UNIVERSITY OF TARTU
Institute of Computer Science

Conversion Master in IT

Katrin Shuyler

A Case Study on Post-editing Machine Translation: Tasks, Challenges, and Attitudes

Master's Thesis (15 ECTS)

Supervisors:
Alexander Nolte, PhD
Mark Fišel, PhD

Tartu 2021

A Case Study on Post-editing Machine Translation: Tasks, Challenges, and Attitudes

Abstract

While neural machine translation is gaining wider commercial traction as a useful tool for translating technical texts, the professional translator community is hesitant about post-editing machine translation. A qualitative case study guided by the socio-technical idea of the importance to achieve a balance between human and technical aspects of a system was conducted with a focus on post-editing machine translation. The eight translators who took part in this study were given an identical task of post-editing machine-translated technical text. Interviews were then conducted allowing translators to express their perspectives. The interview questions were designed to address how translators approached the study task, what caused them difficulties about post-editing, and their attitudes towards machine translation. The analysis of the data collected for this case showed that while there was variation to how the translators approached the post-editing task, there appeared to be a workflow shared by the majority of participants. Regarding the challenges translators faced in post-editing, the analysis suggests the key factors affecting the translators were decision making and knowledge of the translation tool. This research contributes to the knowledge about translators' attitudes by showing a difference of opinion between professional and personal machine translation use.

Keywords: neural machine translation, post-editing, human-computer interaction, socio-technical systems

CERCS: P170 (Computer science, numerical analysis, systems, control)

Ülesanded, väljakutsed ja hoiakud masintõlke järeltoimetamisel

Lühikokkuvõte

Neuromasintõlge leiab üha laialdasemat kasutust, kuid samal ajal on tõlkijad masintõlke järeltoimetamise suhtes kahtleval seisukohal. Lähtudes sotsiotehnilisest ideest süsteemi inimlike ja tehniliste aspektide tasakaalu kohta, valmis kvalitatiivne juhtumiuuring, mis seadis keskmesse tõlkijate vaatenurga masintõlke järeltoimetamisele. Sellest võttis osa kaheksa tõlkijat, kellele saadeti sama tehnikateksti masintõlke järeltoimetamise ülesanne ja kellega hiljem tehti intervjuud. Intervjuuküsimused koostati sooviga leida vastus sellele, milline oli tõlkijate lähenemisviis järeltoimetamise ülesandele, mis valmistas neile selle juures raskusi ning millised on nende hoiakud masintõlke suhtes. Selle uurimistöö jaoks kogutud andmete analüüs viitas erinevustele tõlkijate lähenemisviisis järeltoimetamisele, kuid siiski läbis suurem osa tõlkijaid peamised töösammud sarnaselt. Analüüsi põhjal nähtus, et järeltoimetamisel kohatud raskused võisid olla seotud otsuste langetamise oskusega ja teadmistega tõlketarkvarast. Käesolev uurimistöö annab oma panuse teadmisse tõlkijate hoiakute kohta, näidates erinevusi tõlkijate suhtumises masintõlke kutsealasesse ja isiklikku kasutusse.

Võtmesõnad: neuromasintõlge, järeltoimetamine, inimese ja arvuti interaktsioon, sotsiotehniline süsteem

CERCS: P170 (Arvutiteadus, arvutusmeetodid, süsteemid, juhtimine (automaatjuhtimisteooria))

Table of Contents

Abbreviations and Terminology

CAT - computer-assisted translation
GT - Google Translate
FM - fuzzy match
LSP - language service provider
MT - machine translation
NMT - neural machine translation
NM - no match
PE - post-editing
PEMT - post-editing machine translation
QA - quality analysis
RBMT - rule-based machine translation
SMT - statistical machine translation
TM - translation memory
WPS - words per second

1 Introduction

As translation technology rapidly develops, it is being widely used in the domain of technical translation to help translate user manuals, product descriptions, software texts, and other consumer texts. Technical translation projects have four traits in common: (1) high rate of repetition, (2) high volume, (3) urgency, and (4) a need for terminological consistency. The first three traits, as listed by O'Brien (2012), address quantity[1], while (4) is a matter of quality. It is widely acknowledged that present-day translation is an interaction between humans and machines, because "/.../ all translators use technology in their translating process in one way or another" (Sakamoto 2019: 210). Without using computer-assisted translation tools (CAT-tools) to meet the needs presented by the four above-mentioned aspects, it is not realistic for a language service provider (LSP) to survive, let alone successfully compete in the rapidly growing[2] translation market. Thus, in order to succeed as a technical translator, in addition to the necessary domain and linguistic knowledge, one needs to have a good command of CAT-tools. Some translation aids, such as translation memory systems are now commonly used by the translation community, whereas machine translation (MT) still has not convinced the majority of professionals (Carl, 2011; Sakamoto, 2019).

The history of MT can be traced back several decades, but with the rise of neural machine translation (NMT) as the new state-of-the-art MT in recent years (Scansani, Bernardini, Ferraresi, and Bentivogli, 2019), MT seems to have entered a new era (Pym, 2019) of finally being ready for widespread commercial use. Several studies show how NMT outperforms the statistical machine translation (SMT), which was the previously predominant MT technology (Castilho et al., 2017; Shterionov et al., 2018). Therefore, more and more LSPs are taking steps to implement MT in their translation workflow. Among other changes in the workflow that come with this change, is that a new task is being introduced to translation professionals: the task of post-editing machine translation (PEMT).

1.1 Problem Statement

When it comes to translating technical texts, it is generally accepted that MT helps translators to increase their productivity compared to traditional translation methods (Plitt and Masselot, 2010). In addition to the increase in productivity, Cadwell, O'Brien, and Teixeira (2018) refer to several studies that prove MT is more helpful than harmful on the final quality of the translated text. Thus, an Estonian translation agency called Technica Translations is preparing to begin using a new translation tool in the domain of the automobile industry, enabling the employment of a new technology: neural machine translation (NMT). This NMT technology is integrated into a CAT program called memoQ as a custom-built plugin. As a result of

[1] The underlying premise here is not to compromise on quality, despite increasing time pressure and number of pages to translate. In this sense, (1)-(3) are related to quality as well.
[2] According to statista.com Global language services: market size 2019 language services market has more than doubled in 2019 (at 49.6 billion US dollars) compared to 2009 (23.5 billion US dollars) and is projected to continue rapid growth.

implementing MT into their workflow, Technica Translations hopes to be able to provide faster translations in higher volumes while simultaneously not suffering any loss in quality.

The reason for using a custom-developed NMT, and not an off-the-shelf plugin is threefold: (1) since Technica Translations' solution will be used domain-specifically, the training corpora will be selected accordingly. Staying within the confines of a single, specific domain and supplying the NMT engine with ample training material will give better translation results than using a general-purpose NMT alternative (Koehn and Knowles, 2017); (2) NMT solutions by third parties (Google MT, Microsoft Translator) generally are not compliant with confidentiality requirements set forth by the LSP's client agreement; and (3) customizable MT solutions by third parties that would comply with confidentiality requirements are costly, making it more feasible in the long-term to implement one's own solution with the added possibility of scaling it to more projects.

In general, the studies on MT are often centered around technical aspects (Zaretskaya, 2017), meaning their focus is typically on evaluating the outcome of MT in terms of fluency, accuracy, intelligibility, or other criteria. Studies paying particular interest to translators' preferences and mindset are harder to come by. As Zaretskaya points out: "/.../ the user perspective of MT has not been explored sufficiently, in particular how MT is used as a translation aid and how it can be improved from the point of view of the user" (Zaretskaya, 2017: 53). Zaretskaya's suggestion about the necessity of research on MT users is in line with the socio-technical value that employees' rights and needs should receive the same amount of attention as technological aspects in the company (Mumford, 2000). Given that MT continues to require human input for the end product to be of adequate quality, there is little hope for success in benefiting from this technology without translation professionals who are willing to and skilled enough to post-edit MT output. However, the professional translator community is not welcoming MT (Sakamoto, 2019). This calls for gaining more knowledge about how translators approach PEMT and their thoughts and feelings about it. The problem situation this thesis addresses is presented in Figure 1.

Figure 1. The Problem Situation of the Current Thesis.

The aim of this study is to explore how translators PEMT, which challenges they meet, and to get an idea about their attitudes towards MT. The main data source in this study is interviews with translators who had all previously completed the same PE task assigned to them as a part of this research. This study attempts to answer the following three research questions:

RQ1. How do translators perform the post-editing task?
RQ1.1. What are the subtasks and in which sequence do they occur?
RQ1.2. How does post-editing compare to translation and editing?
RQ2. What are the challenges translators encounter when post-editing machine translation?
RQ3. What are translators' attitudes towards machine translation?

1.2 Structure of This Thesis

Chapter 1 introduces the topic of this thesis, giving an overview of the problem this thesis is addressing, and the respective research questions are formulated.
In chapter 2, the field of research and theoretical framework are presented with a literature review.
In chapter 3, the research design and methodology are explained.
Chapter 4 presents the results of the thematic analysis of the interviews.
Chapter 5 contains the results from supporting data sources.
In chapter 6, the research findings are discussed and the important results are brought out. This chapter also contains implications for future research and practice that are motivated by the results of this study. Additionally, limitations of this study are addressed.
In chapter 7, conclusions about the important results are made.

2 Background and Related Studies

This chapter maps the field of related research and the key concepts it draws on. After presenting a brief historic overview of MT development (2.1), there will be an overview of studies looking into post-editor's tasks and skills (2.2), translation and PE in the theoretical framework of socio-technical systems (2.3), and studies addressing the measurement of translators' professional experience and PE effort (2.4).

It is important to stress that NMT is relatively new and has only recently been established as a state-of-the-art method[3]. Considering this novelty, some of the literature that is referred to focuses on the previously-predominant SMT. This adds to the necessity for conducting more research on NMT technology.

2.1 Brief History of Machine Translation and Its Three Main Paradigms

Hutchins defines machine translation as follows: "The term 'machine translation' (MT) refers to computerized systems responsible for the production of translations with or without human assistance" (Hutchins, 1995: 431). Whereas the use of mechanical dictionaries was first proposed in the 17th century, the first MT system was developed in the 1950s in the United States of America and it translated some preselected sentences from Russian into English — a choice guided by the wish to attract funding during the Cold War era (Hutchins, 1995). In the early days of MT no less than fully-automatic translation was expected from MT solutions. Since the aim was ambitious, the concerned parties were disappointed with the MT results and the MT program was mostly halted in 1966 due to the conclusions drawn in the infamous ALPAC report (O'Brien, 2012). The 1980s witnessed a new rise in MT which is continuing to this day: new MT systems have been developed and MT research has been expanded.

Based on their working principle, MT systems can be divided into three paradigms: rule-based MT (RBMT), statistical MT (SMT), and neural MT (NMT). RBMT relies on a set of linguistic rules and resources, such as syntactic parsers, morphological generators, dictionaries, etc. (Sreelekha, 2017). SMT is grounded in text corpora, aligning two languages based on a statistical model. NMT makes use of an artificial neural network to predict the likelihood of a sequence of words (DeepAI, 2019).

The futuristic ideal of the 1960s was to achieve fully automated translation, but nowadays the expectations are somewhat lower and in most cases MT output needs to be processed by a human in order to achieve the quality desired. Post-editing machine translation (PEMT) refers to the revision of machine translated content by a human. It can be either full PE to comply with higher quality requirements or light PE for texts which do not need to be perfect. If the editor only sees MT output, then they are dealing with monolingual PE. If the source text is present as well, then it is called bilingual PE.

[3] Google Translate becoming a NMT engine in November 2016 could serve as a milestone here.

2.2 Tasks and Skills in Post-editing

The literature reviewed in this section was helpful in asking and answering *RQ1: How do translators perform the post-editing task?; RQ1.1: What are the subtasks and in which sequence they occur?*; and *RQ1.2: What is specific to post-editing compared to translation and editing?*

Although the role of a post-editor can be considered to be a relatively new one in the translation industry, the concept is not new. In his writing nearly seven decades ago, Yngve (1954) mapped what factors into being a post-editor in *The machine and the man* as follows:

> "A post-editor is a person skilled in the output language but who may be entirely ignorant of the input language. His task is to take the imperfect output from the machine and edit it into a polished or at least easily comprehensible document. This puts man in the role of partner with the machine. /.../ Although man has been reduced to a link in the chain, he does not have to solve the large number of routine problems, but can concentrate on the real difficulties. It has been shown that the post-editor is better able to do his job if he also knows the input language; thus we have the bilingual posteditor. It has also been shown that the post-editor is better able to do his job if he is an expert in the particular field of knowledge. /.../ It seems obvious that the amount of work done by a post-editor depends upon the ultimate purpose for which the translation is being made. If the purpose is to provide a translation in a literary style that could be published in a journal, possibly with large circulation, the post-editor might have a big job. If the purpose is to provide a rough copy that can be used by experts to determine whether or not the material is of interest to them, the post-editor would have a smaller job, or might not be needed at all. His utility depends upon how perfect a translation the machine makes and how perfect a translation is desired" (Yngve, 1954: 21).

The themes mentioned by Yngve can all be found in present-day literature about PE: proficiency in both source and target languages, PE up to human quality, human-computer interaction ("partnering with the machine"), automating human work, reducing repetitive tasks and letting human translators concentrate on real difficulties in translation, domain knowledge, choice between light or full PE based on the purpose of the text, and MT quality. Why does PE continue to be referred to as something new and still worth exploring if it has been around for so long? Although discussed already in the mid 20th century, the tasks and skills related to PE should be explored in connection to the increasing use of NMT while considering the socio-technical aspects of the 21st-century translation industry.

