Chapter 5

Investigating post-editing: A mixed-methods study with experienced and novice translators in the English-Greek language pair

Maria Stasimioti & Vilelmini Sosoni
Ionian University

In recent years, Post-Editing (PE) has been increasingly gaining ground, especially following the advent of neural machine translation (NMT) models. However, translators still approach PE with caution and skepticism and question its real benefits. This study investigates the perception of both experienced and novice translators vis-à-vis PE, it compares the technical, temporal and cognitive effort expended by experienced translators during the full PE of NMT output with the effort expended by novice translators, focusing on the English-Greek language pair and explores potential differences in the quality of the post-edited texts. The findings reveal a more negative stance of the experienced translators as opposed to novice translators vis-à-vis Machine Translation (MT) and a more pragmatic approach vis-à-vis PE. However, the novice translators' more positive attitude does not seem to positively affect the temporal and cognitive effort that they expend. Finally, experienced translators have a tendency to overcorrect the NMT output, thus carrying out more redundant edits.

1 Introduction

In recent years, the translation industry has seen a growth in the amount of content to be translated and has received pressure to increase productivity and speed at reduced costs. To respond to these challenges, it has turned to machine translation (MT) "which appears to be moving from the peripheries of the translation

field closer to the centre" (Koponen 2016: 131). The most common and widely expanding scenario – especially for certain language pairs and domains – involves the use of MT output to be then post-edited by professional translators (Koponen 2016). This practice is generally termed post-editing of machine translation (PEMT) or simply post-editing (PE) and is subcategorised into two types according to the required quality: full post-editing, which is expected to improve the final product to publishable quality, and light/rapid post-editing, which aims to correct the text for accuracy, but not style and fluency (Allen 2003).

PE has been increasingly gaining ground (O'Brien et al. 2014; O'Brien & Simard 2014; Lommel & DePalma 2016; Nunes Vieira et al. 2019), especially following the advent of neural machine translation (NMT) models which have been proven to consistently outperform statistical machine translation (SMT) models in shared tasks, as well as in various project outcomes (Bojar et al. 2016; Toral & Sánchez-Cartagena 2017; Castilho, Moorkens, Gaspari, Calixto, et al. 2017; Castilho, Moorkens, Gaspari, Sennrich, et al. 2017). In fact, NMT has been widely hailed as a significant development in the improvement of the quality of MT, especially at the level of fluency (Castilho, Moorkens, Gaspari, Calixto, et al. 2017; Castilho, Moorkens, Gaspari, Sennrich, et al. 2017), and the PE of NMT output has been found to be faster than translation from scratch (Jia et al. 2019).

However, translators still approach PE with caution and skepticism and question its real benefits (Gaspari et al. 2014; Koponen 2012; Moorkens et al. 2018). Their skepticism is directly related to the nature of PE which involves "working by correction rather than creation" (Wagner 1985: 2), to the belief that PE is slower than translating from scratch and to the view that MT is a threat to their profession (Moorkens et al. 2018: 58). Several studies were carried out aiming at identifying the extent to which attitudes to MT and PE affect PE effort. A positive attitude to MT has been found to be a factor in PE performance (De Almeida 2013; Mitchell 2015). Experienced translators have been found to exhibit rather negative attitudes to PE as opposed to novice translators (Moorkens & O'Brien 2015) and to be rather reluctant to take on PE jobs, while novice translators appear to be more positive towards MT and PE and more suited for PE jobs (García 2010; Yamada 2015). Previous research has shown that professional translators and novices generally exhibit different translation behaviour (Carl & Buch-Kromann 2010; Carl et al. 2010; Hvelplund 2011; Dragsted & Carl 2013; Moorkens & O'Brien 2015; Nitzke 2019; Schaeffer et al. 2019), while cognitive effort has been found to be greater for novice than for professional translators (Göpferich et al. 2011). Yet, there is still a lack of empirical studies about post-editing by different profiles (Mesa-Lao 2014). Under the light of the above, the

aim of this study is threefold: it seeks to investigate the perception of both experienced and novice translators vis-à-vis PE; it aims to compare the technical, temporal and cognitive effort (Krings 2001) expended by experienced translators with the effort expended by novice translators during the full PE of NMT output, focusing on the English-Greek language pair; finally, it aims to explore potential differences in the quality of the post-edited texts.

2 Related work

The comparison of the translation, revision and PE behaviour of student translators and professional translators has been the subject of several empirical studies.

Carl et al. (2010) compared the revision behaviour of 12 student translators with that of professional translators and found that professionals tend to have a longer revision phase after they have completed a first draft of the translation. Irrespective, though, of when the revision is made, students and professionals revised the same parts of the translations, presumably due to the fact that they face the same problems during translation.

Carl & Buch-Kromann (2010) analysed the user activity data (eye movements and keystrokes) of 12 students and 12 professional translators translating two small English texts into Danish. The human translations, as well as a machine translation produced by Google Translate, were evaluated and compared both automatically with BLEU and manually with human scores for fluency and accuracy. The results revealed that although both professionals and students produce equally accurate translations, professionals seem to be better at producing more fluent texts more quickly than students.

Dragsted & Carl (2013) asked 12 professional full-time translators with at least two years professional of experience in translation between Danish and English, and 12 MA students at the Copenhagen Business School, specialising in translation between Danish and English to translate three small English texts into Danish. Eye tracking and keylogging data were used in order to identify typical behaviour in translation students and professionals. The authors found differences between the two groups. In particular, contrary to the earlier findings by Jakobsen (2002), students carried out more initial planning than professionals, while professionals carried out more end revision than students. In addition, they found a tendency for professionals to prefer local orientation in the initial planning phase and during drafting and a more global perspective in the revision phase.

Moorkens & O'Brien (2015) carried out a study with nine expert translators (having on average 11.3 years of translation experience and 4 years of PE experience) and 35 undergraduate student translators, and studied the productivity/speed of the participants, the edit distance of their final product, and their attitude towards PE. The participants were asked to carry out two PE tasks from English into German. The professional group was much faster than the students, while the students tended to edit less of the MT output and they had a more positive attitude towards PE.

Nitzke (2019) asked twelve professional translators (university degree holders with at least some professional work experience) and twelve semi-professional translators (university students with limited professional work experience) to translate two texts from scratch, to bilingually post-edit two machine translated texts and to monolingually post-edit two machine translated texts. No significant difference between students and professionals was reported in terms of task duration and keystrokes (insertions and deletions).

