Ehrensberger-Dow, Hunziker Heeb 2016).

In translation process research, scholars use the following methods to observe translators: think-aloud (TAP), keystroke logging, screen recording, eye-tracking, retrospective interviews, and direct observation of the translation process (Risku 2019; Károly 2022; Lesznyák 2024), and then try to identify the regularities related to the translation process. TPR is usually not individual research, as data collection and analysis require collaborative work (Risku 2019), and quantitative and qualitative methods are often combined.

2.2. The use of translation resources as an element of translation competence

Types and use of translation resources, cognitive effort, and patterns of interaction between professional translators and online resources are widely studied topics (Gough 2018; Lesznyák 2008; Hvelplund 2016, 2017; Prieto Ramos 2021 etc.). Different classifications of translation resources can

be found in the terminological and lexicographical literature (Fóris 2019; Gaál 2016; Sermann 2021; Tamás 2014).

The skills related to the use of translation resources are included – either explicitly or implicitly – in the different models of translation competence. The PACTE Research Group (Hurtado Albir 2017) classifies the use of resources as an instrumental sub-competence, while in the latest version of the EMT model (2022), information mining competence does not appear as a separate element, but the correct use of search engines and the use of corpus-based resources are included in technological competence. The TransCert model (Budin et al. 2013) includes information mining and technological competence, while the eTransFair model (2016) includes information retrieval and terminology competence. In summary, the effective use of translation resources plays a role in some form in each of the translation competence models examined.

2.3 Translation-oriented research activity (TRA)

We based our study on Hvelplund's research about translators' use of digital resources during translation (Hvelplund 2016, 2017) and Gough's research about translation-oriented research activity (Gough 2016, 2018, 2019, 2023). Hvelplund's findings (2017) show that digital resource consultation constitutes a considerable amount of the overall translation task time (20 per cent), and translators mostly used general language dictionaries, bilingual dictionaries and Internet search engines for both literary and technical translation. Gough (2019) used a quasi-naturalistic, observational method via screen recordings to study professional translators' research activity. The author identified four resource behaviour types: Dictionary Enthusiast, Parallel Text Fan, Mixed User and MT Adopter. Her results suggest that domain expertise or a lack of it plays a role in the resource behaviour profile.

In a previous study, we investigated the types of online translation resources used by first- and second-year students enrolled in a Master of Translation and Interpreting programme when translating and post-editing the same legal text from English into Hungarian. Results show that first-year students were more likely to use monolingual and bilingual dictionaries, while second-year students opted for more corpus-based resources. The choice of the online translation resources was not determined by the mode of translation (HT or PE), but by the translator's level of expertise (Sermann 2023).

In the present study, based on the screen videos of students' workflow, we tried to find correlations between translator trainees' resource dominance, the time on task and the success of finding the correct target terms. Resource dominance was operationalized as the number of the types of different resources used and the frequency of consulting these sources, which led to two categories: (a) termino-lexicographic and (b) text-based.

3. THE SZEGED TRANSLATION COMPETENCE RESEARCH PROJECT – AIMS AND OBJECTIVES

The research reported here is part of the Translation and MT Post-Editing Competence Research Centre's overarching research project (University of Szeged, Hungary), which aims to investigate the role that elements of translation competence (PACTE) play in human translation and the post-editing of machine translation. In the project, we have been working with the variables of source language skills (English), L1/target language (Hungarian) competence, professional background knowledge related to the content of the source text (bilingualism, copyright), declarative knowledge of translation, the source language text type, translation experience, work mode (HT or PE), and students' perceptions of the advantages and disadvantages of working with each method (Lesznyák, Bakti, Sermann 2023, 2024). Having been involved in product-oriented research, we are now presenting the first results of our translation process research.

In the investigation reported here, our research questions were the following:

1. What are the resource preferences of the different subgroups (1st and 2nd year students, HT and PE students)?
2. How is students' resource dominance related to the success of finding the target equivalents of the key terms?
3. Is there a correlation between resource dominance and time on task?
4. Is there a correlation between the success of finding the correct target terms and time on task?

4. METHODOLOGY

4.1. Participants and material

14 first-year students (at the beginning of their studies) and 12 second-year master's students of translation (at the end of their studies) formed the sample of the present study. It is important to stress here that at the University

of Szeged, in Hungary, translator training takes place at the master's level exclusively, within the framework of a four-semester translator and interpreter training program. Therefore, first-year master's students could be viewed as highly skilled language learners or users with no prior translation experience. Additionally, it's worth noting that the students did not receive any specific training in post-editing at the time the data was collected, but second-year students had already completed the courses 'Basics of Law' and 'Legal Translation'.