In an attempt to find out what makes a translator a good post-editor, Rico Pérez and Torrejón (2012) explicitly position themselves in the tradition of authors who find PE to be a task for translators. They start with mapping the tasks and processes of a post-editor based on Krings' research (Krings, 2001) as tasks and processes related to the source text, MT, target text production, target text evaluation, reference work, physical writing, and PE task management

(Rico Pérez and Torrejón, 2012: 168-169). Rico Pérez and Torrejón's work relies mainly on the analysis of related literature, their insights, and a set of MT errors. Based on identifying the PE tasks and processes, they then gather PE competences and skills into three groups: (1) "Core competences" (addresses competences of attitude and strategy), (2) "Linguistic skills" (addresses competences of communication and text, cultural and intercultural, and domain), and (3) "Instrumental competence" (addresses competences of knowledge on MT, term management, maintaining MT dictionary, and basic skills in programming) (Rico Pérez and Torrejón, 2012: 169). It is notable that Rico Pérez and Torrejón see a need to have basic programming skills as an instrumental competence of a post-editor. Good command of CAT-tools and MT knowledge is usually mentioned by other authors as well (Nitzke, Hansen-Schirra, and Canfora, 2019), however, programming skills are beyond the usual required skill set. According to Rico Pérez and Torrejón, these skills are needed "for creating macros or scripts for automated correction" (Rico Pérez and Torrejón, 2012: 170). They relate good instrumental competence and its components to a positive attitude in interacting with the machine. That is to say, knowledge about MT and being technically skillful helps with the core competence of attitude. In the context of the present study: acquiring necessary skills (RQ1) to overcome challenges in PE (RQ2) helps to improve attitude (RQ3).

Ginovart Cid (2020) takes a look at the professional profile of a post-editor based on a survey conducted among LSPs and linguists. Ginovart Cid's aim was to examine the competency gap between professionals in the field and new MT-related tasks required by the industry. In order to help linguists compete in the job market, Ginovart Cid sought to better define the newer and more technological roles in the language industry and the respective skills needed. The top three skills in PE were the capacity to post-edit up to human quality (full PE), the ability to PE according to PE guidelines, and the potential to identify MT errors (Ginovart Cid, 2020: 178). Among the skills recognized as less important were applying the right correction strategy, reaching a good enough PE quality (light PE), deciding whether to edit or translate from scratch (Ginovart Cid, 2020: 178). Three skills that LSPs seek in a post-editor candidate are revision and proofreading skills, subject field knowledge or specialization, and CAT tool knowledge (Ginovart Cid, 2020: 179). The skills of the post-editor are addressed by RQ1.2 in this study.

Nitzke et al. (2019) examined which competences are required from a post-editor in addition to the ones required from a translator (addressed by RQ1.2 in this study). They addressed translation risk management in connection to the PE task and proposed a PE competence model. Related to risk management, Nitzke et al. (2019) discussed aspects of post-editor's work, such as awareness of information security and being able to decide between light or full PE.

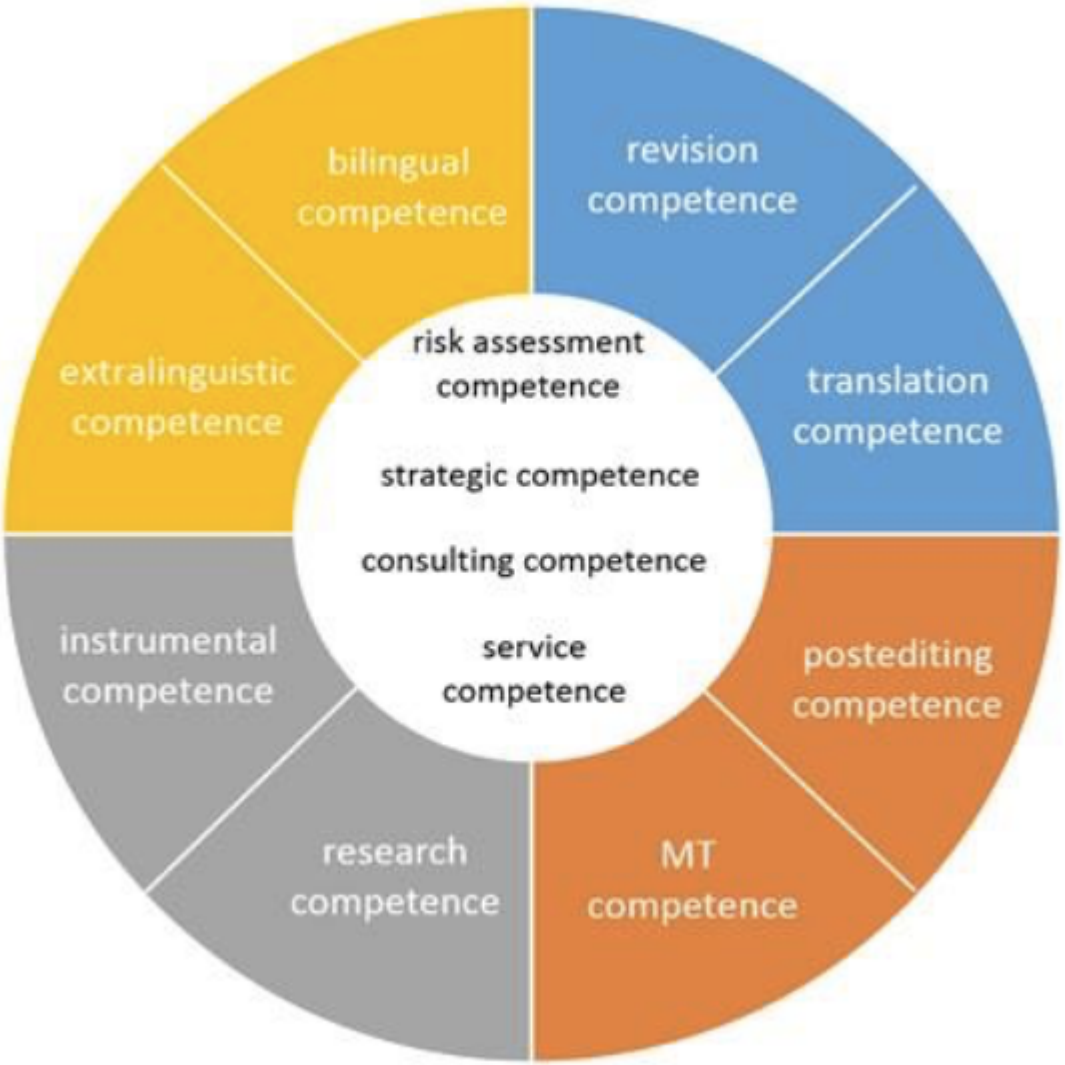

Figure 3. PE competence model with 4 core competences and 8 sub-competences from the study of Nitzke et al. (2019: 250).

Their (Nitzke et al., 2019) competence model is based on linguists' ability to assess risks and know about their trade. Post-editors should have both linguistic and cultural knowledge, be able to translate and revise, be acquainted with PE processes and know the working mechanisms of MT in order to spot NMT-specific errors, be able to use CAT-tools, and engage in independent research on the subject matter (Nitzke et al., 2019: 249-250). The scope of these competences is wide, ranging from being able to assess PE-related risks and negotiate with the client to engaging in research and knowing specifics about NMT technology. Because PE is a demanding task, the respective difficulties and prerequisites for translators to transition into post-editors need a more thorough exploration, which is the aim of the present research. Exercising a wide perspective on the post-editor's role, Nitzke et al. (2019) point out that success in PE is also affected by the following factors: "psycho-physiological components, the post-editor's self perception, the PE brief including guidelines for the task, an affinity towards technology and computer" (Nitzke et al., 2019: 251). To address this demanding task, a holistic approach, such as the socio-technical theoretical framework discussed in the next section, becomes useful.

As a point of reference, the European Masters of Translation rely on the translation competence model which brings together linguistic, domain-specific, technological, research, intercultural, and translation workflow related expertise (Gambier, 2009) as seen in Figure 4 below.

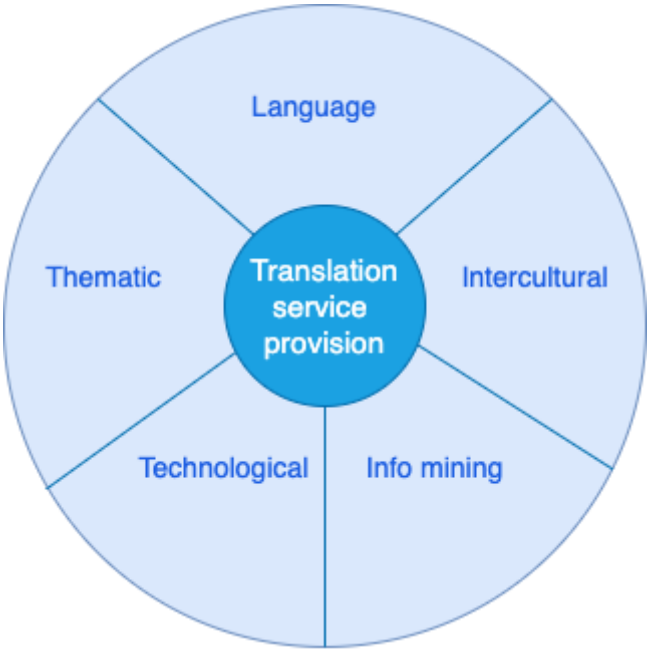

Figure 4. The EMT model of translation competence (Gambier, 2009: 4).

There are overlaps in translation competences and PE competences (addressed by RQ1.2 of this study) in the models presented in Figures 3 and 4. Namely, the post-editor seems to need all the competences of a translator in addition to competences related to (1) risk assessment (when to choose light or full MT, or when one cannot choose to use MT at all, price negotiation based on PE strategy), (2) knowledge about MT technology, and (3) MT tools. Based on these additions it becomes clear the post-editor's tasks are more intertwined with technology, involve more risks, and include more decision making.

2.3 Socio-Technical Perspective on Challenges and Attitudes in Post-editing and Translation

The literature reviewed in this section was helpful in asking and answering *RQ2: What are the challenges translators encounter when post-editing machine translation?* and *RQ3: What are translators' attitudes towards machine translation?*

Due to widely used CAT-tools (such as translation memories (TMs) and term bases (TBs)) and the use of MT, translation is an increasingly reciprocal action between the translator and the machine. As such, it can be viewed as a human-computer interaction (HCI) and more widely in the context of socio-technical systems. SpringerLink Encyclopedia of Database

Systems defines human-computer interaction as "/.../ the study of the way in which computer technology influences human work and activities" (Dix, 2009). O'Brien (2012) offers an even more specific term for describing the relationship between the translator and the machine: translator-computer interaction (TCI).

The socio-technical concept stems from post-WW2 research projects in the British coal mining industry that addressed work processes and organizational arrangements due to miners leaving the industry for better work opportunities (Trist, 1981). Analogously, some authors of translation studies (O'Brien, 2012) have voiced their fear regarding MT scaring professionals away from their trade. Socio-technical research tries to figure out the best equilibrium between human and technical aspects of a system, and views these two aspects to closely affect each other (Trist, 1981). While MT in general is not a recent technology, it has gained wider commercial traction in companies thanks to the advances of NMT. Adopting NMT brings change in tasks, people, and possibly even structure of the organization (Leavitt, 1965). As pointed out by Karamanis, Luz, and Doherty: "Re-educating translators to work with MT implies organizational change" (Karamanis et al., 2011: 51). This means that for successful use of MT it is pertinent to consider factors other than only the aspects arising directly from translators interacting with MT tools. The following model in Figure 5 places the translator's interaction with the computer into the wider theoretical framework of the socio-technical system.

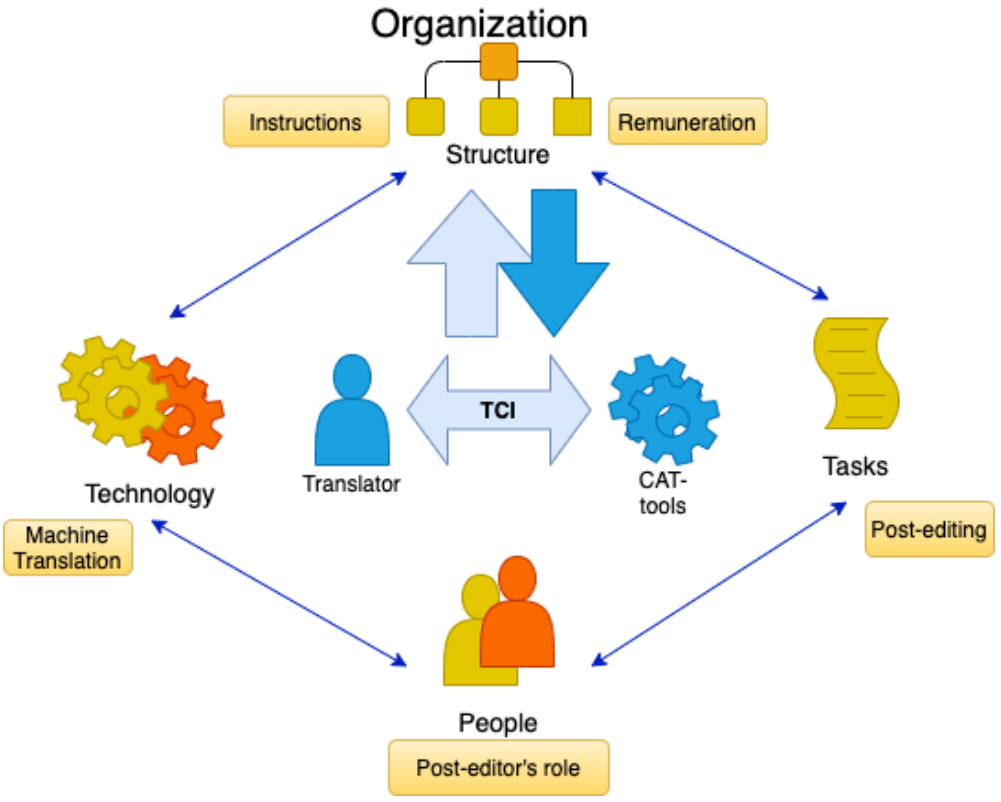

Figure 5. Translator-computer interaction placed in Leavitt's diamond model of organizational change (Leavitt, 1965).

In O'Brien's 2012 article, in which she famously characterized contemporary translators' work as translator-computer interaction, she took a look at both advances and problems that arise in such interactions. She pointed out a relevant factor that translators often view MT as a "black box", because they do not know much about the underlying mechanisms that make it work (O'Brien, 2012). She claimed this in 2012 when the field was dominated by SMT, but perceiving MT as a black box could be even more true now, in the age of NMT, since the latter is more complex to understand by wide audiences. With no thorough understanding of the technology and thus being left without opportunities to meaningfully affect the processes of MT, the translator might have less trust in MT. She suggests that in order to overcome such mistrust and dehumanizing effects of MT, translators should be closely involved in the processes connected to MT, starting from the testing process (O'Brien, 2012).