Schaeffer et al. (2019) compared two different translation expertise levels, i.e. translation students and professional translators, and investigated the process of the revision of six English-German human pre-translated texts using eye-tracking and keystroke logging data. The results revealed that translation students correct significantly fewer errors and take longer to correct these errors as compared to professional translators.

The question of the role of translation expertise on PE has hardly been addressed. To the best of our knowledge, only two studies investigated the influence of translation expertise on PE (Moorkens & O'Brien 2015; Nitzke 2019), while only a handful of studies focused on the influence of the perception of MT and PT on PE (Moorkens & O'Brien 2015). In addition, no studies to date involved the role of translation expertise and Greek or focused on Greek translators' perception vis-à-vis MT and PE. We will therefore help fill this research gap by investigating PE carried out by experienced and novice translators in the English-Greek language pair.

3 Experimental setup

As already pointed out, the study adopts a mixed-methods approach and triangulates findings from different methods. For the investigation of the participants' perception of PE, pre-assignment questionnaires are used. For the assessment of PE effort and following Koponen et al. (2019), the study uses both process-based approaches and product-based approaches. In particular, keystroke logging and

eye-tracking data are used to measure the temporal, technical and cognitive effort expended by translators during the PE of NMT output generated by the NMT system developed by Google (Google Translate NMT system). Moreover, the MT output is compared with the final post-edited text to determine the number and type of changes using the edit distance metric WER; this has been widely used as an indicator of MT quality and PE effort. Finally, an analysis of the number and the type of each edit is carried out to evaluate the quality of the post-edited texts.

The PE experiments were carried out in March 2018 at the HUBIC lab[1] (Raptis & Giagkou 2016) of the Athena research center[2] in Athens.

A detailed consent form was signed by all participants prior to the execution of the experiments, while all stored data were fully anonymised in accordance with Greek law 2472/97 (as amended by laws 3783/2009, 3917/2011 and 4070/2012).

3.1 The participants

Twenty translators – ten experienced translators and ten novice translators – participated in the experiments, in which their eye movements and typing activity were registered with the help of an eye-tracker and specialised software (see Section 3.2). In the present study, following Whyatt & Kościuczuk (2013) and Colina & Angelelli (2015), experienced translators are selected among those with more than five years of translation experience and novice translators among those with less than five years of translation experience. Experienced translators are considered to be the "products of long years of deliberate practice" who exhibit "consistently high levels of performance" (Shreve 2002: 151), while they employ some "cognitive routines" occurring "automatically" (Kaiser-Cooke 1994: 137). On the other hand, novice translators are considered to "lack the routine processes acquired by experience" (Palumbo 2009: 130) and thus face more problems during a translation task. It should also be noted that in the present study neither experience nor training in PE was a prerequisite for participating in the experiments.

Their selection followed a call for participation, which was sent to the members of the two biggest Greek associations of professional translators, i.e. the Panhellenic association of translators[3] (PEM) and the Panhellenic association of professional translation graduates of the Ionian University[4] (PEEMPIP), and was

---

[1]http://www.hubic-lab.eu/

[2]https://www.athenarc.gr/en

[3]http://www.pem.gr/el/

[4]http://peempip.gr/el/

shared on social media. Translators expressed their interest for participating in the study by filling in a Google form; they subsequently received an e-mail with details on the aim of the research and guidelines for the PE task along with some educational material on PE (see Section 3.3). In addition, they were asked to fill in a questionnaire before the actual experiment. The questionnaire consisted of 34 questions: 22 closed-ended (multiple choice) and 12 open-ended, all of which aimed at defining the profile of the participants and their perception of MT and PE.

The following graphs and tables provide information regarding the participants' profiles.

Table 5.1: Participants' gender, age distribution, education level and degree type

		Translators	
		Experienced	Novice
Gender	Male	0	1
	Female	10	9
Age distribution	20−30	3	8
	30−40	4	1
	40−50	3	1
Education level	Undergraduate	4	4
	Postgraduate	6	6
Degree type	Translation	7	8
	Language/linguistics	2	1
	Other	1	1

It should be noted that all participants had normal or corrected to normal vision. Five wore contact lenses and four wore glasses, yet the calibration with the eye-tracker was successful for all of them.

All experienced translators had many years of translation experience (see Table 5.2) compared to novice translators who had either 0 years of translation experience (3), 1 year of translation experience (5), 2 years of translation experience (1) or 3 years of translation experience (1).

Table 5.2: Participants' years of translation experience

	Translators	
Years of translation experience	Experienced	Novice
1−5	0	10
5−10	5	0
10−20	4	0
Over 20	1	0

Table 5.3: Participants' years of experience in PE

	Translators	
Years of experience in PE	Experienced	Novice
0 years	1	10
1 year	2	0
2 years	2	0
3 years	3	0
4 years	0	0
5 years	1	0
over 5 years	1	0

As far as experience in PE is concerned, none of the novice translators had experience in PE, compared to almost all of the experienced translators who had at least 1 year of experience in PE[5] (see Table 5.3).

As can be seen in Table 5.4, although novice translators had no experience in PE, half of them had received relevant training in the past. Half of the experienced translators had also received previous training in the past. More interestingly, as it emerges from Tables 5.5 and 5.6, all the novice translators would be interested in attending a PE course in the future, considering it to be either important or very important, while 7 experienced translators would be interested in attending a PE course in the future, considering it fairly important (5), important (4) or very important (1).

---

[5]One year of experience in PE corresponds to one year full-time equivalent (FTE) and it represents the number of working hours that one full-time worker completes during a year or during any other fixed time period.

Table 5.4: Participants' previous training in PE

	Translators	
Previous training in PE	Experienced	Novice
Yes	5	5
No	5	5

Table 5.5: Participants' interest in PE training

	Translators	
Interest in PE training	Experienced	Novice
Yes	7	10
No	3	0

Table 5.6: Participants' view on the importance of PE training

	Translators	
Importance of PE training	Experienced	Novice
Very important	1	5
Important	4	5
Fairly important	5	0
Slightly important	0	0
Not important	0	0

3.2 Description of the experiment

A Tobii TX-300 eye-tracker and the Translog-II software (Carl 2012) were used to register the participants' eye movements, keystrokes and time needed during the PE tasks they were asked to carry out. The participants were asked to work at the speed at which they would normally work in their everyday work as translators; therefore, no time constraint was imposed. However, they did not have Internet access and were not allowed to use online or offline translation aids as this could lead to a reduction in the amount of recorded eye-tracking data. In the case of offline translation aids (e.g. dictionaries), the participants might look away from the screen resulting in a reduced amount of analysable eye-tracking data; in the case of online translation aids (Internet), the eye-tracking data would partially reflect gaze activity that does not directly reflect source text (ST) processing, target text (TT) processing or parallel ST/TT processing (Hvelplund 2011: 86).