The sample was divided into two subgroups (translator or post-editor), and they were asked to translate or post-edit a 350-word long legal text (a part of a copyright agreement, see Appendix). The source language was English, the target language was Hungarian and English was the B or C language of the translator trainees, and Hungarian was their A language. The MT output for the post-editors was produced by eTranslation, the machine translation tool of the EU.

4.2. The data collection procedure

Data collection took place in Autumn 2022 and in Spring 2022. All the translations were prepared in a classroom setting, and there was a time limit of 120 minutes for the task, but that limit was never actually reached. Students had internet access and were allowed to use whatever sources and web pages they wanted to use. Nevertheless, students in the HT condition were instructed to refrain from MT. Students were asked to produce a target text of publishable quality (i.e., full post-editing of the MT output), without using TM software. Students worked in Translog, as data was collected on the translation/post-editing process, too. In addition, OBS Studio was used to record the computer screen while students were working on the translation/post-editing task. To establish time on task, video recordings were used. The length of the translation/post-editing process was counted from the moment the 'start logging' button was pushed in Translog until the 'stop logging' button was pushed. Because of a technical breakdown, some data was lost, as a result, we could use 22 recordings in total.

4.3. Methods of data analysis

Based on the screen videos of students' workflow, measures on resource use were created. First, each resource students consulted was categorized either as termino-lexicographic (LEX) or as text-based (TXT). Within

each group two categories were created: (1) the number of the specific types of resources used (e.g. 3 dictionaries) and (2) the total frequency of consulting these sources (e.g. 11 searches /in the 3 dictionaries). This way four variables were created that were used in the quantitative analysis: LEX number, LEX frequency, TXT number, and TXT frequency.

Afterwards, 10 key terms were selected from the source text (see Appendix), and their equivalents were evaluated in the target texts, which were translated or post-edited by the students. Seven of the terms belong to the legal domain and three of them to the academic domain. The correct Hungarian equivalents of the legal terms were determined by an expert in copyright law.

For the statistical analysis, SPSS v. 26 was used. The analysis focused on determining the resource dominance of the subgroups with the help of paired samples t-tests, comparing the resource dominance of different subgroups with independent samples t-tests, and we also tried to find correlations between resource dominance, success in finding appropriate terminology and time on task. Because of space restrictions, we will refrain from sharing large amounts of data with non-significant results and will focus on significant findings.

5. RESULTS

5.1. Differences in resource dominance within the subsamples

Our first research question aimed at finding out what type of resources the different subgroups preferred. First, paired samples t-tests were carried out to compare the preferences within the first-year and second-year samples. On the one hand, we compared how many different types of resources they used, and on the other hand, the frequency of consulting these resources was contrasted.

Within the first-year sample ($n = 12$), it was found that students used significantly more termino-lexicographic ($M = 2.42$, $SD = 1.68$) than text-based resources ($M = 1.33$, $SD = 1.23$, $t(11) = 3.463$, $p = .005$). In line with this, they also did significantly more searches in termino-lexicographic ($M = 18.00$, $SD = 13.86$) than in text-based sources ($M = 5.17$, $SD = 5.80$, $t(11) = 3.736$, $p = .001$).

In the second-year sample ($n = 10$), no significant differences were found in the types of sources used by the students (termino-lexicographic $M = 1.90$, $SD = 0.74$, text-based $M = 1.40$, $SD = 1.27$, $t(9) = 1.246$, $p = .244$).

In other words, the termino-lexicographic and text-based sources were used to approximately the same extent, and the minor difference in mean values may be due to measurement error. However, analysing the frequency of searches has shown that second-year students, just like their first-year peers, were engaged in a significantly higher number of termino-lexicographic searches ($M = 17.2$, $SD = 11.10$) than in text-based searches ($M = 5.4$, $SD = 5.83$, $t(9) = 3.190$, $p = .011$). As can be seen, the standard deviation for text-based searches was higher than the mean, which is a signal of high variation in and abnormal distribution of the data. Consequently, medians were also compared for the frequency of searches. The paired samples Wilcoxon-test showed that the medians were significantly different, too, with the effect size being large (termino-lexicographic MDN = 14.50, text-based MDN = 3.50, $Z(9) = -2.497$, $p = .013$, $r = 0.79$).