Cadwell et al. (2018) sought to find out why professional translators resist MT. They also suggest the proliferation of NMT to be inevitable and thus, the translators will have to surrender to this tendency. Their research was conducted using focus groups from two different institutions. Translators were presented with different reasons for using or not using MT. In both institutions, translators thought of speed or productivity gains as the main reason for using MT. It is striking that hundred percent (100%) of translators in one institution found wider MT use to be inevitable, whereas the corresponding percentage in the other institution was zero (Cadwell et al., 2018: 310). The three main reasons for not using MT that translators in both institutions mentioned were: (1) MT is ineffective for some text types; (2) MT quality is not satisfactory for some language pairs; (3) MT has a negative effect on translator's abilities (Cadwell et al., 2018: 311). Other common concerns were listed by translators, such as fear, cognitive load, translator's shift of roles into a more inferior one, distrust, and loss of creativity (Cadwell et al., 2018: 311). According to the authors' findings, translators in one institution were more likely to use MT and they suppose this is because of how the socio-technical deployment of MT had been better in the respective institution (Cadwell et al., 2018). This gives wider empirical support to what O'Brien (2012) argued about translators' involvement based on her experience. In connection to using MT, Cadwell et al. (2018) point both towards challenges in using MT (addressed by RQ2 of this research) and attitudes about MT (RQ3).

Ehrensberger-Dow and Massey (2017) viewed different client-related, colleague-related, resource-related, and tool-related aspects of the profession of translator in the theoretical framework of socio-technical systems. Their conclusion, that translators' reluctance to adopt new technology might be caused by not feeling involved in the decision-making processes, is in line with the conclusion of Cadwell et al. (2018) about better socio-technical deployment resulting in more motivation to use MT. Based on in-depth interviews, Ehresberger-Dow and Massey concluded that translators' "/.../ perceived self-determination is more important to the success of socio-technical change than the technological developments" (Ehrensberger-Dow and Massey, 2017: 104). The present study also explores translators' perspectives and seeks to map different aspects that factor into the socio-technical changes that occur at the onset of the use of MT.

In their article about CAT-tool features that irritate translators, O'Brien, Ehrensberger-Dow, Hasler, and Connolly (2017) point out that although CAT-tools have been around and widely used for around 20 years, many translators still treat them with "suspicion or disinterest" (O'Brien et al., 2017: 145). The authors (O'Brien et al., 2017) conducted a large-scale survey among professional translators to elucidate which tools are frustrating to use or simply unavailable. Based on both quantitative and qualitative analyses of their findings, the authors concluded that translators find the complexity of the user interface and segmentation to be most problematic (O'Brien et al., 2017). The authors address cognitive friction (could be seen as an interruption in the seamless workflow) and cognitive load (commonly assessed in PEMT research) in translating with CAT-tools (O'Brien et al., 2017). The current research examines the cognitive issues of PEMT through asking about PE effort and problems in the interviews, in addition to the analysis of the PE editing time and PE edit distance data (RQ2).

Olohan (2011) examined how the technological and social aspects are intertwined in translation using translation memories. In doing so, she relied on Andrew Pickering's notions of "mangle of practice" as a "/.../ dialectic of resistance and accommodation, which is brought to bear on scientific and technological advances /.../" and the "dance of agency"[4] (Olohan, 2011: 344). This dance of agency means that humans are trying to control the non-human, such as technological and natural factors, and non-human elements move towards gaining control over humans (Olohan, 2011). To find out more about what she calls interaction between translator and TM software, Olohan qualitatively analyzed forum posts where translators had written about their problems and experiences with SDL Trados software. Based on her findings, she points out that different translators can have contrasting perceptions of translation software and the different ways of interacting with it (Olohan, 2011). She compares two translators' perceptions of CAT-software issues: Translator 1 attributed the blame for problems to the software, whereas Translator 2 viewed the problems more as human shortcomings (Olohan, 2011). The current study qualitatively explores the difficulties translators face when using MT (RQ2). Inspired by Olohan's (2011) approach, the author of this study read translators' forum posts about MT, but only for gaining better background knowledge on translators' opinions about MT from an additional source.

Bundgaard, Christensen, and Schjoldager (2016) addressed translator-computer interaction in the process of using a TM suite with integrated MT technology. Their (Bundgaard et al., 2016) empirical research follows Olohan's theoretical framework (Olohan, 2011) and is conducted with a life-like setting among eight translators in a Danish translation bureau as MT had begun to be implemented in their usual TM tool (SDL Studio). Bundgaard et al. (2016) used several data collecting methods ranging from screen recording to a post-experimental questionnaire. The authors found CAT-tools to be of major importance in the work of a translator as well as how CAT-tools have both a helping and inhibiting effect on the work. One of the inhibiting factors mentioned was the translator's feeling of entrapment by the suggestions of MT (Bundgaard et al., 2016). This can be interpreted as losing one's agency to the machine by "accommodating resistance" (Bundgaard et al., 2016: 114),

[4] In this context, "agency" is not limited to intentional agency (Olohan, 2011: 345).

highlighting once again, the need for better socio-technical deployment of MT. The present study also takes the problem of losing agency into consideration (RQ2).

Another study following the footsteps of Olohan (Olohan, 2011) was conducted by Ruokonen and Koskinen (2017) who focused on translators' emotions brought out by using CAT-tools. In order to capture how translators feel about machine-aided translation (addressed in RQ3 of the present research), they used an uncommon research method of letting translators "/.../ write a short 'love letter/break-up letter' to a tool, application or aspect of work" (Ruokonen and Koskinen, 2017: 310). Seventy percent (70%) of these letters referred to technology, showing it is an essential part in the work of a translator. In these letters, narratives of resistance were dominating but not without hope for a better future (Ruokonen and Koskinen, 2017).

Scansani et al. (2019) explored how much translators trust MT. They divided translation students into two groups, both of which received the same text, but participants in one group were told it was a human-translated text that needed editing, while participants in the other group were informed it was a machine-translated text which needed PE. Scansani et al. (2019) viewed trust "/.../ as strictly related to productivity: when post-editors/revisers do not trust a text, they are likely to carry out time-consuming and potentially unnecessary searches, or perform unnecessary edits" (Scansani et al., 2019: 73). Attitude towards MT (RQ3) in the present research examines the issue of trust, mediated by exploring translators' assessments of the quality of MT, willingness to use MT in their work, and motivation to seek more information about MT. According to the findings of Scansani et al. (2019), neither translation edit rate nor editing speed demonstrated a difference between the two groups, suggesting human translation and MT were trusted equally.

Guerberof (2013) asked 24 translators about their opinions on PE and MT via an online questionnaire and a later debrief. Among others, topics like PE pricing, PEMT in comparison to translation and editing, trust, knowledge about MT, PE instructions, and MT quality were brought up in the debriefings (Guerberof, 2013). This is similar to the present study's attempt to answer RQ3. Guerberof points out that studying translators' opinions on MT is rare: "It is not very often that translators are asked their opinions about PE and MT in the localisation industry. This could be out of fear of an adverse or negative response or due to the fact that translators are often invisible in the localisation work-flow, and we feel this invisibility is increasing as process automation increases, and all aspects related to technology seem to acquire more relevance than the act of translating itself" (Guerberof, 2013: 75). It is important to give a voice to translators to balance the technology-centered approach to MT.

2.4 Measuring Post-editing Effort and Professional Experience

The literature reviewed in this section helped to answer *RQ2: What are the challenges translators encounter when post-editing machine translation?*

According to Krings's theory (Krings, 2001), PE effort consists of three types of effort: (1) temporal, (2) cognitive, and (3) technical. When measured, (1) reflects the time used for PE, (3) can be represented via the number or share of edits/changes made during PE (edit distance), while (2) poses difficulties in the means of direct measurements. Temporal effort can be viewed as the sum of cognitive and technical effort, although there are overlaps between the two (Tatsumi, 2010). In a generalized way, PE effort could be explained in the following way (refer to Figure 6): the post-editor thinks and decides (cognitive effort), makes necessary changes into the target text (technical effort), and the time the post-editor takes for their cognitive and technical processes results in temporal effort (Tatsumi, 2010).

It does not seem to be achievable to divide the technical and cognitive types of effort into two discrete units, because they are intertwined and as a result there are "overlaps" (Tatsumi, 2010: 40). However, as temporal and technical efforts are more straightforward to measure, they can serve as a basis for assumptions about cognitive effort to support qualitative information on PE challenges collected via interviews (RQ2).

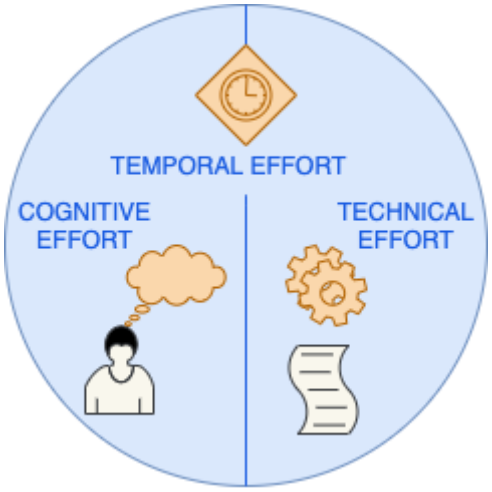

Figure 6. The constitution of PE effort based on Krings (2001) and Tatsumi (2010).

De Almeida (2013) explored the influence of translation experience on PE speed and quality based on a study conducted with 20 participants who remotely completed a PE task. De Almeida summarizes the profile of the best (in terms of the combination of the PE speed and quality) post-editors: "The participants who presented the best PE performance had both previous translation and PE experience and performed few or no searches on the Web (presumably as a result of their previous professional experience)" (De Almeida, 2013: 200-201). She measured both translation and PE experience in years. The present research borrows from her research by measuring translation and editing experience in years and comparing this to total PE time (De Almeida, 2013). However, considering the nature of the data collected, PE experience in the present study will only be measured as existing and non-existing.

3 Case Study

This chapter discusses the design and methodology used for this study. The study is first placed into a wider theoretical framework and then its specific methodological aspects are explained.

3.1 General Research Design

Research design is needed for planning and structuring the research in a way that allows the research questions to be suitably answered. This study attempts to answer the following research questions:
RQ1: How do translators perform the post-editing task?
RQ1.1 What are the subtasks and in which sequence they occur?
RQ1.2 How do translators compare post-editing to translation and editing?
RQ2: What are the challenges translators encounter when post-editing machine translation?
RQ3: What are translators' attitudes towards machine translation?

Yin (2003) suggests using a case study design "/.../ when "how" or "why" questions are being posed, when the investigator has little control over the events, and when the focus is on a contemporary phenomenon within some real-life context" (Yin, 2003: 1). This research corresponds to these criteria. It seeks to gain knowledge about translators' reception of the new task of PE in a small-sized translation agency (Technica Translations). As such, this research is of an exploratory nature and its questions are broad. The unit of the analysis is translators that took part in this study. Qualitative data and methods were used in this study with some supportive addition of quantitative data.

3.2 Participants

According to Lazar, Feng, and Hochheiser (2017), users who participate in a HCI study have to be "/.../ representative in terms of age, educational experience, and technical experience, but also in terms of the task domain /.../" (Lazar et al., 2017: 6). The translators who took part in the study were of different ages; some male and some female; all had higher education; possessed translation experience between 6 and 25 years, varying rate of task domain experience, either some previous experience in PEMT or none, and they had varying experience with the CAT-tool memoQ which used in this study. The translators were remunerated for their effort.

Table 1. Information about translators' relevant experience

Trans-lator	**Translation experience**	**Editing experience**	**Domain experience (car industry share of total work estimate)**	**Main domain of expertise**	**Professional PEMT experience**	**memoQ experience**
T1	6 yrs	14 yrs, often	Yes, 30%	Technical user manuals	5 projects	6 yrs
T2	17 yrs	rare	Yes, 40%	Technical user manuals	none	Has not used often
T3	14 yrs	rare	Yes, less than 5%	Legal texts	none	Has not used often
T4	25 yrs	rare	Yes, less than 5%	Technical user manuals	none	4 yrs
T5	17 yrs	rare	Yes, 5-10%	Technical user manuals	none	5 yrs
T6	16 yrs	rare	Yes, 5-7%	Technical user manuals	Less than 5 projects	10 yrs
T7	21 yrs	10 yrs, often	Yes, 10%	Technical user manuals	More than 5 projects	17 yrs
T8	6 yrs	no	Yes, 30%	Marketing materials	none	6 yrs

The relatively small number of German to Estonian technical translators was the main limitation of finding suitable participants for this study. However, the chosen unit size of eight seems to be in line with several previous studies in the field, where the number of participating translators has been around ten (O'Brien, 2006; Guerberof Arenas, 2008; Plitt and Masselot, 2010; Läubli, Fishel, Masseu, Ehrensberger-Dow, and Volk, 2013). Information about participants' relevant experience can be seen above in Table 1.

3.3 Study Setting

The study was conducted remotely to allow more flexibility for the participants to choose a suitable time slot for the PE task and interview, and because of limitations caused by the outbreak of COVID-19. Technica Translations provided a server for setting up the PE task in the desktop version of memoQ (refer to Figure 7). All participants used their own equipment (computers, monitors, keyboards, mouses) which contributed to realistic results.