The experiment consisted of one session for each participant. Before the sessions, the participants were informed by email about the nature of the experiments, the task requirements and the general as well as task-specific guidelines they had to follow (see Section 3.3). The session started with a warm-up PE task to familiarise each participant with the procedure. After the warm-up task, the participants had to carry out full PE of the NMT output of the same two semi-specialised texts – in accordance with the detailed task-specific guidelines they received (see Section 3.3.2); the two texts were presented to them in the same order. During the experiment, the ST was displayed in the Translog-II software at the top half of the screen and the MT output at the bottom half, as suggested by previous studies (Hvelplund 2011; Mesa-Lao 2014; Carl et al. 2011; 2015; Schmaltz et al. 2015). Translators worked directly on the MT output. To facilitate eye-tracking measurements, texts were fully displayed to avoid any need for participants to scroll in either the ST or the TT window. For the purposes of this study, each ST and each TT was considered an Area of Interest.

The STs used in this study were short educational texts selected from OER Commons[6], which is a public digital library of open educational resources. Three[7] excerpts of around 140 words each were selected from various courses on social change and the endocrine system and the titles of the courses were retained as context information for the participants. The texts were chosen with the following criteria in mind: they had to be semi-specialised and easy for participants to

---

[6]https://www.oercommons.org/

[7]One text was used exclusively for the warm-up session and is not included in the ensuing analysis and discussion.

post-edit without access to external resources and they also had to be of comparable complexity. The texts chosen had comparable Lexile® scores (between 1300L and 1400L), i.e. they were suitable for 11th/12th graders. The Lexile Analyzer[8] was used as it relies on an algorithm to evaluate the reading demand – or text complexity – of books, articles and other materials. In particular, it measures the complexity of the text by breaking down the entire piece and studying its characteristics, such as sentence length and word frequency, which represent the syntactic and semantic challenges that the text presents to a reader. The outcome is the text complexity, expressed as a Lexile measure, along with information on the word count, mean sentence length and mean log frequency.

Table 5.7: Lexile® scores for the source texts used in the study

	Text 1–T1	Text 2–T2
Lexile® measure	1300L–1400L	1300L–1400L
Number of sentences	5	6
Mean sentence length	28.60	22.67
Word count	143	136
Characters without spaces	647	1761

The NMT-core engine used to produce the raw MT output was Google Translate (output obtained March 24, 2018). The MT output was evaluated using the bilingual evaluation understudy (BLEU) metric. Generally speaking, a score below 15 percent means that the engine is not performing optimally and PE is not recommended as it would require a lot of effort to finalise the translation and reach publishable quality, while a score above 50 percent is a very good score and means that significantly less PE is required to achieve publishable translation quality. The BLEU score was calculated for the two texts using the Tilde custom machine translation toolkit[9]. Both texts had a very good score. In particular, BLEU score for Text 1 was 51.33 and for Text 2 was 60.62. This means that PE could be used to achieve publishable translation quality in both cases.

3.3 PE task guidelines

In the PE task, the participants were asked to fully post-edit the raw output generated by the NMT system. Since previous training and experience in PE was

---

[8]https://lexile.com/
[9]https://www.letsmt.eu/Bleu.aspx

not a prerequisite for participating in the study, the participants received brief training in PE before executing the task. The training included a video and a presentation on PE. In addition, they received general as well as task-specific guidelines which they were instructed to follow with the aim of achieving consistency and in order to avoid interference with the eye-tracker connection. The task-specific guidelines, i.e. the guidelines for the full PE of the NMT output, were based on the comparative overview of full PE guidelines provided by Hu & Cadwell (2016) as these were proposed by TAUS (2016), O'Brien (2010), Flanagan & Paulsen Christensen (2014), Mesa-Lao (2013) and Densmer (2014).

3.3.1 General guidelines

- Your hair should not block your eyes.
- Do not wear mascara.
- Avoid touching your eyes (e.g. rubbing your eyes, removing and wearing eyeglasses, etc.).
- During the PE tasks, look exclusively at the computer screen in front of you.
- Try to keep your head as steady as possible.
- External resources (dictionaries, Internet, etc.) cannot be used

According to O'Brien (2009), the quality of the eye-tracking data may be affected by several factors, such as participants' optical aids, eye make-up, lighting conditions, noise, unfamiliarity, user's distance from the monitor, etc. In an effort to minimise the implications of some of these factors, the participants were given the above-mentioned general guidelines, while a controlled environment for the experiment was set up. In particular, a quiet room was selected, blackout blinds were used to reduce the amount of natural light, the same artificial light was used during all the experiments, and a fixed chair was used, so that the participants could not easily move about and increase or decrease the distance to the monitor (Hvelplund 2011: 103).

3.3.2 Task-specific guidelines

- Retain as much raw MT translation/output as possible.
- The message transferred should be accurate.
- Fix any omissions and/or additions (at the level of sentence, phrase or word).
- Correct mistranslations.

- Correct morphological errors.
- Correct misspellings and typos.
- Fix incorrect punctuation if it interferes with the message.
- Correct wrong terminology.
- Fix inconsistent use of terms.
- Do not introduce stylistic changes.

The training material was sent to the participants five days before the execution of the tasks. In an effort to ensure that they had actually studied the material and that there were no questions or doubts, the participants were interviewed prior to the execution of the tasks and were specifically asked about the training material and also about the guidelines they had received.

3.4 Process and product analysis

As already pointed out, one of the aims of this paper is to compare the technical, temporal and cognitive effort (Krings 2001) expended by the experienced translators for the full PE of NMT output with the effort expended by the novice translators, focusing on the English-Greek language pair. Another aim is to investigate the quality of the post-edited texts and to explore potential differences in the work produced by the experienced and the novice translators.