In the next step, the sample was divided into human translators and post-editors, and resource preferences were analysed with paired samples t-tests once again. In the human translator group ($n = 10$) no significant difference was found between the number of the two types of resources used (termino-lexicographic $M = 2.70$, $SD = 1.16$, text-based $M = 1.90$, $SD = 0.99$, $t(9) = 1.809$, $p = .104$). Nevertheless, figures showed that students did significantly more searches in termino-lexicographic ($M = 27.8$, $SD = 9.36$) than in text-based sources ($M = 8.5$, $SD = 5.87$, $t(9) = 4.636$, $p = .001$).

Results of the t-tests indicated that the post-editors used significantly more termino-lexicographic ($M = 1.75$, $SD = 1.36$) than text-based resources ($M = 0.92$, $SD = 1.24$, $t(11) = 2.802$, $p = .017$). As the standard deviation for text-based sources was larger than the mean indicating abnormal distribution, the paired samples Wilcoxon-test was used to compare the medians. The test results showed that there was a significant difference between the medians, although the probability of measurement error was slightly higher (termino-lexicographic MDN = 2, text-based MDN = 0, $Z(11) = -2.226$, $p = .026$, $r = 0.643$), but the effect size was still moderate to strong.

The paired samples t-test showed that post-editors carried out significantly more searches in termino-lexicographic ($M = 9.17$, $SD = 7.03$) than in text-based sources ($M = 2.58$, $SD = 3.99$, $t(11) = 4.026$, $p = .011$). Because of abnormal distribution, the paired samples Wilcoxon-test was applied again, to check whether medians were significantly different. Results of the tests were significant again, suggesting that the difference between the median number of searches in termino-lexicographic sources (MDN = 8)

and in text-based sources (MDN = 0) is significantly different ($Z(11) = -2.805$, $p = .005$, $r = 0.810$) with the effect size being large.

5.2. Resource dominance differences between the subsamples

After characterizing the individual subsamples, we compared the groups, that is, first-year and second-year students, on the one hand, and human translators and post-editors, on the other hand, with each other. No significant differences were found between first-year and second-year students' research activities on either of the measures.

However, when human translators and post-editors were compared, all the differences in the frequency of searches were significant and the differences concerning the types of sources used were marginally significant (see table 1).

Table 1. Resource dominance differences between human translators and post-editors. Results of the independent samples t-tests

	HUMAN TRANSLATORS (N = 10)		POST-EDITORS (N = 12)		T-TEST	P
	M	SD	M	SD		
Number of termino- lexicographic resources	2.70	3.08	1.16	1.36	1.74	.096
Number of text-based resources	1.90	0.99	0.92	1.24	2.02	.057
Total number of resources used	4.60	1.65	2.75	2.38	2.08	.051
Frequency of searches (LEX)*	27.80	9.60	9.17	7.03	5.33	.000
Frequency of searches (TXT)*	8.50	5.87	2.50	3.99	2.81	.011
Frequency of searches (total)*	36.30	8.41	11.75	9.93	6.18	.000

Note: * = significant differences

5.3. Number of correct equivalents of terms found

To answer research questions 2 and 4, the success of finding the correct target equivalents had to be determined. Mean values (HT = 1.40, PE = 1.67, 1st year = 1.42, 2nd year = 1.70) show that students were not particularly

successful in finding the equivalents. Most students could get one or two equivalents right out of 10. Three students could not find any of the terms at all, and the maximum value, indicating the largest number of hits was 4. The differences between the first- and second-year students, on the one hand, and between human translators on the other hand were not significant.

5.4. Time on task

Research questions 3 and 4 focused on the temporal aspects of translation, consequently the mean time spent on the translation task itself had to be calculated for the subgroups. Mean translation and post-editing times can be seen in Table 2. It is also apparent that no significant differences were found between first-and second-year students, but post-editors worked significantly faster than human translators. Differences in standard deviation are also remarkable and Levene’s test has shown that the variances are also significantly different between the two groups ($F = 12.990$, $p = .002$), suggesting that the post-editor group shows a much larger variety concerning the time spent on the task than the HT group.

Table 2. Differences between human translators and post-editors and between first-year and second-year students on time on task. Results of the independent samples t-tests

	HUMAN TRANSLATORS (N = 10)		POST-EDITORS (N = 12)		T-TEST	P
	M	SD	M	SD		
time on task (min)	70.99	4.67	59.76	11.76	3.033	.008
	FIRST-YEAR STUDENTS (N = 12)		SECOND-YEAR STUDENTS (N = 10)		t-test	p
	M	SD	M	SD		
time on task (min)	65.99	10.94	63.52	10.83	0.529	.603

5.5. Correlations between resource dominance and success in finding appropriate terminology

The second research question aimed at revealing whether there is a relation between resource dominance and success in finding appropriate terminology. To answer this question, Pearson correlation coefficients were computed between indices of resource dominance, on the one hand, and

the number of correct equivalents found by the students. In the total sample, no significant correlations were found at all, and the same was true for the first year, the second year and the human translator subgroup, too. However, in the post-editor subgroup, a moderate to strong, significant correlation was found between the frequency of consulting text-based sources and the success of finding the right terms ($r = 0.676$, $p = .016$) suggesting that it pays off to use text-based sources during post-editing.