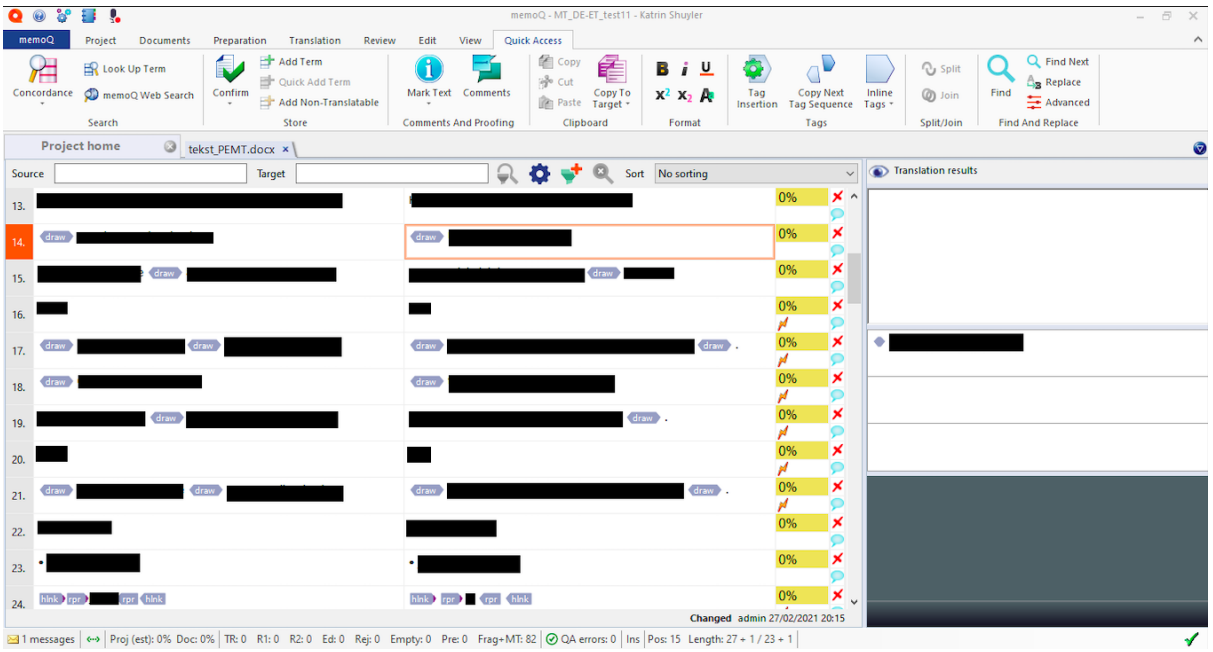

Figure 7. An extract from the study task in memoQ with redacted text.

The source text of the study task was in German, from the domain of the automobile industry, and contained material about features and functions of cars. Sentences of different lengths were chosen and were not edited before MT. It is not known if controlled language rules were used in composing the text by the manufacturer. Since the same text was given to all translators and the unit of this study are translators, it was not necessary to focus on specifically determining the characteristics of the source text.

The instructions accompanying the study task were kept minimal due to the exploratory nature of this research — to seek information about how translators approach the task. The instructions were to post-edit up to human quality, to post-edit the task all at once, to confirm all segments, to submit the results of the time measurement function in memoQ, to generate the editing time report after finishing the task, and additionally mark down the time spent on the task when starting it and ending using a timepiece.

3.4 Data Sources

Data sources used in research should be chosen to answer the research questions and give suitable data. Interviews with translators (for answering all RQs) and the post-edited file in

memoQ (incl. data on editing time and edit distance) (for answering RQ2) were chosen as suitable data sources for this study. According to Yin (Yin, 2003), data triangulation can be considered a strength for a case study, as data with different origins can support each other. The downside would be that data from different sources are harder to compare, particularly for those new to academic research (Fusch and Ness, 2015). Fusch and Ness (2015) view data triangulation from the angle of reaching data saturation: "There is a direct link between data triangulation and data saturation; the one (data triangulation) ensures the other (data saturation). In other words, data triangulation is a method to get to data saturation" (Fusch and Ness, 2015: 1411). This study triangulates data to some extent: in order to support the interview findings about translators' difficulties in PE, data about PE time, and edit distance was collected via memoQ's built-in tools. The main source of data in this study is interviews and other data serve a supportive role.

The interview questions were designed for semi-structured interviews. Semi-structured in this case means the interviewer tried to get answers to all the questions in the interview guide and left openness for exploring further as she saw fit. To test the design of the study, the study task was assigned to one translator and a pilot interview was conducted with them. This translator did not take part in the main study. The pilot interview gave valuable feedback on interview questions and task instructions. The purpose of the interviews was to find out how translators approached the PE task (RQ1), the challenges they faced during PE (RQ2) and their perception of MT (RQ3). Examples of interview questions to achieve this goal follow (the full interview guide can be found in the Appendix I):
Please describe your activities during the post-editing task. (RQ1.1) *What was different compared to translating in memoQ?* (RQ1.2)
How would you assess the effort it took to post-edit this text? How was it to spot the MT mistakes (compared to editing human translation)? (RQ2)
What is your opinion on translation becoming widely accessible thanks to MT platforms like Google Translate? Would you use MT in your everyday work? (RQ3)
As an introductory part of the interview, some experience-related information about the translators was collected (see Table 1 in section 3.2) to serve as a basis for comparisons mainly addressing RQ2. As Lazar et al. (2017) point out, interviewees might forget about relevant details when removed from the work situation. To mitigate this risk, the translators were given access to the PE study task once again before and during the interview, so they could refresh their memory and refer to specific details, if needed. All interviews were conducted in Estonian and lasted between 30 and 60 minutes.

For the purpose of supporting interview findings on RQ2, additional information about technical and temporal PE effort was collected in this study. Data about the technical PE effort was collected via an integrated memoQ tool Edit distance report (memoQdocs: Create edit distance report). The Edit distance report (refer to Figure 8) is an automatic metric showing the amount of changes made to the translation.

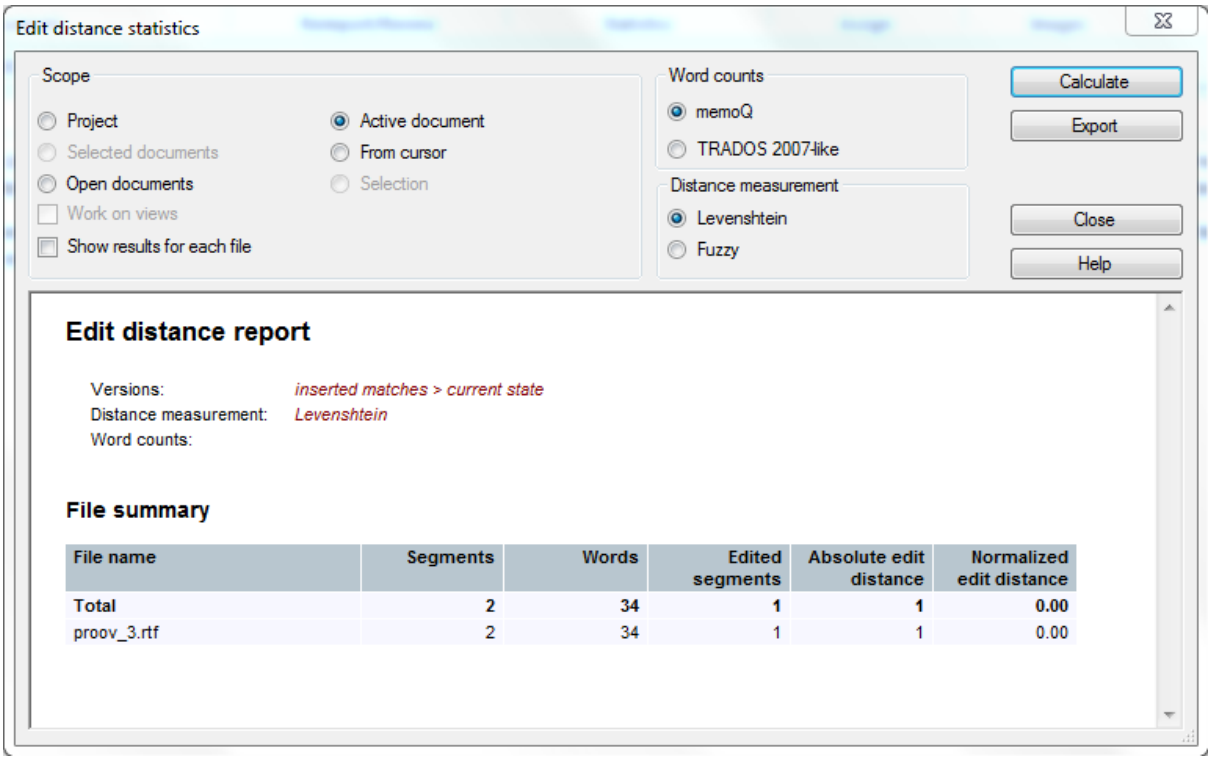

Figure 8. Edit distance statistics in memoQ reflecting the change of 1 character compared to MT text.

For this research, Levenshtein measurement is chosen and presented in normalized edit distance. Levenshtein distance is defined as "/.../ a minimum number of Levenshtein edit operations needed to transform one string into the other /.../" (Garabık, 2006: 2). Such operations are insertion, deletion, and substitution; all on a character level. Normalized edit distance used in this study ranges from 0 to 1 and reflects the amount of changes in percentages. If nothing was changed, the normalized edit distance would be 0 (0%) and if everything was changed it would be 1 (100%).

To gain information about temporal PE effort, the translators were asked to provide an editing time report. For this purpose, an integrated memoQ tool for measuring editing time (memoQdocs: Create editing time report) was used. The Editing time report is an automatic metric which records time spent on editing a document. MemoQ offers an editing speed measurement in the form of characters or words per hour and the total time spent on the task in hh:mm:ss. The total time used in this study is in minute-level precision.

To find out more about PE subtasks to help answer RQ1.1, in addition to interviews also PE texts are taken into account, since tracked changes reveal information about physical writing processes, i.e. the types of edits done by translators. However, this gets analyzed only at a very general level and precise conclusions about the nature of the edits remain out of the scope of this research.

In order to validate that the post-edited texts have a necessary quality level (that translators made needed changes into the MT to edit it up to human quality), a typology of PE mistakes was needed. Since the SAE J2450 metric is designed specifically for the automobile industry (SDL Producthelp: SAE J2450 QA Model), it was chosen as a guiding quality metric for this study. Two professional editors worked on the quality analysis (QA): one found the mistakes in PE texts and the other one categorized them based on SAE J2450. The QA served to validate the necessary quality level of the PE texts while more detailed conclusions drawn on quality fall out of the scope of this study.

3.5 Analysis Procedure

To find answers to the research questions from the interviews, the method of thematic analysis was chosen. This choice was motivated with thematic analysis being a flexible method independent from "/.../ a particular epistemological or theoretical perspective" (Maguire and Delahunt 2017: 3352). For the most part, thematic analysis guidelines given by Braun and Clarke (2006) were followed. The interviews were transcribed, so that "/.../ the transcript retains the information you need, from the verbal account, and in a way which is 'true' to its original nature /.../" (Braun and Clarke 2006: 88). The transcripts were entered into a program called Obsidian.md which then served as the tool for the coding, categorizing, and mind-mapping that followed. The steps of the thematic analysis are presented in Figure 9.

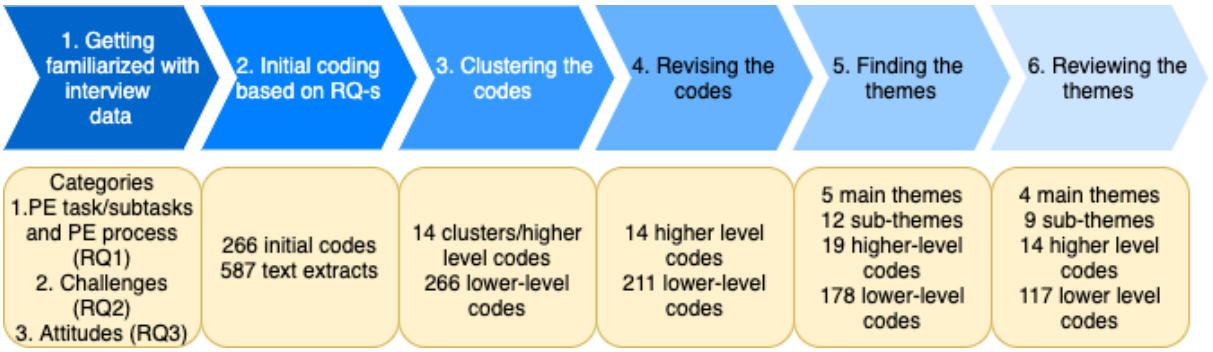

Figure 9. The process of working with interview data with respective results.

The approach to create the initial codes was top-down theoretical — the interviews were coded in regard to research questions with three wide categories in mind: PE task/subtasks and PE process (RQ1), challenges (RQ2), and attitude (RQ3). What was considered to remain out of the scope of these categories, either did not get coded or in case of pertinent information, got temporarily coded under Miscellaneous, with the possibility of placing it in the coding scheme at a later time. The researcher tried to "/.../ pay attention to the process of "zooming in and zooming out" in the coding process to avoid de-contextualizing the data (Sotiriadou et al., 2014) /.../" (Chandra and Chang 2019: 102), meaning the researcher coded sentences or passages of the transcript rather than single words or short phrases and used precise initial codes arising from the data, for example: "PEMT of this task similar to editing human translation" (RQ1), "memoQ slow when a lot of partial matches" (RQ2), "MT will be used more widely" (RQ3). Initial coding of the eight interviews was an iterative process

where codes got added, deleted, merged, split, and changed. Next, codes were divided into 14 clusters, which serve as higher-level codes.