According to Krings (2001), there are three categories of post-editing effort: (i) the temporal effort, which refers to the time taken to post-edit a sentence to a particular level of quality, (ii) the technical effort, which refers to keystroke and mouse activities such as deletions, insertions, and text re-ordering and (iii) the cognitive effort, which refers to the "type and extent of those cognitive processes that must be activated in order to remedy a given deficiency in a machine translation" (Krings 2001: 179). The cognitive effort is directly related to the temporal effort and the technical effort; however, these do not inform how PE occurs as a process, how it is distinguished from conventional translation, what demands it poses on post-editors, and what kind of acceptance it receives from them (Krings 2001: 61). Furthermore, temporal, technical and cognitive effort do not necessarily correlate, since some errors in the MT output may be easily identified, but may require many edits, while other errors may require a few keystrokes to be corrected, but involve considerable cognitive effort (Krings 2001; Koponen 2012; Koponen et al. 2019).

For our process-based analysis we used eye-tracking and keystroke logging data to measure the temporal, technical and cognitive effort. As far as the temporal effort is concerned, we measured the total time (in minutes) the participants

needed to post-edit the NMT output. It should be noted that the start time of the PE task was calculated from the moment we opened the project (i.e. when we pressed the "start logging" button); the task was considered completed when we pressed the "stop logging" button in the Translog-II. Technical effort is generally measured by the number of keystrokes, i.e. insertions and deletions. Finally, cognitive effort has been generally measured by calculating fixation count, fixation duration and gaze time (Sharmin et al. 2008; Carl et al. 2011; Mesa-Lao 2014; Jia et al. 2019; Doherty et al. 2010; Elming et al. 2014). In view of that, in the present study we measured the average fixation count, the mean fixation duration (in milliseconds), as well as the average total gaze time (in minutes), i.e. the sum of all fixation durations, to compare the cognitive effort expended by experienced translators and novice translators when post-editing the NMT output.

For our product-based analysis and similarly to previous studies (Carl & Buch-Kromann 2010; Moorkens & O'Brien 2015; Koponen et al. 2019; Koponen & Salmi 2017), we used an edit distance metric, i.e. word error rate (WER), to analyse the PE product and each participant's final post-edited text was used as a reference text. WER is based on the Levenshtein[10] distance, and calculates the edits performed to measure the distance between the MT output and its post-edited version. However, these edits do not always reflect actual errors (Koponen et al. 2019). Previous studies (De Almeida 2013; Koponen & Salmi 2017; Koponen et al. 2019) have shown that post-editors either over-edit the MT output making preferential choices or they under-edit it leaving errors uncorrected, while sometimes they also introduce new errors. For that reason, following the calculation of WER, each edit operation was annotated manually by one annotator – a professional translator with 10 years of translation experience – with one of the following categories suggested by Koponen & Salmi (2017) and Koponen et al. (2019):

unedited: no change;
form changed: different morphological form;
word changed: different lemma;
deleted: word removed;
inserted: word added;
order: position of a word changed.

It should be noted that, following Koponen & Salmi (2017) and Koponen et al. (2019), in case a word had been affected by more than one edit type, it was annotated with both categories. For example, some words had both their morphological form and their position changed (form + order) or a different lemma was used and its position was also changed (word + order).

---

[10]http://www.levenshtein.net/

Each word-level edit was then assessed for correctness of meaning (accuracy) and language (fluency) as well as for necessity, i.e. for establishing whether the edit was necessary to correct the meaning and/or the language or whether it was a preferential edit in terms of style or word choice. According to Koponen & Salmi (2017) and Koponen et al. (2019), each edit could be either correct or incorrect and either necessary or unnecessary. In some cases, no edit may be required in the MT output, while in some other cases post-editors may leave errors uncorrected or make preferential changes. Therefore, we decided in our study to cater for such cases by adding additional options. In particular, as far as correctness is concerned, we added the "correct no edit" option, which means that the post-editor was right to leave the MT output unedited given that there was no error in the MT output; we also added the "edit missing" option for cases when an error in the MT was not corrected by the post-editor; finally, we added the "redundant edit" option, which indicates a preferential change made by the post-editor. As far as necessity is concerned, we added the "necessary no edit" option, which means there was no error in the MT output and therefore no edit was required, and the "edit required" option, which means that there was an error in the MT which was not corrected by the post-editor. Summarising the above, the options for correctness and necessity are the following:

Correctness
- correct no edit
- correct edit
- incorrect edit
- edit missing
- redundant edit

Necessity
- necessary no edit
- necessary edit
- unnecessary edit
- edit required

4 Findings and discussion

4.1 Perception analysis: Pre-assignment questionnaire

As it emerges from the participants' answers in the pre-assignment questionnaire in relation to their attitude towards MT and PE, novice translators appear to be more positive towards MT and PE compared to the experienced translators confirming, thus, the findings of previous studies (Moorkens & O'Brien 2015; García 2010; Yamada 2015).

4.1.1 Perception of MT

In the open-ended question regarding their view of MT, only one experienced translator characterised MT as "very positive". Three experienced translators

gave negative responses characterising it as a "necessary evil" (P01), "still very inadequate" (P03) and "not very helpful" (P05). The rest of the experienced respondents were neutral and pointed to the deficiencies of MT or the improvements or conditions that are required for its efficient use by translators (i.e. the use of good quality data, domain specialisation and deep learning). For instance, respondent P06 wrote: "If the machine has been adequately trained, it's OK. In any case, those machines should be continuously provided with new texts in order to be in a position to better 'understand' the structure of each language." Respondents P09 and P10 referred to the MT quality for the English-Greek language pair, noting respectively: "MT is not very developed for Greek yet. It does, however, seem to be highly developed for other languages. Overall, I have not seen substantial results yet concerning the quality of the produced texts, but it is gradually getting better" and "It certainly needs improvement in the language combinations including Greek". Finally, respondent P09 made a very interesting comment about MT and its impact on translators: "I don't think it will 'replace' translators, but I do see it acting as a first stage in the localisation process, in the future, followed by PE and other quality assurance tasks".

The novice translators, on the other hand, found MT "very helpful" (P01), "a vital part of the translation process" (P02), "extremely useful" (P04, P10), "necessary" (P06), and noted that "it helps translators save time and finish a project faster" (P08) and "saves money and time for clients" (P02). Only two novice translators expressed reservations saying: "MT is a tool that can be both useful and useless, depending on the extent it is used and the language knowledge, as algorithms see only words not meanings, and there, a human translator comes to the rescue" (P03) and "I find it quite challenging". Finally, it is interesting that MT is not viewed as a threat by the novice translators, who appear to be more confident as regards the central role played by humans in the translation process. As respondent P06 mentioned, "MT can never replace human translation".