5.6. Correlations between resource dominance and time on task

The next step in the analysis was to compute Pearson correlation coefficients between indices of resource dominance and time on task. Table 3 shows a recurrent pattern for the total sample and all subsamples except for human translators. The pattern shows that there is a large number of positive, moderate to strong significant correlations between the number of termino-lexicographic (LEX) resources used and time, between the frequency of termino-lexicographic (LEX) searches and time and the frequency of searches in total and time.

Table 3. Correlations between indices of resource dominance and time on task in the total sample, the first-year sample, the second-year sample, the post-editor group and the human translator group (Pearson r)

		NO OF LEX RESOURCES	NO OF TXT SOURCES	NO OF TOTAL RESOURCES	FREQUENCY OF LEX SEARCHES	FREQUENCY OF TXT SEARCHES	FREQUENCY OF TOTAL SEARCHES
Time on task	First-year (n=12)	.62* (p=.03)	n.s.	n.s.	.66 (p=.02)	n.s.	.69** (p=.01)
	Second-year (n=10)	.70 (p=.03)	n.s.	n.s.	.68 (p=.03)	n.s.	.67 (p=.04).
	Human Translator (n=10)	n.s.	-.69 (p=.03)	n.s.	n.s.	n.s.	n.s.
	Post-editor (n=12)	.62 (p=.03)	n.s.	n.s.	.63 (p=.03)	n.s.	.60 (p=.04)
	Total sample (n=22)	.61 (p<.01)	n.s.	.52 (p=.01)	.66 (p<.01)	n.s.	.68 (p<.01)

5.7. Correlations between success of finding terms and time on task

The fourth research question tried to find out whether there is a correlation between the success of finding the correct equivalent of terms and the time spent on the total translation/post-editing task. Interestingly, no significant correlations were found either in the total sample or in any of the subgroups, suggesting that merely spending more time on the task will not contribute to finding the correct terms.

6. DISCUSSION

6.1. Differences in resource dominance within the subsamples

Our first research question aimed at describing students' resource dominance. Paired samples t-tests showed that first-year students used significantly more termino-lexicographic sources than text-based ones, and they also did significantly more searches in termino-lexicographic than in text-based sources. Second-year students demonstrated similar tendencies to first-year students in the frequency of searches, but not in the types of resources used.

Human translators could be characterized by a higher number of searches in termino-lexicographic sources than in text-based sources, but the number of termino-lexicographic sources was not significantly higher than the number of text-based sources. In the post-editor group, both the number of sources and the number of searches were significantly higher in the termino-lexicographic category than in the text-based category.

Generally, these findings suggest that students' translation-related research can be characterized by a strong termino-lexicographic dominance. Text-based sources (that is parallel texts) are consulted in a smaller number and less frequently. This is in line with Hvelplund's (2017) previous findings showing that translators heavily relied on dictionaries in the translation process.

The problem with interpreting this finding is that we do not have any benchmark data on what would be the ideal orientation. Most probably, there is no "perfect method", the proportion of termino-lexicographic or text-based sources that is rewarding in a particular context is dependent on several factors. These include the language pair, the quality of the termino-lexicographic sources in the language pair, the expertise of the translator/post-editor and perhaps on several other factors. As we will see below, however, students' failure to find most of the correct equivalents of the legal and educational terms suggests that overreliance on termino-lexicographic

sources may not have been the most fruitful strategy with this assignment. The dominance of termino-lexicographic sources reflects what practising translation trainers often see during training: students' trust in dictionaries (in whatever form) and reluctance to move on to more time-consuming and complicated research.

6.2 Resource dominance and the success of finding the target equivalent

The results indicate that students were relatively unsuccessful in finding the correct equivalents of terms, getting one or two terms right on average out of ten. Correlation analysis has brought no significant results, either, except post-editors and the frequency of consulting text-based sources. This implies that resource dominance is not directly linked to the success of finding the appropriate terms during translation. This makes sense, as it is not only what we use, but how we use the sources that can make a real difference. Unfortunately, without think-aloud data or retrospective interviews, it is impossible to discover what went wrong and at what stage of the research process.