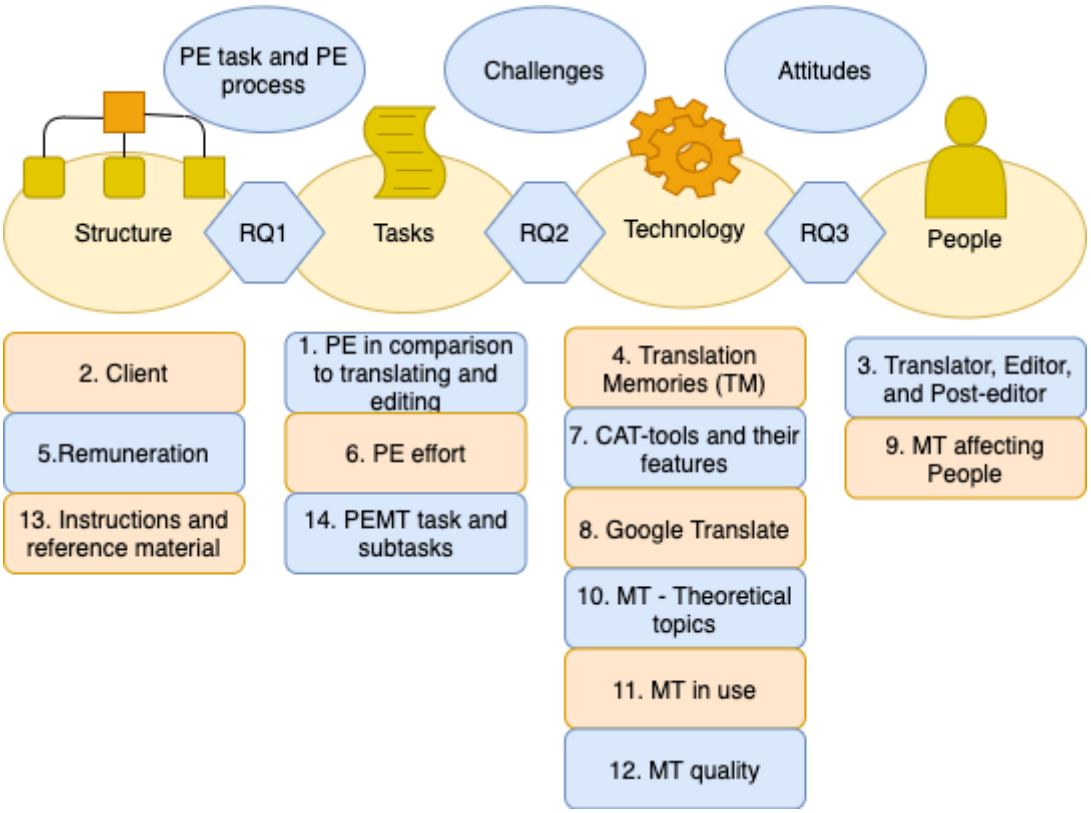

Figure 10. Socio-technical categories and numbered cluster subjects.

When the codes were divided into clusters, the author placed the socio-technical categories of people, technology, tasks, and structure under the three RQ-motivated categories as seen in Figure 10. In short, this framework could be described as: the structural need shapes the PE task (*RQ1 addresses the PE task and process in detail*) which was achieved with MT and CAT-tools (*RQ2 addresses the difficulties that arise from doing the task in its technological context*) by translators (*RQ3 addresses how translators perceive the MT technology)*. This clustering was taken as a theory-driven measure to cognitively process and group all the codes before reducing the multitude of lower-level codes, and finding themes and subthemes. The three RQ-motivated and four socio-technical categories formed a more abstract level to organize the cluster subjects (refer to Figure 10) from a theory-guided perspective that a researcher always has, given that "a naïve realist view of qualitative research" (Braun and Clarke 2006: 80) is a misconception.

After one more iteration of revising the codes and reducing the number of fine-grained lower-level codes, the search for themes could start. According to Braun and Clarke "[a] theme captures something important about the data in relation to the research question, and represents some level of patterned response or meaning within the data set" (Braun and Clarke 2006: 82). Thus, the criteria used in this research for finding themes can be described as considering how closely a topic relates to research questions and counting how many

translators mentioned this topic. Saldaña points out that "[a] theme is an outcome of coding, categorization, and analytic reflection, not something that is, in itself, coded /.../" (Saldaña, 2015: 13). This research takes after Saldaña's suggestion by starting to determine themes only after careful work with data and codes.

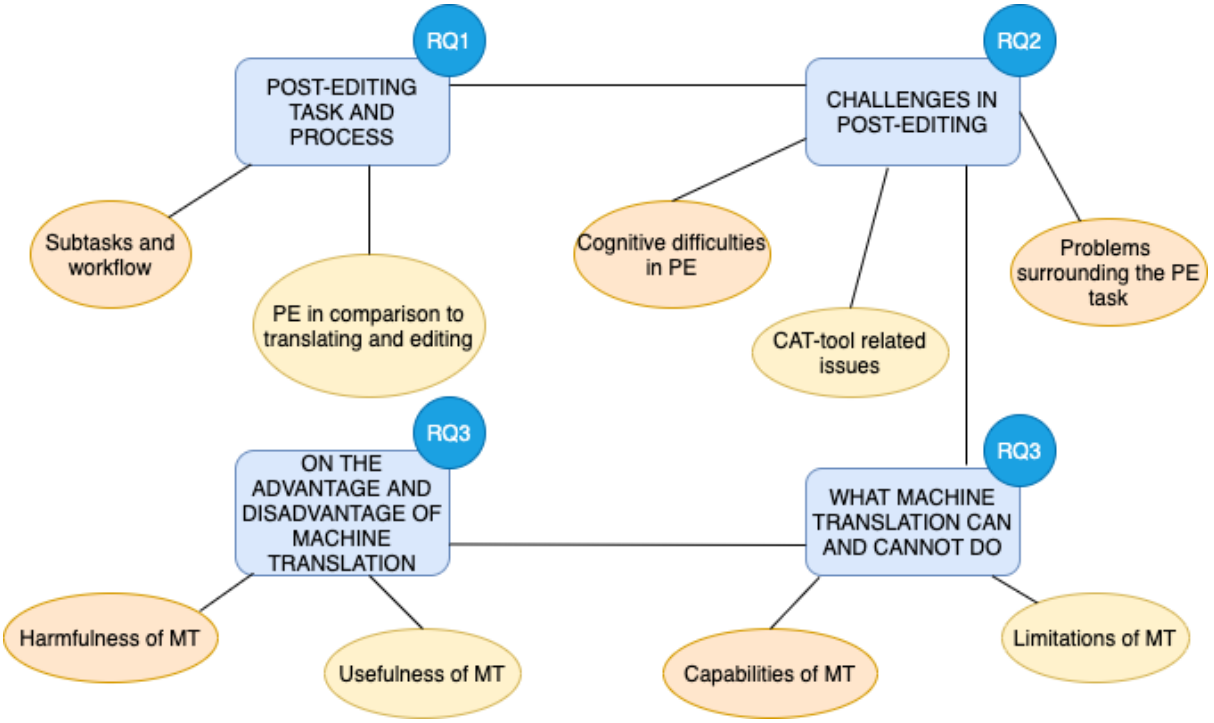

Figure 11. Mind map with the final four main themes and their subthemes.

The interest of this research is to find out how translators performed the PE task assigned them during this study (what their strategy executing it was, including what subtasks they performed and in which sequence; how they described it being different from their translator and editor experience) (RQ1); what difficulties translators faced while completing the PE task (RQ2); what are translators' attitudes towards MT (RQ3). In the search for themes, the attempt was to have one or two themes per research question. In the phase of reviewing themes, it was necessary to assure what Braun and Clarke (2006) point out as 'internal homogeneity' and 'external heterogeneity' of the themes (Braun and Clarke, 2006: 91). The final set of themes and subthemes in relation to research questions is presented in a mind map in Figure 11.

3.6 Ethical Considerations

The study uses information about translators' professional experience, attitudes and their work performance. To protect translators' identities, every translator was granted a participant number. Translators' interview quotes used in this thesis do not contain personally identifiable information, such as their names, company names, project names, specific locations, etc.

4 Results of Thematic Analysis

This chapter presents the results of the thematic analysis of the interviews. The analysis is built upon the three research questions of this thesis:
RQ1. How do translators perform the post-editing task?
RQ1.1. What are the subtasks and in which sequence do they occur?
RQ1.2. How does post-editing compare to translation and editing?
RQ2. What are the challenges translators encounter when post-editing machine translation?
RQ3. What are translators' attitudes towards machine translation?

The analysis results are structured in themes and subthemes according to the thematic mind map in Figure 11 and with a reference to the research question they address in the caption. Subtheme sections contain higher-level codes and their definitions. Some of the subthemes used to be higher-level codes and this is noted respectively.

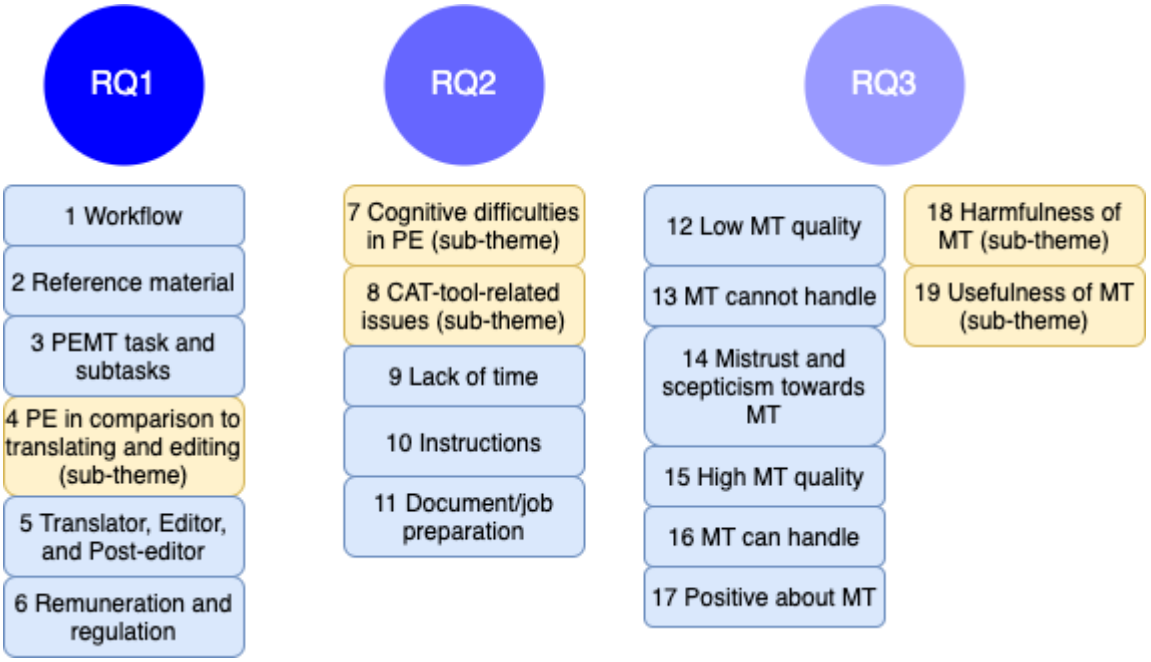

Figure 12. Final higher-level codes and higher-level codes turned into subthemes (marked yellow).

The list of lower-level codes and the number of times they were mentioned can be found in the Appendix II. Translators' quotes are linked to the respective translator numbers (refer to table under section 3.2). All original quotes are in Estonian and translated for this thesis by the author. Final higher-level codes in connection to research questions can be found in figure 12.

4.1 Post-Editing Task and Process (RQ1)

This theme addresses how the PE task as a part of this study was approached, what its subtasks were, and in which order these subtasks were performed. Based on the study task and translators' general experience in the field, skills and knowledge necessary for PE, and

the rights and responsibilities of translation professionals were discussed. PE was compared to translating and editing.

4.1.1 Subtasks and workflow

The subtheme of subtasks and workflow is directly related to the PE task given to the translators. It contains codes about steps the translators took to complete the PE task assigned to them for this study, specific examples in the task that caught their attention, and their thoughts about the process of PE. Related to necessary tasks for PE, the topic of remuneration emerges. The following higher-level codes were thematized under this topic:

- **Workflow** - The order in which tasks are performed.
- **Reference material** - Sources used by translators to refer to while PE.
- **PEMT task and subtasks** - Details about the study task and its subtasks.
- **Remuneration and regulation** - PE fares and standards.

Workflow. The approach to the PE task was not uniform among participating translators. Four translators mentioned that they needed to take some time first to get to know the task to "/.../ *get a better overview of the text, a clearer overview of what it is about*." (T5) Getting the overview was either more thorough (reading source and target through) or less thorough (just glimpsing at the text). One translator said they looked at the length of the task, then confirmed all the machine-translated segments and started editing the target segments step-by-step after confirming them. In contrast to this approach, two translators started PE segment-by-segment without confirming the edited segment before starting to edit the following one. One of the two translators who chose the strategy of not confirming the segments realized after some time that it did not work in their favor, because non-confirmed segments are not being saved in the translation memory and not displayed as reference material. The other translator who chose to edit without confirming the segments followed it through until the end of the task and confirmed all segments at once after reading the translation through one more time (the final check). The third, and most common approach, was to confirm the segment after editing it, which was the path that the rest of the translators chose. The prevalent strategy for working within segments that three translators mentioned was to first read through the source, then the target, and then find out what needs correcting and correct it.

Reference material. While working with the task, the translators did not stay within memoQ during the entire task. There were terms, such as "Bedienelement" ("control element") and "Multimedia-System" ("multimedia system") that needed more clarity so they googled (mentioned by 6 translators), used web dictionaries (mentioned by 2), and visited the car manufacturer homepage (mentioned by 3). One translator emphasized approaching the reference material source-critically and disregarded some popular forums and social media. The translators mentioned googling car models, terms, but also in order to understand the meaning of a sentence in a wider manner than just searching for terminology. On the car manufacturer home page the translators referred to figures and terminology. They did it both for the sake of understanding the text and in an attempt to keep the terminology consistent.

One translator admitted to having consulted with someone who knows more about cars. Two translators mentioned getting confused because the machine translated “mehr” (“more”) as “>” (“greater than”), whereas depending on the context “refer to more” would have been the correct translation.

PEMT task and subtasks. All eight translators performed a final check which varied from quickly scanning the text to see if they noticed any obvious inaccuracies to reading the post-edited text more carefully. One translator mentioned using memoQ QA check to help with performing the final check and another translator said they used spellcheck (using QA and spellcheck were not required for this task). One translator said they would have liked to let the client know about some problematic passages and ask for more information about the hyperlinks (if they should end with DE or ET), among other things.

Two translators mentioned that this task took them more time than they expected and one said they did not want to accept the task because “/.../ *I know that such things take time and there is no particular financial result*.” (T6) One of the translators mentioned the first 20 segments taking more time than the rest of the task. Another translator said that refreshing their knowledge about memoQ took some extra time. Two translators explicitly mentioned that there was a lot to correct in this task.

Remuneration and regulation. In terms of remuneration, there was some uncertainty because PE is a relatively new task and the time and effort spent on it are difficult to estimate. It was mentioned that some LSPs allow optional usage of MT as a reference that one does not have to use, but if MT gets used, the price will be calculated differently. Two translators pointed out that PE is not just proofreading/language editing and the rate should include thoroughly comparing the translation to the source text. One translator showed concern that with the wider spread of MT the translation agencies will start offering rates based on time and not on the amount of words.