4.1.2 Perception of PE

In the open-ended question regarding their view of PE, the experienced translators appear to adopt a more pragmatic approach, accepting its necessity and its future dominance in the translation market. Six respondents highlighted its importance and/or its necessity saying: "It may be the solution to some problems of the industry" (P02), "It is necessary" (P03, P08) "It is very useful and time-saving" (P04) "it will be widely used in the future so familiarising myself with it will bring added value to my work" (P01) and "It is the future" (P10). Four experienced translators adopted a more reserved stance to refer to PE. Respondent P01 wrote: "Too much effort, too little time", while respondent P09 noted "In my

personal experience, it makes my job harder, since MT is not very developed for Greek yet. However, it is now a hot trend in the localisation field and does seem to be gaining some ground. It could help with some aspects of the field, but does create problems in others." Respondent P06 highlighted its usefulness in certain fields: "Useful only in standard fields, e.g. technical, automotive, maybe medical. Certainly not suitable for marketing, or other creative work" Finally, one experienced translator clearly expressed their preference for translation as opposed to PE saying: "I prefer translating from scratch" (P05).

As far as novice translators are concerned, nine out of the ten underscore the importance and crucial role of PE with the exception of one. They write: "It is important in order to achieve a legible and comprehensible text, whether to be published or not" (P02), "It is really necessary" (P01, P05, P09), "It is of primary importance", "It is important as the first time you see a document, you are not always 100 percent focused and in that way you revise the text for making sure it is correct" (P03), "Post-editing contributes to the speed of the translation and controls the quality of the product" (P04), "Post editing is essential when it comes to delivering a satisfying translation. If this translation is a product of MT and not Human Translation. Post editing can also help improve MT in the way that AI systems 'learn' through their experience" (P07), "It could save time for the translator, as long as it is applied to technical texts" (P08). Finally, respondent P10 acknowledged the fact that it can be useful, albeit time-consuming: "it can be helpful but sometimes its processing requires more time than translation from scratch".

4.2 Process analysis: Measuring PE effort

4.2.1 Temporal effort

As far as the temporal effort is concerned, we measured the average time (in minutes) experienced translators, on the one hand, and novice translators, on the other hand, needed to post-edit the two texts. As it emerges from Table 5.8, experienced translators needed less time ($M = 7.17, \mathrm{SD} = 1.55$) to post-edit the NMT output compared to novice translators ($M = 8.84, \mathrm{SD} = 2.98$). According to a two-tailed two-sample t-test, that difference in average task time between experienced and novice translators is statistically significant $t(29) = -2.22$, $p = 0.02$. Our findings corroborate the findings of Carl & Buch-Kromann (2010), Moorkens & O'Brien (2015) and Schaeffer et al. (2019) who also reported professional translators to be faster than students. They come in contrast, though, with the findings by Nitzke (2019), given that in her study no significant difference was observed between professionals and students in terms of task duration.

Table 5.8: Temporal effort per group of participants: Mean and standard deviation values of the task duration (both texts averaged)

	Task duration (in mins)	
Participants	Mean	SD
Experienced translators	7.17	1.55
Novice translators	8.84	2.98

4.2.2 Technical effort

As it emerges from Table 5.9, the experienced translators performed more keystrokes ($M = 453, \text{SD} = 210$) compared to the novice translators ($M = 318, \text{SD} = 179$). Similarly to the temporal effort, a statistically significant difference $t(37) = 2.18, p = 0.02$ was reported. Nitzke (2019), who investigated also the technical effort, found no significant difference between professionals and students.

Table 5.9: Technical effort per group of participants: mean and standard deviation values for the total number of keystrokes, insertions and deletions (both texts averaged)

	Total keystrokes		Insertions		Deletions	
Participants	Mean	SD	Mean	SD	Mean	SD
Experienced translators	453	210	237	108	216	104
Novice translators	318	179	173	95	145	85

4.2.3 Cognitive effort

As it emerges from Table 5.10, the novice translators triggered more fixations ($M = 1208, \text{SD} = 374$) and longer gaze time ($M = 6.85, \text{SD} = 2.14$) than the experienced translators ($M = 1002, \text{SD} = 153$ and $M = 5.80, \text{SD} = 1.11$). The differences in fixation count and total gaze time were both statistically significant ($t(25) = -2.28, p = 0.02$ and $t(29) = -1.95, p = 0.03$). No statistically significant difference was reported for mean fixation duration $t(37) = 0.51, p = 0.30$. Our findings corroborate the findings of Pavlović & Jensen (2009), who also found students to invest more cognitive effort into their translations than professionals.

Table 5.10: Cognitive effort per group of participants: Mean and standard deviation values of the fixation count, mean fixation duration and gaze time (both texts averaged)

	Fixation count		Mean fixation duration (msec)		Total gaze time (mins)	
Participants	Mean	SD	Mean	SD	Mean	SD
Experienced translators	1002	153	348.62	46.61	5.80	1.11
Novice translators	1208	374	341.63	39.15	6.85	2.14

4.3 Product analysis: Edits' analysis

As it emerges from Table 5.11, the average WER score was lower for the novice ($M = 0.19, \text{SD} = 0.07$) than the experienced translators ($M = 0.28, \text{SD} = 0.10$), confirming the finding of Moorkens & O'Brien (2015) who found students to post-edit less the MT output. It should be noted that the difference in average WER score between novice and experienced translators was found to be statistically significant $t(34) = 3.15, p = 0.002$.