As indicated in the previous paragraph, a moderate to strong, significant correlation was found between the frequency of consulting text-based sources and the success of finding the right terms in the post-editor group. Again, the finding sounds self-evident, as text-based sources are indispensable props of the translation and post-editing process. They are particularly useful when specialized dictionaries are not of the best quality or when the translator lacks subject knowledge and cannot appropriately evaluate hits in termino-lexicographic sources. Although neither of these conditions were empirically assessed, both of them look true in the research context. It must be added, that at this point of our research, it cannot be decided whether text-based sources provided the correct term(s) or they just helped the evaluation process of the student post-editor.

In addition, with the above explanation in mind, it is not clear why human translators had not profited from consulting text-based sources the way post-editors had. A possible explanation is that human translators were much more overwhelmed by research (which we have evidence for) and presumably, by other aspects of the translation task, too. In consequence they had no cognitive capacities left for deep processing information from text-based sources. Again, verbal data from the subjects would be needed to confirm this hypothesis.

6.3. The role of time on task

Research questions 3 and 4 concentrated on the role of time on task in resource dominance and in the success of finding the equivalent terminology. Our results confirm Hvelplund's (2017) findings although with a different methodology. The positive correlation between the frequency of total searches and time on task in the total sample and most subsamples indicates that the more searches a student did, the longer they translated/post-edited. In other words, total translation/post-editing time is closely related to resource consultation, particularly to using termino-lexicographic sources.

Surprisingly, however, time on task has not shown any significant correlations with finding the correct terms in the target language. This means that students who devoted more time to translation-related research did not do better in terms of terminological correctness. Once again, this implies problems of using resources productively.

6.4. Differences between human translators and post-editors

Although no research question focused directly on the differences between translators and post-editors, our findings offered some valuable insights into it. Comparisons of resource dominance provided evidence that human translators used a wider variety of sources and engaged in a considerably larger number of searches than post-editors. This can be interpreted as a larger effort on behalf of the human translators. In addition, it was also shown that human translators worked significantly longer than post-editors, which again, can be seen as a sign of increased effort. Paradoxically, the increased effort has not led to better performance in terms of terminology. Putting the pieces of the puzzle together, it looks like it takes human translators more time and more research effort to produce the same quality concerning the terminology of a translation task as post-editors. In other words, post-editing appeared to be more efficient in this experimental setting.

7. CONCLUSION

The study aimed to reveal what resources translation trainees prefer when translating or post-editing a legal text, and how resource use dominance is linked to performance indices like time on task and correct terminology use. Our findings show, that students tend to prefer termino-lexicographic sources over text-based ones. It is also striking that human translators make significantly more efforts both in terms of time and research than

post-editors but they are not more successful in finding the appropriate terms. Post-editors, on the other hand, engage in a smaller number of searches, thus, they achieve the same quality with less effort. Nevertheless, the reduced research activity can be seen as a negative tendency too, as it is very much in line with what we, as trainers see in post-editing classes, that is, students' unfounded trust in MT. One could argue that with a bit more effort and appropriate research techniques, students could improve the terminological quality of post-edited target texts.

The findings reported in the study shed light on a pattern of failed effort, too: some students are inclined to spend more time on tasks (whether from scratch translation or post-editing), which is closely linked to more termino-lexicographic searches. Nevertheless, they are not more successful than the students who save time, among others, by not engaging in so many searches. A probable explanation for this failure was offered above: the way students evaluate the information they find may not be appropriate. In addition, the lack of (legal) background knowledge can be linked to this problem, as students without sufficient thematic knowledge may not be in the position to understand, interpret and evaluate whatever they find in termino-lexicographic or in text-based sources. This leads us to the question (of how far) /re/search can compensate for missing thematic knowledge. It is a question that future research could focus on. The correlation found between the use of text-based sources and terminological success in the post-editing condition may offer a way out of this trap suggesting that it is worth investing in consulting text-based sources.

The research results have implications for translator training. Although the students participating in the study had terminology classes, they seemed to lack research skills. Consequently, it can be assumed that the methodology of teaching research behaviour should be carefully designed. Students should understand that it is not (only) the quantity of searches that helps them find the correct equivalents but the “quality of the search activity”, too, that is, what they do with the information they find. Presumably, evaluating sources and hits in sources should be practised in class. In addition, the importance of text-based sources should be emphasized and students must be encouraged to learn about the topic of the translation. Otherwise, they will not be able to make informed decisions about terms.