One translator pointed out the field of MT should be more regulated: “*If it is in a way that everybody is tinkering in their own program X, what will the quality be? It should be regulated by universal standards. If the MT should spread and become universal, which I think will happen, then there should be specific legal regulation and rules the MT should comply with and it should be determined which program will be used for producing this MT*. ” (T8)

4.1.2 Post-editing in Comparison to Translating and Editing

This subtheme is a higher-level code turned into a subtheme and it views how translators compare PEMT to translating and editing human translation. Additionally, there is one more higher level code under this subtheme:

- **Translator, editor, and post-editor** - Insights about what it takes to be a post-editor and ideas about what translation professionals should be liable for.

PE in comparison to translating and editing. It was pointed out that translating in the contemporary translation market already includes a great proportion of editing — there are translation memories linked to translation jobs with the work of previous translators on the same project that need editing and there are fuzzy matches produced by the system. As one translator put it: "*Inevitably, the earlier translation was also semi-editorial or contained strong editorial elements. You need to look at what the system has to offer to you and then accept or change it accordingly*." (T5)

PEMT gets compared to using translation memory. Translators have experience with better and worse TM-s. In case of a good TM, its 100% matches might need less editing than machine-translated matches. But in the case the TM is of a low quality it was deemed to be less helpful than MT. The same logic applies to fuzzy matches; and in addition to TM quality, MT engine's quality also factors in: "*In case there is an MT engine, a good one, then there is no big difference, if you edit FM or MT. FM has some content differences, whereas good MT is 100% compliant to the source text. Thus, it depends on the quality of MT*." (T7) Three translators mentioned editing fuzzy matches to require more effort than editing MT.

Translators expressed different thoughts about machine errors in comparison to human errors: "*MT errors are definitely different from colleagues' "translation gems" and mistakes*." (T5) and that in PEMT "*[i]t is easier to see what to correct.*"(T2) It was also mentioned that MT is expected to fail and thus it is easier to find mistakes: "/.../ *I think that MT is easier in this regard, because you know it is MT and you expect there to be errors. It is a big trap /.../ that you have a translation that is done seemingly very professionally and then there is some wrong term*." (T3) Additionally, MT errors were said to be more unexpected and more automatic compared to human errors which were said to be more logical.

Half of the translators found PEMT of this task to be similar to editing human translation. More specifically, they said it to be similar to a translation of an inexperienced and substandard translator. It was also pointed out that, although PE is similar to editing human translation, PE requires paying more careful attention to the source. It was mentioned that in case of a poor MT, editing MT could be more stressful than translating. Most translators expressed their preference of translating over PE: "*I /.../ found MT to be interesting and useful, but reached the conclusion /.../ not want to work either on light or on full post-editing projects*." (T4)

Translator, editor, and post-editor. There was no agreement between translators if PE is a task more suitable for an editor or for a translator or requires the best of both of them. One translator said about editors doing PEMT: "*Yes, if the editor has a command of the source language and even if they do not have direct translation experience. Although, editing experience is already as good as translation experience*." (T5) Another translator said: "*If PE will be left only to an editor, then it requires domain experience. /.../ otherwise it will be all sorts of nonsense*." (T4) It was said about whose task PEMT should be: "*I don't think there is too much difference, but the prerequisite is that both should be very good at what they do. A translator who is post-editing MT should be a very good translator.*" (T7) For some

translators, PE seems to bear more resemblance to editing and for some more to translating. However, it appears that a post-editor should think like a translator ("*By editing MT one must approach it like a translator. One needs to understand exactly what there is. This is easier by traditional editing*." (T1)) and work like an editor ("*An editor would be a safer bet in my opinion, they would know even better about the grammar of their mother tongue*." (T8)).

Considering that translation in combination with TM already contains a lot of editing, as was already mentioned, one translator says: "*But where does the border between a translator and an editor run? I think of myself as a translator, because my tasks are called translation*." (T5) Mentioned by three translators, an integral responsibility of both an editor and a post-editor is that they are the last person who needs to guarantee the job is flawless and ready for the client/publication. This requires a certain amount of decisiveness and confidence: "*No, I am not an editor, I am not 100% sure of myself and I am the one asking questions and checking everything from the EKI (Estonian Language Institute) home page.*" (T8)

Two translators mentioned that a post-editor should guarantee terminological consistency not only in the task that they are post-editing but also in the entire project and thus refer to client materials and earlier translations. Three translators think that a post-editor should feel responsible for correcting MT errors. Two translators stated achieving high quality and getting the text to flow more naturally as their goal in PE: "*I tried to find passages where I could make additions to give the text more marketing appeal, make it less mechanical*." (T8)

4.2 Challenges in Post-editing (RQ2)

This theme consolidates subthemes related to difficulties of a different nature and origin that translators faced in connection to PE both during this task and from their previous experience.

4.2.1 Cognitive Difficulties in Post-editing

Current subtheme is a higher-level code turned into a subtheme and it takes a look into mental matters that translators deal with while PE.

Three translators said PE poses difficulties for them based on its novelty and them not being used to it. One of these three said that seeing the right column filled with text created a strange situation for them: "*Based on my earlier translation experience I would prefer the right target column to be empty*." (T5) Two translators mentioned, it was difficult to find balance between one's own preferences and things that truly need correcting in MT. One translator said that MT might appear nice at first glance, but on a closer look it reveals problems with conveying the meaning of the source text. Another translator mentioned, it is tiresome to force oneself to doubt everything while PEMT. One translator said MT made them feel they needed to correct for the sake of correcting. After making some of such forced corrections they changed the affected text back to a machine-translated version. Another translator said MT errors were harder to spot due to being more unexpected than mistakes

made by humans. One translator mentioned that it was difficult to post-edit a short text due to less context and taking time to get familiarized with the text does not feel as rewarding as in the case of a long text. One translator pointed out that in the case of PE it is not possible to discuss translation choices with a translator. For one translator it took a lot of effort to compare machine-translated text to the source.

Three translators mentioned that MT can interfere with the train of thought: "*MT is starting to interfere with my train of thought. If you read it through, then maybe you will stick too much to what the machine has done and then try to change it as little as possible. But if you can think on your own, then the sentence builds up in your head and comes more naturally. Maybe it is a matter of getting used to it.*" (T1) and another translator said: "*It depends on the translator, but I know translators who get a mental block and stick to machine-translated sentences. /.../ For me as well, if there are some parts given to you, you will continue thinking about them.*" (T3)

In Table 2 is the cognitive effort assessment translated into numbers.

Table 2. Translator's assessment on cognitive effort of this study task

Translator	**Verbal assessment of the cognitive effort of this task**	**Rating interpretation**	**Numeric rating**
T6	"In this case it was very good to have pretranslation, it was much easier this way."	Low effort	1
T2	"Pretranslation makes it easier to correct. /.../ MT simplifies work a little bit. So it is positive."	Low effort	1
T4	"Regular brainwork."	Medium effort	2
T1	"Medium effort. It was not too difficult or too time-consuming."	Medium effort	2
T3	"It was not too big of an effort, it was like editing in general."	Medium effort	2
T5	"/.../ the cognitive effort is higher." (*in comparison to translating*)	High effort	3
T7	"The effort to compare the MT to the source is really high."	High effort	3
T8	"I had cognitive issues with this task. /.../ The cognitive battle as an editor about how much to change and if I am doing the right thing at all."	High effort	3

4.2.2 CAT-tool Related Issues

This higher-level code turned into a subtheme takes a closer look at what translators say about CAT-tools and their different features for not only post-editing the study task, but also in a wider sense for translating and editing. This wider perspective was discussed because PEMT in memoQ is approached similarly to translating and particularly editing — there are two columns (one with the source and one with the target language) and the translation professional has the same toolkit plus MT matches. Predominantly there were challenges mentioned related to CAT-tools, but also some helpful aspects get mentioned under this subtheme in contrast to difficulties.

Three quarters of the translators shared positive opinions of memoQ as an environment for PE. Among other things they said it to be "/.../ *the most logical and easiest compared to Studio and Wordfast*." and "/.../ *this program is good in terms of usability*." The translators mentioned different memoQ features they use in working with translations:

A. the option to open the original document as reference (not possible in this study task),
B. QA check,
C. track changes (part of default task setting),
D. spell check,
E. adding comments,
F. using several translation memories at once (not possible in this study task),
G. key shortcuts,
H. memoQ internal search to achieve term consistency.

Two translators mentioned liking PE in Memsource more for better convenience and functions compared to memoQ. One translator mentioned preferring memoQ to all other translation tools and another one preferred it over Across but not over Memsource. Five translators said that in their day-to-day translation work they prefer Trados over memoQ.

Following problems and limitations in connection to memoQ were mentioned:

A. QA is problematic and inconvenient (mentioned by 2), while Trados has icons for indicating possible errors already during the translation process (these specifics about QA were mentioned by 1),
B. performs slowly (compared to Trados - 1; compared to Memsource - 1; in general - 1) (3),
C. the text in memoQ is chopped up more (harder to follow and get an overview) (1),
D. creating translation memory is more difficult than in Trados (1),
E. it is not possible to change text colors (as it is in Trados) (1),
F. automatic capital letter at the beginning of segments (2),
G. formatting superscript and subscript is problematic (1),
H. markups/tags are a hassle (also in other CAT-tools) (1),

I. using buttons is a hassle (keyboard shortcuts help) (1),
J. difficult to find settings (1).

Besides considering the different aspects of usability and functionality listed above, one important reason for preferring a certain CAT-tool is habit: "*They are both very good, but the one you are more used to seems as if it is more convenient. If one would use memoQ more, this would be convenient to use as well.*" (T6) And another translator said: "*I am a little bit lazy, too. Once I have learned to use one program and am fine with it and it is suitable for the client as well, then I am not looking around anymore. Maybe this is a mistake, maybe I would find something better, if I would look for it*." (T3)

4.2.3 Problems Surrounding the Post-editing Task

In addition to cognitive difficulties that arose from the task and problems with CAT-tools, there were other issues mentioned that surround the PE task. Some of the issues mentioned were directly about the study task, while others were more general. This subtheme covers the findings under the following higher-level codes:

- **Lack of time** - Problems that are connected to not having enough time.
- **Instructions** - Problems with PE instructions.
- **Job preparation** - The PE task was not prepared according to the translator's expectations.

Lack of time. Four translators mention the lack of time ("*For me as someone who is constantly busy /.../*"; "*/.../ takes time and often there is no time*" (T7)), while only one translator said to have had enough time to take as long as they needed for the task.

Job preparation. Four translators would have wanted to see the original file/reference ready for them while performing the task. One translator would have preferred some segments/terms to be locked according to the project needs (client preferences in terminology and proper names that are not to be translated). As a general remark about all translation jobs and not just PE, formatting the layout of the translation document was mentioned by two translators as a task that should be done by someone else (by a technical editor, for instance).

Instructions. In connection to the option of locking terms or segments, it was pointed out that project instructions should include how to approach links, proper names, company names, among other aspects. Three translators mentioned that the instructions for the study task lacked the information about the translation's purpose of use or target group. Unrelated to this study task, two translators said that light PE instructions and its purpose are unclear to them: "*Who would need this partly post-edited text that is not correct, but just gives an idea of the text?*" (T4) One translator mentioned that often there are no instructions accompanying the PE task at all. One translator said they would prefer instructions to be written, because seminars take time.

4.3 What Machine Translation Can and Cannot Do (RQ3)

This theme addresses MT's perceived limitations and capabilities that shape the translators' attitudes towards MT.

4.3.1 Limitations of Machine Translation

This section takes a closer look at aspects in translation that MT is not able to perform that well or is not able to do at all. These aspects are grouped under three higher-level codes:

- **Mistrust and scepticism towards MT** - Insights about the limitations of MT.
- **MT cannot handle -** Examples of things that MT cannot do or cannot do well.
- **Low MT quality** - Mentions of MT having poor quality.

Mistrust and scepticism towards MT. Although several translators admit using MT in situations where they are very busy or need to translate from a language they do not understand, or simply to see what the machine has to offer, five of them have expressed a more or less critical position towards using MT professionally. For example: "*I am able to use the MT module /.../ [and it] is offering me MT options, but they are mostly useless. I would say 95% is my own work, if not more*." (T4) and "*In my work I have not used MT /.../, unless it is a PEMT task*." (T1)

Three translators said MT needs human editing to achieve the necessary quality or a human touch. Three translators did not believe MT could replace human beings, either not yet, considering its current state, or ever, since the machine is limited in the capacity of understanding context. Five translators approached MT very carefully, expecting there to be errors, referring thoroughly to the source, and keeping their guard high to be sure they do not miss anything: "/.../ *you will need to be conscious and cautious, not to take it at face value*." (T5) Another translator expressed their doubt: *"In the case of MT, I have a big fear that there might be something translated in a very wrong way*." (T7) Two translators mentioned trusting another human more than a machine, with the reservation that this human should be a translation professional with known high work standards: "*I would trust my good colleagues and collaboration partners who can be trusted based on previous experience. Not everybody*." (T5) So, in terms of trust, the machine was seen to beat a total stranger but remain inferior to a trusted colleague.

MT cannot handle. The rather sceptical attitude towards professional use of MT seen in the previous passage can be linked to several limitations that translators pointed out in connection to MT:

A. MT cannot translate **literature** (mentioned by 6): "*The machine is not able to produce creative, belletristic text, and convey it as the original does. It is not able to provide this aesthetic quality*." (T7)
B. MT cannot handle **sentence structure** (6): "/.../ *MT sentence structure reflects the source rather than target language*." (T5 - said based on this study task specifically).

C. MT does not understand **context** (3): "*There are things that the machine cannot understand, such as the nuances of language, context, other background that a human can take into account*." (T1)
D. MT does not take **client terminology** or **terminological consistency** into account (3): "*There was terminological inconsistency, the same thing was translated in different ways, although the context remained the same*." (T4 - said based on this study task specifically).
E. MT does not handle **legal texts** well (2): "*It should be an extra good MT that could differentiate these terms and aspects, conform to the idea and convey it.*" (T3 - said about the ability to differentiate between common law and statutory law.)
F. MT does not handle **long sentences** well (1): "/.../ *nice long complex sentence with coordinating conjunctions, then it fails*." (T5)
G. MT does not handle **metaphors** and **proverbs** (1): "/.../ *it cannot handle metaphoric expressions and proverbs*." (T5)
H. MT has no **human touch** or **feeling** (1): "*About some sentences I felt they were too alien. /.../ I cannot help the feeling about MT that it takes everything word for word and there is no human touch or feeling in there*." (T8 - said based on this study task specifically).
I. MT cannot handle **word structure** (1): "*And there was congruence, inflection, there were some incorrect translations*." (T4 - said based on this study task specifically).
J. MT is not able to make text **user friendly** and **marketable** (1): "*Making the text to be more user friendly, so that it would not be that automatic. Making text softer and more marketable, these things it cannot do*." (T8 - said based on this study task specifically).