Table 5.11: Average WER for both texts per group

	WER	
Participants	Mean	SD
Experienced translators	0.28	0.10
Novice translators	0.19	0.07

Looking at the edit categories in Table 5.12, we observe that the number of unedited lexical items is higher in the case of the novice translators ($M = 116$, $\text{SD} = 8$) compared to the experienced translators ($M = 106, \text{SD} = 9$), with that difference being statistically significant ($t(37) = -3.37, p = 0.0009$). The novice translators were also found to be more reluctant to change the word order ($M = 4, \text{SD} = 4$) and the word form ($M = 7, \text{SD} = 4$) in the text, as well as to rephrase the text ($M = 8, \text{SD} = 6$) compared to the experienced translators ($M = 6, \text{SD} = 5$ and $M = 11, \text{SD} = 4$ and $M = 13, \text{SD} = 8$ respectively). Unlike the difference in changing the word order ($t(38) = 1.29, p = 0.10$), the differences

in changing a word form and in using a different lemma were statistically significant, ($t(38) = 2.46$, $p = 0.009$ and $t(34) = 2.21$, $p = 0.02$). Our findings seem to be in line with the observation made by Depraetere (2010) that during PE students follow the instructions given and do not rephrase the text if the meaning is clear, but do "not feel the urge to rewrite it" (ibid: 4), potentially leaving errors that should be corrected according to the instructions. Depraetere points out that this indicates a "striking difference in the mindset between translation trainees and professionals" (ibid: 6). In addition, as Yamada (2019) observed, a low error correction rate during PE the NMT output may be due the fact that NMT systems produce human-like errors, which make it more difficult for novice translators to post-edit.

Table 5.12: Average number and percentage of edits for both texts per edit category

	Translators	
	Experienced	Novice
unedited	106 (72.61%)	116 (80.10%)
form changed	11 (7.17%)	7 (5.13%)
word changed	13 (8.43%)	8 (5.23%)
deleted	5 (3.90%)	4 (2.55%)
inserted	6 (3.98%)	6 (4.05%)
order	6 (4.12%)	4 (2.94%)

Evaluating the correctness and necessity of each edit (Table 5.13), we noticed that the experienced translators perform more correct and necessary edits ($M = 26$, $SD = 13$) compared to the novice translators ($M = 17$, $SD = 6$) who tend to leave in their final post-edited text more errors that should be corrected (edits missing and required) ($M = 18$, $SD = 5$), confirming, thus the findings by Depraetere (2010). In addition, the experienced translators seem to have a tendency to overcorrect the NMT output, thus carrying out more redundant edits ($M = 8$, $SD = 5$) compared to novice translations ($M = 5$, $SD = 7$). It should be noted that these differences were statistically significant $t(28) = 3.13$, $p = 0.002$, $t(38) = -3.88$, $p = 0.0002$ and $t(34) = 1.99$, $p = 0.03$ respectively.

Table 5.13: Average number and percentage of correctness and necessity for both texts per edit category

	Translators	
	Experienced	Novice
correct and necessary edit	26 (17.65%)	17 (11.09%)
correct and necessary no edit	95 (64.90%)	100 (68.08%)
incorrect and necessary edit	3 (2.18%)	5 (3.09%)
incorrect and unnecessary edit	3 (1.85%)	4 (2.64%)
edit missing and required	11 (7.72%)	17 (12.05%)
redundant and unnecessary edit	8 (5.69%)	4 (3.05%)

5 Conclusion and future work

The study confirms the findings of the previous studies on the more positive attitude of the novice translators vis-à-vis PE. Experienced translators exhibit a more negative stance vis-à-vis MT and adopt a more pragmatic approach to PE. However, this does not seem to affect the effort expended by the experienced translators when post-editing. In particular, the experienced translators expend less time and less cognitive effort during PE as opposed to the novice translators. On the other hand, the technical effort is found to be decreased in the case of the novice translators. This is due to the fact that they do not sufficiently rephrase the MT raw output as they are reluctant to change the word form, the word order and the syntax of the MT output and are not adequately critical of the content, thus leaving errors in the edited text. Finally, experienced translators have a tendency to overcorrect the NMT output, thus carrying out more redundant edits. These findings have several implications for the training of translators and their continuous professional development as they point to the need for a different approach when designing and delivering courses in PE. Experienced translators' training should aim at helping them appreciate the benefits of MT and PE and develop a more positive stance vis-à-vis MT and PE. In addition, their training should aim at helping them avoid the overcorrection of MT output. This can be achieved through extensive practical exercises, which focus on the identification of errors that require correction depending on the given PE guidelines, i.e. for full or light PE. Novice translators' training, on the other hand, should aim at helping them avoid the undercorrection of the MT output. Such training may include several practical exercises in error detection and correction.

Although the conclusions are clear and they point to several suggestions for the training of translators as outlined above, there are a number of limitations to this study. The main limitation involves the low ecological validity of the study, i.e. the fact that the experiments were carried out in an "artificial", experimental situation rather than in a "natural", real-world situation. In particular, the participants were asked to carry out the tasks at a research institute, the Hubic Lab in Athens, i.e. in an environment that differed from their usual work environment; they could not use any resources during the PE tasks, i.e. they could not use online or offline resources, such as dictionaries, termbases, parallel texts, etc., while they did not work in a translation memory environment. Finally, the error analysis was carried out by only one annotator and the sample size was small and consisted only of female participants. It is our intention in the future to address the present study's limitations and carry out a more extensive research with a higher ecological validity and a more comprehensive and refined analysis of the edits performed by translators in order not only to understand the cognitive mechanisms behind their performance and the differences between the experienced and the novice translators, but also to improve work conditions and performance and thus enhance human-computer interaction.

Acknowledgements

References

Allen, Jeffrey H. 2003. Post-editing. In Harold Sommer (ed.), *Computers and translation: A translator's guide*, 297–317. Amsterdam: John Benjamins. DOI: 10.1075/btl.35.19all.

Bojar, Ondřej, Rajen Chatterjee, Christian Federmann, Yvette Graham, Barry Haddow, Matthias Huck, Antonio Jimeno Yepes, Philipp Koehn, Varvara Logacheva, Christof Monz, Matteo Negri, Aurélie Névéol, Mariana Neves, Martin Popel, Matt Post, Raphael Rubino, Carolina Scarton, Lucia Specia, Marco Turchi, Karin Verspoor & Marcos Zampieri. 2016. Findings of the 2016 conference on machine translation. In *Proceedings of the first conference on machine translation. Volume 2: Shared task papers*, 131–198. Berlin, Germany: Association for Computational Linguistics. DOI: 10.18653/v1/W16-2301.

Carl, Michael. 2012. Translog-II: A program for recording user activity data for empirical reading and writing research. In *Proceedings of the eighth international conference on language resources and evaluation (LREC'12)*, 4108–4112. Istanbul, Turkey: European Language Resources Association (ELRA). http://www.lrec-conf.org/proceedings/lrec2012/pdf/614_Paper.pdf.