The study has some limitations, too, one of which is the small sample size. Another weakness of the research is that no professional translators

participated in the data collection. Furthermore, only one text, a legal one was used in the study and participants came from one institution. These factors limit the generalisability of the results, although significant findings may indicate valid tendencies for settings similar to that of the study. The lack of verbal data from participants prevented the interpretation of certain findings, although it is clear that one must set limits to the complexity of the research design, too.

The limitations of this study indicate the directions for future research. The study could be repeated with a larger sample involving professionals and students from other universities or countries. A different text type could be used and other types of process data could be collected.

Moreover, the data set of the present study has remained partly unexplored. Time spent on research has not been measured or related to indices of research behaviour or performance. The use of translation resources in the different stages of HT and PE (drafting, revision) could be studied, too, just like the number of terminology changes post-editors make and their relation to research strategies.

Despite its limitations, the study certainly adds to our understanding of translation trainees' resource use. Nevertheless, more work needs to be done in this field to discover what research strategies and behaviours are needed in the digital era where post-editing will likely dominate translators' work.

REFERENCES

- Angelone Erik 2010: Uncertainty, Uncertainty Management and Metacognitive Problem Solving in the Translation Task. – *“Translation and Cognition”*, ed. Shreve G. M., Angelone E., Amsterdam: American Translators Association Scholarly Monograph Series XV, 17–40.
- Budin Gerhard, Krajcso Zita, Lommel Arle 2013: The TransCert Project: Ensuring That Transnational Translator Certification Meets Stakeholder Needs. – *The International Journal for Translation & Interpreting Research* 5(1), 143–155. DOI: 10.12807/ti.105201.2013.a08.
- Dragsted Barbara 2005: Segmenting in Translation: Differences across levels of expertise and difficulty. – *Target* 17(1), 49–70. DOI:10.1075/target.17.1.04dra.
- Ericsson K. Anders 2010: Expertise in Interpreting: An expert-performance perspective. – *“Translation and Cognition”*, ed. Shreve G. M., Angelone E., Amsterdam: American Translators Association Scholarly Monograph Series XV, 231–262.
- EMT 2022: *Updated version of the EMT competence model framework*. Available at: https://commission.europa.eu/news/updated-version-emt-competence-framework-now-available-2022-10-21_en.
- eTransFair 2018: *Competence Card for Specialised Translators*. Available at: <https://etransfair.inyk.bme.hu/>.
- Fóris Ágota 2019: Új tendenciák és módszerek a fordítási folyamatban: a terminológia és a dokumentáció szerepe. – *“Diszciplínák találkozása a fordításban”*, ed. Szoták Sz., Budapest: OFFI, 211–219.
- Gaál Péter 2010: A hagyományos szótárak vizsgálati szempontjainak alkalmazhatósága az online szótárak esetében. – *“Az alkalmazott nyelvészet ma: innováció, technológia, tradíció. XX. Magyar Alkalmazott Nyelvészeti Kongresszus”*, ed. Boda I. K., Mónós K., Budapest/Debrecen: MANYE–Debreceni Egyetem, 432–437.