Low MT quality. Four translators said that the MT in this study task was of bearable quality: "*It was OK. /.../ On a five-point-scale, I would say 3, it was not that bad*." (T1) and "/.../ *I think this level of translation was relatively bearable*." (T7) and "*If I would grade it on a five-point-scale, then [I would give it] average 3. I would not give it a higher grade*." (T4) Two translators who said the quality to be grade 3 brought this scale up themselves for describing the quality of the study task MT. One translator said the quality of this study task to be poor and after specifically asked, rated this with 2 on a five-point-scale.

When discussing MT with someone, to avoid talking about Google Translate (GT) would be like not mentioning an elephant in the room. Interview questions about the use and opinion on GT were asked a part of this study. Six translators expressed GT being of questionable quality: "/.../ *there is a need to observe the quality. If someone believes Google Translate will give you an excellent text, then the actual result could be all sorts of gibberish*." (T3) One translator pointed out that GT is like a synonym of MT for many and as such it is not a synonym of quality: "*It has been in web browsers for a long time. Knowledge and assessments of it are accompanying us everywhere, be it personal life or reading comments in social media or public media. It is not a synonym of quality, yet. When someone brings up Google Translate, then they rather say that it is of questionable value*." (T5)

4.3.2 Capabilities of Machine Translation

In contrast to the previous subtheme, this one addresses what MT is able to do. Although there are fewer aspects mentioned here, it is notable how different translators can have contrary opinions about some aspects of what MT is able to do or not able to do. This subtheme brings together the following higher-level codes that reflect the code structure in previous section:

- ➢ **High MT quality** - Mentions of MT having good quality.
- ➢ **Positive about MT** - Affirmative insights about MT.
- ➢ **MT can handle -** Things that MT can do well.

High MT quality. According to five translators, the quality of MT in this study task was either rather good or good. As we saw in the limitations section, according to five translators, the quality was bearable or poor. How can this be, if there are 8 participants total? It is important to point out that one person can express different opinions about the same topic within one interview, depending on the exact context or other factors (even change of mind can occur). As mentioned in section 3.5, the approach to mitigate the risks of misunderstanding the interviewee was to code longer text extracts. Three out of the five translators who expressed positive opinions on the quality of this study task said they were positively surprised to see the good quality. This means they were expecting less from MT. It was said they have seen worse: "*It was good for MT. I have seen something different ... Even if you use Google Translate, it might not be that good at all*." (T3)

MT can handle. One translator pointed out that the sentence structure in this study task was good. This opinion contrasts the opinion of six other translators who found MT sentence structure to be lacking. In the limitations section (4.3.1), there were other aspects besides sentence structure listed about grammar that MT is not able to handle or handle well. Inversely to this, one translator said MT is developed so well nowadays that it can deal with grammar fairly well.

Three translators believe MT manages technical translations well: "*Where it is sure possible to use machine translation are user manuals, technical texts, why not medicine where the terminology is very specific*." (T8) Two translators said that MT handles short sentences well in contrast to longer ones that are known to be among MT limitations: "/.../ *easy sentences, such as "Die Löschtaste ist unter dem Bedienheit", if they are not long mammoth sentences, then I think the machine handles text like this well and helps to save time*." (T8) One translator pointed out that MT can sometimes offer words that the translator or editor has forgotten.

Positive about MT. One translator said that MT might have done better at some places in this study task than they did. Two translators thought this study task MT had better quality than some translators could achieve. One translator thought that MT would be good enough for using in-house to give people an idea what the text is about.

Three translators expressed their positive opinion on custom MT engines in comparison to generic ones: "*The free ones I would not use in my work, they are not as trustworthy. Yes, these [custom] ones, they are much better*." (T1) and "*It is important that it would contain one specific theme, this is much more reliable*." (T6) Two translators shared kind words about generic MT as well, they said GT is able to convey the meaning. Half of the translators said MT is constantly developing: "/.../ *it is getting better all the time. I imagine in the future it would rarely happen that MT gives something which is absolutely impossible and you have to rewrite every sentence and all the terms are wrong*." (T3)

Translators listed different information sources about MT, such as language events, conferences, seminars, training sessions, help materials from LSPs, print and online media articles, academic articles, lectures.

4.4 On the Advantage and Disadvantage of Machine Translation (RQ3)

This theme is about how MT is perceived by translators either as a useful or a harmful tool.

4.4.1 Usefulness of Machine Translation

This higher-level code turned into a subtheme looks into how translators find MT to be useful on a personal and also on a wider level.

Whereas the list of MT limitations was longer than the list of capabilities, the list of useful aspects of MT is longer than the list of harmful ones. Four translators expressed optimistic thoughts about professional use of MT: "/.../ *I have used MT for specific tasks in the last 6 months or so and I use it there all the time. Yes, it is useful, very good*." (T2) Another translator said about potentially using it in their daily work "*I think I would use it, if someone would prove to me this is a very good platform, what the statistics are in there, how the texts have formed, how the quality is.*" (T8) This statement shows willingness to use MT, when there is more knowledge about it and its quality is proven to be good. One translator said they enjoyed the study task. The same translator also mentioned they do not feel affected by the proliferation of MT, because they received more translation offers than before. Two translators mentioned MT being helpful when they have been under time pressure.

Four translators shared their insight of MT getting used more widely: "*This is the direction the translation market is going and maybe it will be prevailing one day*." (T5) and "*Yes, the direction is towards it, it pays off to keep oneself informed*." (T1) Five translators said they have used MT for personal purposes. It came out from most interviews that although there was prevalent scepticism about professional use of MT, translators had a positive attitude towards MT (and GT in particular) being used by individuals for private purposes: "*If one is in trouble, then one does not have to suffer under suspense, one can use Google Translate and get an overview, there is no need to seek out a translator. /.../ If we are talking on a professional level, I do not think that industries will start using only Google Translate any*

time soon. On the individual level I welcome Google Translate and these other platforms." (T8) Three translators mentioned GT as their first choice for personal tasks: "*My first choice is Google Translate, but if I hear about some translation engine some company has built, then I have went to see it and try it out, to find out what the quality is they are offering*." (T7) Three translators have found GT to be helpful when translating via a third language.

Six translators expressed the thought that MT helps to save time and effort: "*It is useful, because it will do the job faster and then post-editing will take less time than translation from scratch*." (T1) Five translators pointed out the usefulness of MT in eliminating repetitive tasks: "*What is the point of 100 translators typing the same sentence every time, for example in safety instructions*." (T6) Another translator said MT allows the translator to focus on interesting aspects: "*When the machine is taking care of the formal part of the work, then in my opinion it is quite nice for the translator to focus on the interesting and substantive parts of the translation*." (T3) A third translator said on the same topic: "*I think machines can do more extensive work, but there is always a need for the human touch*." (T8)

Six translators expressed thoughts about MT simplifying life and communication: "*It is positive, because this way a lot of people who might not understand some language very well, could understand some text through it, more or less, or understand what it is about. It is certainly not bad*." (T1) Another translator said: "*The positive thing is that I can use Google Translate to sometimes get an overall picture or just to simplify my work*." (T3) And according to a third translator: "*These MT platforms have made life easier indeed*." (T2)

Some translators also thought about global matters in connection to MT. One translator said MT helps technology to develop our understanding of the human brain. Another translator believes MT helps to connect the world turned upside down by COVID-19 crisis.

4.4.2 Harmfulness of Machine Translation

Translators mentioned harm that could come to them personally, their trade in general, or even to the entire world as a result of the proliferation of MT. This is, again, a higher-level code turned into a subtheme.

Six translators mentioned that MT creates price pressure in the translation industry: "/.../ *LSPs rush to acquire clients with great advertising excitement, stating everything to be much cheaper now, because the machine will work for us* /.../." (T5) and "*Sure the client will have more money or the translation agency will get rich*." (T8) and "*I think it will make prices drop, because it is not on the same level as translating*." (T1) and "*It is useful for clients, it will take the price of this service down*." (T4) One translator hopefully added that although it is probably making translation rates decrease, one is able to post-edit more quickly compared to translating, so maybe the salary will not be affected that much. The other side of the price pressure problem is MT taking work away from translators. This aspect was also mentioned by six translators and often in clear connection to translation agencies and clients gaining from the use of MT: "*For translators, editors, and post-editors it is definitely not useful,*

because their salary depends on the number of words or characters /.../ If one would make a log file now, an analysis on what the number of 100% matches is, then with MT the share of 0% matches undoubtedly declines, as will people's salary." (T4) Another translator suggested there might be a need for less translators in the future. One translator expressed the thought that maybe translators who worry about losing their jobs to MT are not good translators. Another translator felt that MT technologies are affecting the field of interpreting more than written translation.

According to one translator, texts produced by some low-quality MT engines "/.../ *waste the work and time of translators*". (T8) Another translator expressed the worry that post-editors accept bad quality and this would allow MT quality to negatively influence the quality of the language itself: "*MT will make language to be more homogenous, corresponding to the rules of the engine. /.../ it will start repeating and gets amplified* /.../." (T7) The same translator who pointed out the problem of MT influencing language negatively, had a bad experience with light PE — after accepting a couple of light PE tasks, they started turning these tasks down, because "/.../ *it created a bad feeling when you send in such a poor product*." (T7)

5 Supportive Results

To provide a supportive data source to answer (mainly) RQ2, the translators were asked to change their editing time settings in memoQ to measure the time spent on editing. Six translators were able to do it successfully, while two did not find the right settings. Additionally, all translators looked at a timepiece and sent the number of minutes spent on the task. The editing time measurement is just a supportive piece of evidence to show how much time translators roughly spent on 515 words (470 words when calculating for repetitions) of PE. It was suggested to complete the task without taking breaks. However, this was not controlled in any way. Editing time in memoQ varied from 21 minutes to 86 minutes with the average editing time being 51 minutes. Edit distance is presented in a normalized manner as a percentage. This measurement shows the share of changes made in the text. Edit distance varied from 13% to 25%. The chart in Figure 13 demonstrates the editing time and edit distance in this study task. Comparing editing time to edit distance shows a very weak correlation of: $r \approx 0.13$.

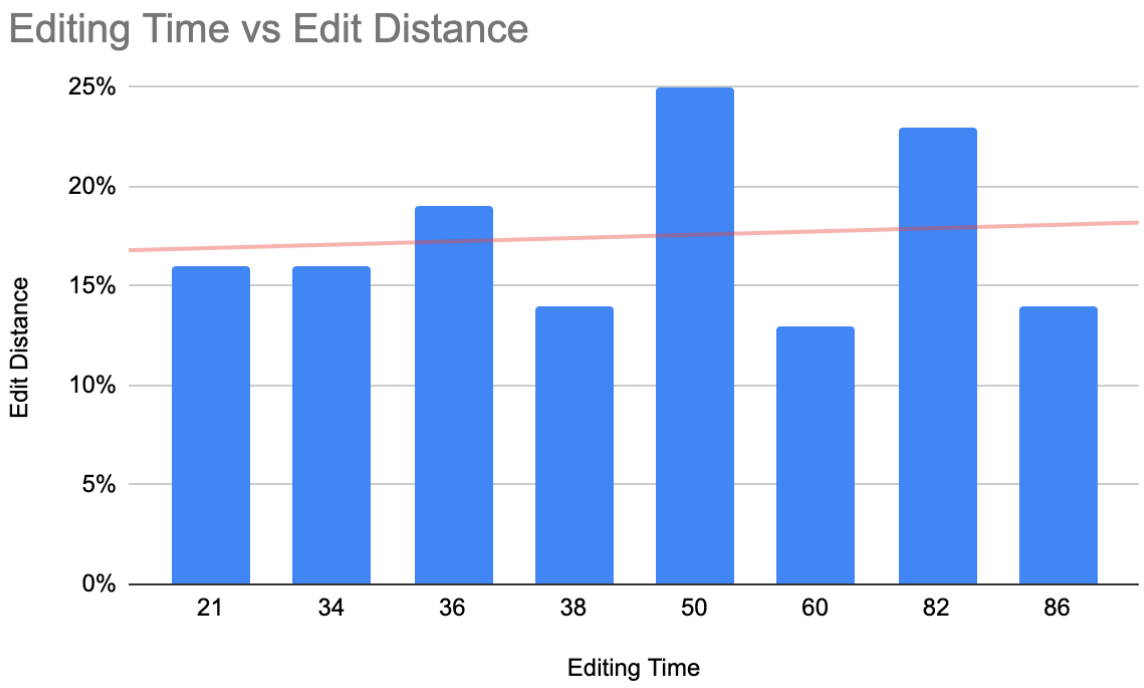

Figure 13. Editing time (minutes) and edit distance (per cents).

An analysis of changes made in the PE text revealed the main strategy chosen by translators was to make use of MT and edit around it, not to delete entire segments and recreate their content from scratch. The editor who corrected the PE texts of the translators deemed the quality of the PE texts to be sufficient. The requirement of editing MT up to the human quality was met.

The translators were asked to assess the effort they put into completing this study task. Based on their replies reflecting perceived cognitive effort, values were assigned to their

assessments (1 = low effort, 2 = medium effort, 3 = high effort; refer to Table 2). When comparing these assessments to the editing time (see Figure 14), the correlation between perceived cognitive effort and the amount of time spent on this task is moderate and negative: $r \approx -0.5$. This is rather counterintuitive and means that translators who spent more time on the task assessed it to be easier than the ones who spent less time.