Carl, Michael & Matthias Buch-Kromann. 2010. Correlating translation product and translation process data of professional and student translators. In *Proceedings of EAMT*. Saint-Raphael, France. http://www.lanaconsult.com/eamt2010/.

Carl, Michael, Barbara Dragsted, Jakob Elming, Daniel Hardt & Arnt Lykke Jakobsen. 2011. The process of post-editing: A pilot study. *Copenhagen Studies in Language* 41. 131–142.

Carl, Michael, Silke Gutermuth & Silvia Hansen-Schirra. 2015. Post-Editing machine translation: Efficiency, Strategies, and Revision Processes in Professional Translation Settings. In Aline Ferreira & John W. Schwieter (eds.), *Psycholinguistic and Cognitive Inquiries into Translation and Interpreting*, 145–174. Netherlands: John Benjamins Publishing Company.

Carl, Michael, Martin Kay & Kristian Tangsgaard Hvelplund Jensen. 2010. Long distance revisions in drafting and post-editing. In Fremlagt på konferencen CICLing-2010, Iaşi, Romania, March 21–27.

Castilho, Sheila, Joss Moorkens, Federico Gaspari, Iacer Calixto, John Tinsley & Andy Way. 2017. Is neural machine translation the new state of the art? *The Prague Bulletin of Mathematical Linguistics* 108(1). 109–120. DOI: 10.1515/pralin-2017-0013.

Castilho, Sheila, Joss Moorkens, Federico Gaspari, Rico Sennrich, Vilelmini Sosoni, Yota Georgakopoulou, Pintu Lohar, Andy Way, Antonio Valerio, Miceli Barone & Maria Gialama. 2017. A comparative quality evaluation of PBSMT and NMT using professional translators. In *Proceedings of machine translation summit XVI*. Nagoya, Japan.

Colina, Sonia & Claudia V. Angelelli. 2015. T and I pedagogy in dialogue with other disciplines. *Translation and Interpreting Studies* 10(1). 1–7. DOI: 10.1075/tis.10.1.01ang.

De Almeida, Giselle. 2013. *Translating the post-editor: An investigation of post-editing changes and correlations with professional experience across two Romance languages*. Dublin City University. (Doctoral dissertation). http://doras.dcu.ie/17732/1/THESIS_G_de_Almeida.pdf.

Densmer, Lee. 2014. *Light and full MT post-editing explained*. http://info.moravia.com/blog/bid/353532/Light-and-Full-MT-Post-Editing-Explained.

Depraetere, Ilse. 2010. What counts as useful advice in a university post-editing training context? Report on a case study. In *EAMT 2010: Proceedings of the 14th annual conference of the European association for machine translation*. Saint-Raphaël, France.

Doherty, Stephen, Sharon O'Brien & Michael Carl. 2010. Eye tracking as an automatic MT evaluation technique. *Machine Translation* 24. 1–13. DOI: 10.1007/s10590-010-9070-9.

Dragsted, Barbara & Michael Carl. 2013. Towards a classification of translation styles based on eye-tracking and keylogging data. *Journal of Writing Research* 5(1). 133–158.

Elming, Jakob, Laura Winther Balling & Michael Carl. 2014. Investigating user behaviour in post-editing and translation using the CASMACAT workbench. In Sharon O'Brien, Laura Winther Balling, Michael Carl, Michel Simard & Lucia Specia (eds.), *Post-editing of Machine Translation*, 147–169. United Kingdom: Cambridge Scholars Publishing.

Flanagan, Marian & Tina Paulsen Christensen. 2014. Testing post-editing guidelines: How translation trainees interpret them and how to tailor them for translator training purposes. *The Interpreter and Translator Trainer* 8(2). 257–275. DOI: 10.1080/1750399X.2014.936111.

García, Ignacio. 2010. Is machine translation ready yet? *Target* 22. 7–21. DOI: 10.1075/target.22.1.02gar.

Gaspari, Federico, Antonio Toral, Sudip Kumar Naskar, Declan Groves & Andy Way. 2014. Perception vs reality: Measuring machine translation post-editing productivity. In *Proceedings of AMTA workshop on post-editing technology and practice*, 60–72. Vancouver.

Göpferich, Susanne, Gerrit Bayer-Hohenwarter, Friederike Prassl & Stadlober Johanna. 2011. Exploring translation competence acquisition: Criteria of analysis put to the test. In Sharon O'Brien (ed.), *Cognitive Explorations of Translation*, 57–86. London: Continuum Studies in Translation.

Hu, Ke & Patrick Cadwell. 2016. A comparative study of post-editing guidelines. In *Proceedings of the 19th annual conference of the European association for machine translation*, 346–353.

Hvelplund, Kristian Tangsgaard. 2011. *Allocation of cognitive resources in translation: An eye-tracking and key-logging study* (PhD series 10.2011). Denmark: Samfundslitteratur.

Jakobsen, Arnt Lykke. 2002. Translation drafting by professional translators and by translation students. In Gyde Hansen (ed.), *Empirical Translation Studies: Process and product*, 191–204. Copenhagen: Samfundslitteratur.

Jia, Yanfang, Michael Carl & Xiangling Wang. 2019. How does the post-editing of neural machine translation compare with from-scratch translation? A product and process study. *The Journal of Specialised Translation* 31. 60–86.

Kaiser-Cooke, Michèle. 1994. Translatorial expertise: A cross-cultural phenomenon from an interdisciplinary perspective. In Mary Snell-Hornby, Franz Pöchhacker & Klaus Kaindl (eds.), *Translation studies. An interdiscipline*, 135–139. Amsterdam, Philadelphia: John Benjamins Publishing Company.

Koponen, Maarit. 2012. Comparing human perceptions of post-editing effort with post-editing operations. In *Proceedings of the seventh workshop on statistical machine translation* (WMT '12), 181–190. Montréal, Canada: Association for Computational Linguistics. https://www.aclweb.org/anthology/W12-3123.

Koponen, Maarit. 2016. Is machine translation post-editing worth the effort? A survey of research into post-editing and effort. *The Journal of Specialised Translation* 25. 131–148.

Koponen, Maarit & Leena Salmi. 2017. Post-editing quality: Analysing the correctness and necessity of post-editor corrections. *Linguistica Antverpiensia, New Series – Themes in Translation Studies* 16. 137–148.