- Gough Joanna 2016: *The patterns of interaction between professional translators and online resources*. PhD Thesis, University of Surrey. Available at: <https://openresearch.surrey.ac.uk/esploro/outputs/99514578402346>.
- Gough Joanna 2018: Investigating the use of resources in the translation process. – “*Trends in e-tools and resources for translators and interpreters*”, ed. Pastor G. C., Durán-Muñoz I., Leiden: Brill, 9–36.
- Gough Joanna 2019: Developing translation-oriented research competence: what can we learn from professional translators? – *The Interpreter and Translator Trainer* 13:3, 342–359. DOI: 10.1080/1750399X.2019.1656404.
- Gough Joanna 2023: *Individual Variations in Information Behaviour of Professional Translators: Towards a Classification of Translation-oriented Research Styles*. Translation Studies. Taylor and Francis Online. Available at: <https://www.tandfonline.com/doi/full/10.1080/14781700.2023.2231933>.
- Ehrensberger-Dow Maureen, Hunziker Heeb Andrea 2016: Investigating the Ergonomics of a Technologized Translation Workplace. – “*Reembedding Translation Process Research*”, ed. R. Muñoz Martín, Amsterdam/Philadelphia: John Benjamins, 69–88.
- Hurtado Albir Amparo (ed.) 2017: *Researching Translation Competence by PACTE Group*, Amsterdam/Philadelphia: John Benjamins. DOI: 10.1075/btl.127.
- Hvelplund Kristian Tangsgaard 2016: Cognitive efficiency in translation. – “*Reembedding Translation Process Research*”, ed. R. Muñoz Martín, Amsterdam/Philadelphia: John Benjamins, 149–170.
- Hvelplund Kristian Tangsgaard 2017: Translators’ Use of Digital Resources during Translation. – *Hermes* 56, 71–87.
- Károly Krisztina 2022: A nyelvi közvetítés empirikus kutatásának módszerei. – “*Bevezetés a fordítás és a tolmácsolás kutatómódszerében I. Általános rész*”, ed. K. Klaudy, E. Robin, O. Seidl-Péché, ELTE FTT–MANYE Fordítástudományi Szakosztály, 93–108.
- Klaudy Kinga 2022: A fordítástudomány dilemmái a 21. század elején. – “*Bevezetés a fordítás és a tolmácsolás kutatómódszerében I. Általános rész*”, ed. K. Klaudy, E. Robin, O. Seidl-Péché, ELTE FTT–MANYE Fordítástudományi Szakosztály, 13–25. DOI: <https://doi.org/10.21862/kutmodszertan1/2>.
- Lesznyák Márta 2008: A szótárhasználat jellegzetességeinek változása a fordítói tapasztalat függvényében. – “*Az I. Alkalmazott Nyelvészeti Doktorandusz Konferencián elhangzott előadások kötete*”, ed. Várádi T., Budapest: MTA Nyelvtudományi Intézet, 67–79.
- Lesznyák Márta 2024: Folyamatorientált kutatások a fordítástudományban. – “*Bevezetés a fordítás és a tolmácsolás kutatómódszerében II. Speciális rész*”, ed. K. Klaudy, E. Robin, O. Seidl-Péché, ELTE FTT–MANYE Fordítástudományi Szakosztály, 71–94.
- Lesznyák Márta, Bakti Mária, Sermann Eszter 2023: Translation and post-editing performance of translation students – a cross-sectional analysis. – *L10N Journal* 2(2), 7–23.
- Lesznyák Márta, Bakti Mária, Sermann Eszter 2024: Reading takes it all? – the role of language competence and subject knowledge in legal translation. – *The Interpreter and Translator Trainer*, 1–20. Available at: <https://doi.org/10.1080/1750399X.2024.2344177>.
- Prieto Ramos Fernando 2021: Translating legal terminology and phraseology: between inter-systemic incongruity and multilingual harmonization. – *Perspectives* 29:2, 175–183. DOI: 10.1080/0907676X.2021.1849940.
- Risku Hanna 2019: Process research. – “*Routledge Encyclopedia of Translation Studies. 3rd Edition*”, ed. M. Baker, G. Saldanha, London: Routledge, 437–441.
- Seeber Kilian G. 2015: Cognitive Load in Simultaneous Interpreting: Measures and methods. – “*Interdisciplinarity in Translation and Interpreting Process Research*”, ed. M. Ehrensberger-Dow, S. Göpferich and S. O’Brien, Amsterdam/Philadelphia: John Benjamins, 18–33.
- Sermann Eszter 2021: *Terminológiai szabványosítás és nyelvi közvetítés*, Szeged: Juhász Gyula Felsőoktatási Kiadó.
- Sermann Eszter 2023: Online fordítói segédesszközök használata humán fordítás és a gépi fordítás utószekeresztése során. – “*XXIX. Magyar Alkalmazott Nyelvészeti Kongresszus*” (accepted for publication).
- Shreve Gregory 2006: The Deliberate Practice: Translation and expertise. – *Journal of Translation Studies* 9(1), 27–42.
- Tirkkonen-Condit Sonja, Jääskeläinen Riitta 2000: *Tapping and mapping the processes of translation and interpreting*, Amsterdam/Philadelphia: John Benjamins. Available at: <https://doi.org/10.1075/btl.37>.
- Tamás Dóra Mária 2014: *Gazdasági szakszövegek fordításának terminológiai kérdéseiről*, Budapest: ELTE.
- Wilss Wolfram 1988: *Kognition und Übersetzen: Zu Theorie und Praxis der menschlichen und der maschinellen Übersetzung*, Tübingen: Niemeyer.

APPENDIX 1.

The source text used for the study:

LICENCE TO PUBLISH

Manuscript #: [name of journal] (“The Journal”)

Title of the contribution: (“The Contribution”)

Author(s): (“the Authors”)

To: Nature Publishing Group (“NPG”), a division of Macmillan Publishers Ltd

1. In consideration of NPG agreeing to publish the Contribution the Authors grant to NPG for the full term of copyright in the Contribution and any extensions thereto, subject to clause 2 below, the exclusive licence (a) to publish, reproduce, distribute, display and store the Contribution in all forms, formats and media whether now known or hereafter developed (including without limitation in print, digital and electronic form) throughout the world, (b) to translate the Contribution into other languages, create adaptations, summaries or extracts of the Contribution or other derivative works based on the Contribution and exercise all of the rights outlined in (a) above in such translations, adaptations, summaries, extracts and derivative works and (c) to license others to do any or all of the above.
2. Ownership of copyright remains with the Authors and provided that, when reproducing the Contribution or extracts from it, the Authors acknowledge first and reference publication in the Journal, the Authors retain the following nonexclusive rights:
 - a) To reproduce the Contribution in whole or in part in any printed volume (book or thesis) of which they are the author(s).
 - b) They and any academic institution where they work at the time may reproduce the Contribution for course teaching.
 - c) To post a copy of the Contribution as accepted for publication after peer review (in Word or Tex format) on the Author’s own website or institutional repository six months after publication of the printed edition of the Journal, provided that they also give a hyperlink from the Contribution to the Journals web site
 - d) To reuse figures or tables created by them and contained in the Contribution of other works created by them.

This Agreement shall be governed by and construed in accordance with the laws of England without regard to the principles of conflicts of law. The parties hereto submit to the non-exclusive jurisdiction of the English courts.

APPENDIX 2.

The 10 target terms selected for the analysis:

licence to publish, contribution, full term of copyright, including without limitation, to grant exclusive licence, to reproduce, derivative work, thesis, academic institution, peer review

VERTĖJŲ IŠTEKLIŲ DOMINAVIMAS IR SĖKMINGA VERTIMO KALBOS TERMINŲ PAIEŠKA VERČIANT ŽMOGUI ARBA POSTREDAGUOJANT MAŠININĮ VERTIMĄ

Santrauka

Vertimo priemonių naudojimas iš esmės pasikeitė per pastaruosius du dešimtmečius – internetiniai skaitmeniniai ištekliai beveik visiškai pakeitė popierines priemones. Šiuolaikinėje technologinėje aplinkoje vertėjai gali naudotis daugybe internetinių skaitmeninių išteklių, tačiau norint pasirinkti patikimas ir veiksmingas priemones, reikia kruopštumo ir disciplinos. Šiame tyrime nagrinėjamas vertėjų praktikantų išteklių dominavimas tą patį teisinį tekstą verčiant žmogui ir postredaguojant ir analizuojami galimi ryšiai tarp išteklių dominavimo, užduoties atlikimo laiko ir sėkmingai rastų teisingų vertimo kalbos terminų.

Čia pateiktas tyrimas yra Segedo vertimo kompetencijos tyrimų grupės bendrojo tyrimo projekto, kuriuo siekiama ištirti vertimo kompetencijos elementų (PACTE) vaidmenį žmogaus vertime ir mašininio vertimą postredagavime, dalis. Kadangi dalyvaujame į produktą orientuotuose tyrimuose, dabar pristatome pirmuosius vertimo proceso tyrimo rezultatus.

Mūsų rezultatai rodo, kad terminografiniai ir leksikografiniai šaltiniai dominuoja su vertimu susijusiuose studentų tyrimuose. Be to, studentams ganėtinai nesisekė rasti teisingų atitikmenų terminams, jie rado vidutiniškai vieną ar du iš dešimties. Be to, nei laikas užduočiai, nei išteklių dominavimas neparodė jokių reikšmingų koreliacijų su tinkamų vertimo kalbos terminų radimu. Vertėjų ir postredaktorių palyginimas parodė, kad postredaktoriai dirbo greičiau ir atliko mažiau paieškų nei vertėjai. Nepaisant to, jų rastų terminų teisingumas buvo panašus į vertėjų.

Akivaizdu, kad tyrimas turi tam tikrų trūkumų, iš kurių vienas yra maža imtis, tačiau, nepaisant jų, išvados tikrai prisideda prie mūsų supratimo apie vertėjų praktikantų išteklių naudojimą.

Gauta 2024-05-03

Eszter Sermann

Department of Italian Language and Literature, University of Szeged, Hungary
6724 Szeged, Pulz utca 1.

E-mail eszter.sermann.lingo@gmail.com

Márta Lesznyák

Institute of English and American Studies, University of Szeged, Hungary
6722 Szeged, Egyetem utca 2.

E mail mlesz@edpsy.u-szeged.hu