Koponen, Maarit, Leena Salmi & Markku Nikulin. 2019. A product and process analysis of post-editor corrections on neural, statistical and rule-based machine translation output. *Machine Translation* 33(1–2). 61–90. DOI: 10.1007/s10590-019-09228-7.

Krings, Hans P. 2001. *Repairing texts: Empirical investigations of machine translation post-editing processes*. Vol. 5. Kent, OH: Kent State University Press.

Lommel, Arle & Donald A. DePalma. 2016. *Europe's leading role in machine translation: How Europe is driving the shift to MT*. Tech. rep. Boston. http://cracker-project.eu.

Mesa-Lao, Bartolomé. 2013. *Introduction to post-editing: The CasMaCat GUI*. http://bridge.cbs.dk/projects/seecat/material/hand-out_post-editing_bmesa-lao.pdf.

Mesa-Lao, Bartolomé. 2014. Gaze behaviour on source texts: An exploratory study comparing translation and post-editing. In Sharon O'Brien, Laura Winther Balling, Michael Carl, Michel Simard & Lucia Specia (eds.), *Post-editing of Machine Translation*, 219–245. United Kingdom: Cambridge Scholars Publishing.

Mitchell, Linda G. 2015. *Community post-editing of machine-translated user-generated content*. Dublin City University. (Doctoral dissertation).

Moorkens, Joss & Sharon O'Brien. 2015. Post-editing evaluations: Trade-offs between novice and professional participants. In İlknur Durgar El-Kahlout, Mehmed Özkan, Felipe Sánchez-Martínez, Gema Ramírez-Sánchez, Fred Hol-

lowood & Andy Way (eds.), *Proceedings of the 18th annual conference of the European association for machine translation*, 75–81. Antalya, Turkey.

Moorkens, Joss, Antonio Toral, Sheila Castilho & Andy Way. 2018. Translators' perceptions of literary post-editing using statistical and neural machine translation. *Translation Spaces* 7. 240–262. DOI: 10.1075/ts.18014.moo.

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

Nunes Vieira, Lucas, Elisa Alonso & Lindsay Bywood. 2019. Introduction: Post-editing in practice – process, product and networks. *The Journal of Specialised Translation* 31. 2–13.

O'Brien, Sharon. 2009. Eye tracking in translation process research: Methodological challenges and solutions. In Inger M. Mees, Fabio Alves & Susanne Göpferich (eds.), *Methodology, technology and innovation in translation process research: A tribute to Arnt Lykke Jakobsen*, vol. 38 (Copenhagen studies in language), 251–266. Copenhagen: Samfundslitteratur.

O'Brien, Sharon. 2010. *Introduction to post-editing: Who, what, how and where to next*. http://amta2010.amtaweb.org/AMTA/papers/6-01-ObrienPostEdit.pdf.

O'Brien, Sharon & Michel Simard. 2014. Introduction to special issue on post-editing. *Machine Translation* 28(3–4). 159–164. DOI: 10.1007/s10590-014-9166-8.

O'Brien, Sharon, Laura Winther Balling, Carl Michael, Michel Simard & Lucia Specia. 2014. *Post-editing of machine translation: Processes and applications*. Newcastle upon Tyne: Cambridge Scholars Publishing.

Palumbo, Giuseppe. 2009. *Key terms in translation studies*. London: Continuum.

Pavlović, Nataša & Kristian T. H. Jensen. 2009. Eye tracking translation directionality. In Anthony Pym & Alexander Perekrestenko (eds.), *Translation Research Projects 2*, vol. 2, 93–109. Intercultural Studies Group.

Raptis, Spyros & Maria Giagkou. 2016. From capturing to generating human behavior: Closing the interaction loop at the HUBIC lab. In *Proceedings of the 20th pan-Hellenic conference on informatics (PCI) with international participation*. Partas, Greece: ACM Digital Library, International Conference Proceedings Series. DOI: 10.1145/3003733.3003814.

Schaeffer, Moritz, Jean Nitzke, Anke Tardel, Katharina Oster, Silke Gutermuth & Silvia Hansen-Schirra. 2019. Eye-tracking revision processes of translation students and professional translators. *Perspectives* 27(4). 589–603. DOI: 10.1080/0907676X.2019.1597138.

Schmaltz, Marcia, Igor da Silva, Fabio Alves, Adriana Pagano, Derek Wong, Lidia Chao, Ana Leal, Paulo Quaresma & Caio Garcia. 2015. Translating and post-editing in the Chinese-Portuguese language pair: Insights from an exploratory

study of Key logging and eye tracking. *Translation Spaces* 4. 145–169. DOI: 10. 1075/ts.4.1.07sil.

Sharmin, Selina, Oleg Spakov, Kari-Jouko Räihä & Arnt Lykke Jakobsen. 2008. Where on the screen do translation students look while translating, and for how long? In Susanne Göpferich, Arnt Lykke Jakobsen & Inger M. Mees (eds.), *Looking at Eyes: Eye-tracking studies of reading and translation processing*, vol. 36, 31–51. Samfundslitteratur.

Shreve, Gregory M. 2002. Knowing translation: Cognitive and experiential aspects of translation expertise from the perspective of expertise studies. In Alessandra Riccard (ed.), *Knowing translation: Cognitive and experiential aspects of translation expertise from the perspective of expertise studies*, 150–171. Cambridge University Press.

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

Toral, Antonio & Víctor M. Sánchez-Cartagena. 2017. A multifaceted evaluation of neural versus phrase-based machine translation for 9 language directions. In *Proceedings of the 15th conference of the European chapter of the Association for Computational Linguistics*, vol. Volume 1: Long papers, 1063–1073. Valencia, Spain: Association for Computational Linguistics.

Wagner, Emma. 1985. Post-editing systran: A challenge for commission translators. *Terminologie et Traduction* 3. 1–7.

Whyatt, Bogusława & Tomasz Kościuczuk. 2013. Translation into a non-native language: The double life of the native-speakership axiom. *mTm. A Translation Journal. Special Issue: Translation in an age of austerity* 5. 60–79.

Yamada, Masaru. 2015. Can college students be post-editors? An investigation into employing language learners in machine translation plus post-editing settings. *Machine Translation* 29(1). 49–67. DOI: 10.1007/s10590-014-9167-7.

Yamada, Masaru. 2019. The impact of Google neural machine translation on post-editing by student translators. *The Journal of Specialised Translation* 31. 87–106.