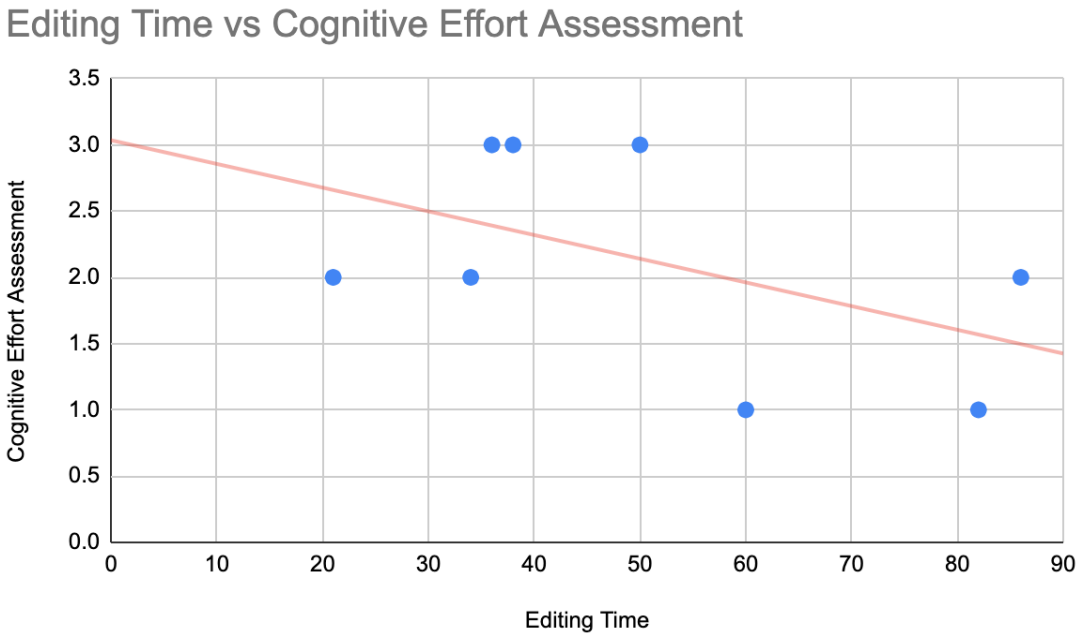

Figure 14. Perceived cognitive effort and editing time.

Although comparison between temporal effort (editing time) and perceived cognitive effort showed less cognitive effort and more temporal effort to correlate, the comparison of technical effort (edit distance) and cognitive effort assessment (see Figure 15) gives a more plausible result of $r \approx 0.17$.

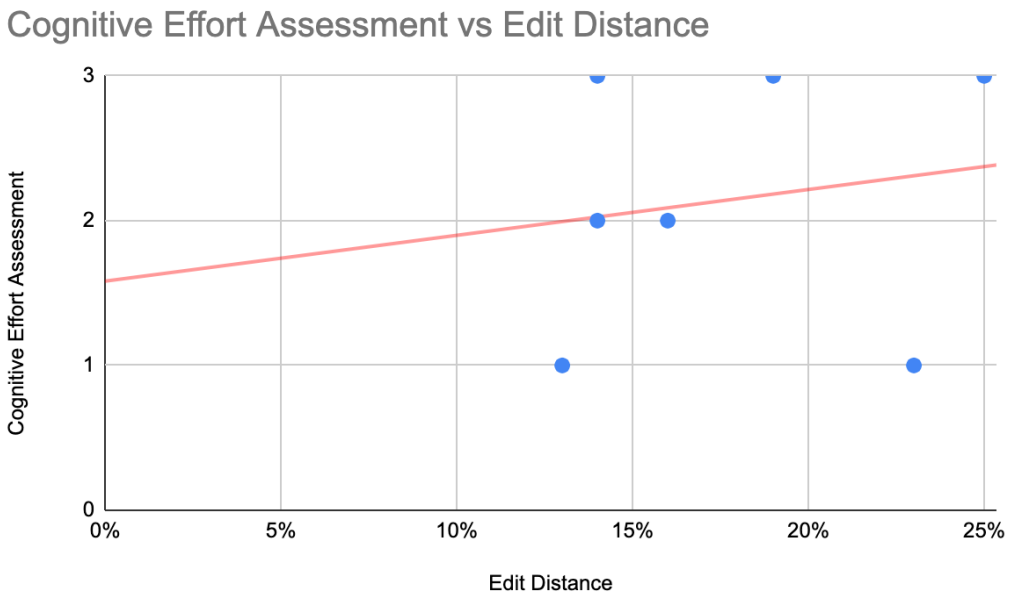

Figure 15. Perceived cognitive effort and edit distance.

Chart in Figure 16 shows a strong correlation between edit distance and having experience with working on PEMT projects assigned by LSPs (1 = existing experience; 0 = no experience): r ≈ -0.72. To see more detailed information about translators' previous experience, refer to Table 1.

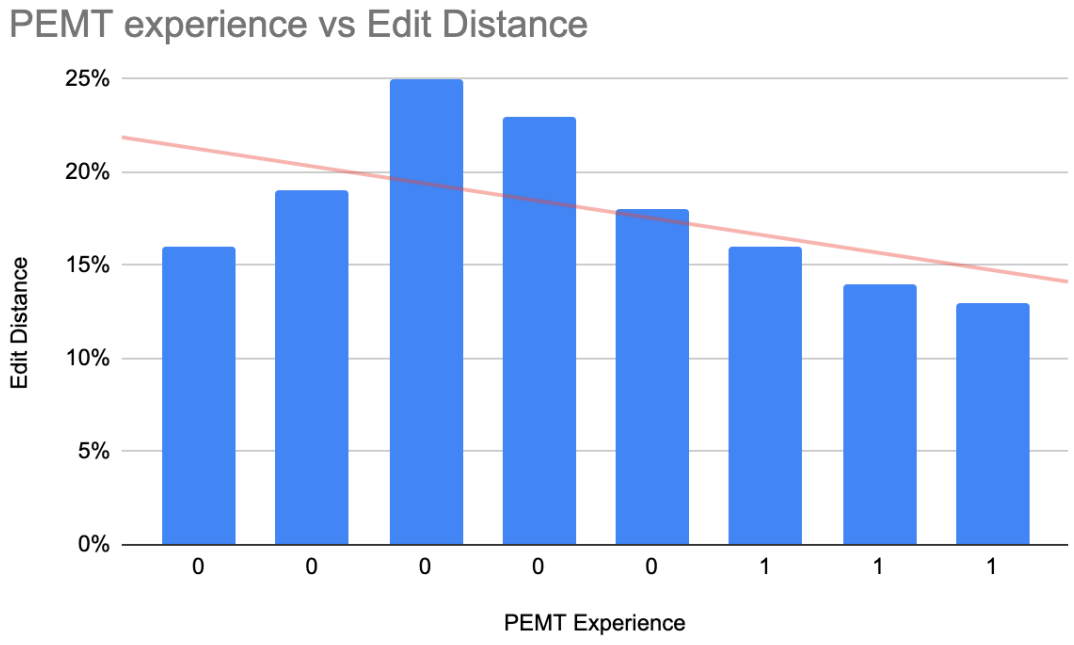

Figure 16. Edit distance and PEMT experience.

The correlation between edit distance and years worked as a translator was very weak and negative, meaning the translators with more translation experience in years made slightly less

changes in the study task: r ≈ -0.08. There is a weak negative correlation between editorial experience (2 = 10 yrs or more; 1 = rare; 0 = none) and edit distance, meaning the translators with editing experience made slightly fewer changes compared to the ones who had little or no editing experience: r ≈ -0.37. The correlation between translation experience and editing time was weak: r ≈ 0.18.

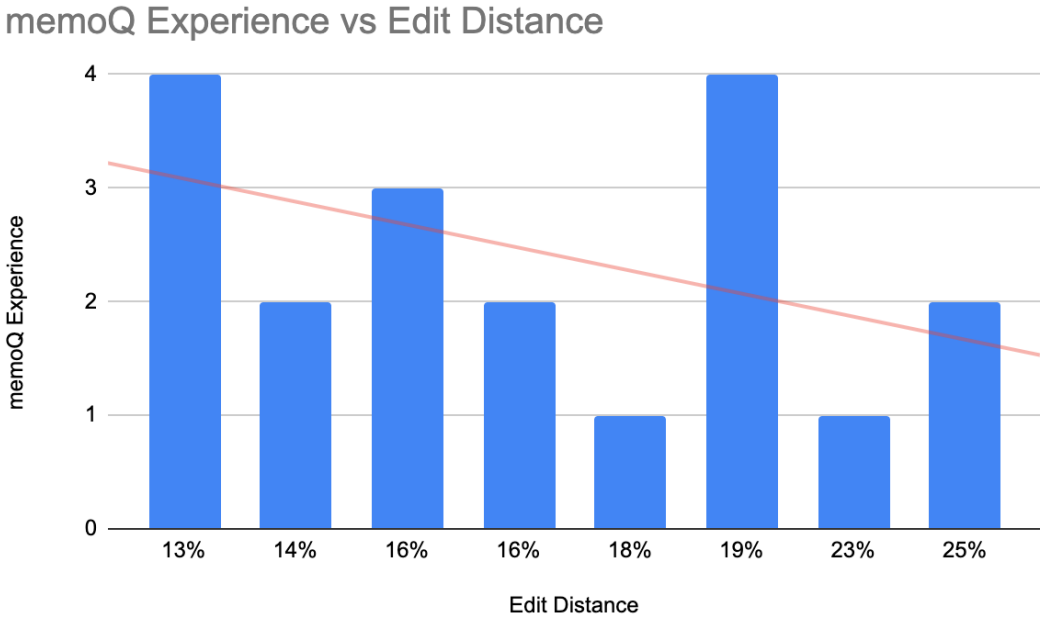

Figure 17. memoQ Experience and edit distance.

The experience with memoQ was divided into 4 groups (10 and more years = 4; 6 years and more = 3, 5 years and less = 2; has not used it often =1). There was a moderate negative correlation towards more experienced memoQ users having a lower edit distance: r ≈ -0.43. There was also a moderate negative correlation towards more experienced memoQ users having a lower editing time: r ≈ -0.55.

6 Discussion

The discussion section interprets the study findings while answering the research questions in the context of related literature. Because of the exploratory design of this study, there were no hypotheses formed to compare with the findings.

6.1 How do Translators Perform the Post-editing Task? (RQ1)

The findings of this research suggest that the participating translators had different approaches to the task of PEMT and they viewed this task in the familiar context of translation and editing. RQ1 is a broad question and in order to answer it more specifically, two subquestions were formed which are discussed in the following two sections 6.1.1 and 6.1.2.

6.1.1 What are the subtasks and in which sequence they occur? (RQ1.1)

Translators mentioned several activities within the PEMT task. The results of the subtasks in this study mostly fall in line with the other research in the field. To view this study's PE subtasks in the context of previous research (refer to section 2.2), they are organized according to the categories of PE-related tasks and processes by Krings (2001) and Rico Pérez and Torrejón (2012: 168-169):

A. Source text-related process includes:
    a. reading the source text before starting to post-edit.
B. MT-related processes include:
    a. reading the MT output,
    b. comparing the MT to the source text,
    c. spotting MT errors,
    d. identifying questionable places in MT (possible errors).
C. Target text production processes include:
    a. correcting MT errors,
    b. leaving comments in memoQ about translation choices,
    c. achieving terminological consistency,
    d. following the instructions to edit up to human quality.
D. Target text evaluation processes include:
    a. final check through reading the target text,
    b. final check through a QA in the CAT-tool.
E. Reference work-related processes include:
    a. using internet search engines (googling),
    b. using web dictionaries,
    c. referring to the car manufacturer homepage,
    d. consulting someone more knowledgeable.
F. Physical writing processes - the amount of edits was measured by edit distance (refer to section 5) and the nature of edits was visible through tracked changes. These

revealed that the translators' preferred strategy was to edit the existing MT. They did not prefer deleting the entire segment to recreate it.

While the subtasks mentioned by the participants are mostly in line with the activities mentioned by Rico Pérez and Torrejón (2012: 168-169), there are some distinctions as well. One of the translators in this study used the opportunity to leave comments in memoQ about the reasons behind their PEMT choices. Depending on the exact workflow, these comments could be directed to an additional editor or project manager who might need to ask for more information from the client. This refers to the need to figure out if and how the commenting function in PEMT should be used. In accordance with the need to make commenting consistent, one translator points out the necessity to regulate the entire field of PEMT better — use standards and quality control. Since PEMT is relatively new, the commercialized use of MT is still rather under-regulated and based on translators' insights in the current study this lack of regulation appears to be at the expense of a translator: editing MT of poor quality and lacking instructions takes time and effort, while it might not be reflected in the remuneration.

According to Rico Pérez and Torrejón (2012), target text evaluation processes are limited to evaluating MT quality. However, both the final check of the post-edited material and performing a QA on it could be a good fit under this category, because it is a step forward from target text production. The final check translators performed in this study was not uniform, some did a more thorough job and some just quickly glimpsed at the text, to see if there were obvious errors. One translator performed a memoQ QA. In the context of this study, the thoroughness of the final check seems to be connected to how much translators edited during the editing segments phase of the task and how confident the translators were with their PE abilities. Two translators who mentioned that they quickly browsed the text for final check said they did substantial editing and carefully read the segments while still in the target text production phase. The only translator who both performed a final check by reading the text and the QA in memoQ happened to be the only translator in this study with no previous editing experience. This translator also expressed strong hesitation about being able to edit the task up to the necessary quality. The minimal instructions of this study task did not include performing a final check, but it seemed to be a matter of course for the translators who took part in this study, since all but one translator mentioned doing it.

Based on how the translators described their work process on the study task, generalized steps of the study task are visualized in a workflow in Figure 18.

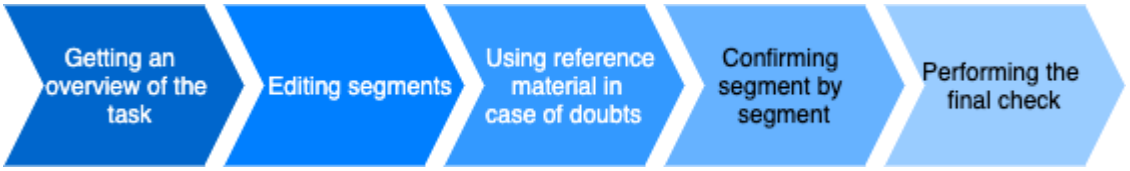

Figure 18. General workflow of the study task.

Getting an overview of the task ranged from reading the source and target through thoroughly to just glancing at the length of the task. Editing segments involved reading the source, then target, comparing them and finding out what needed correcting, and then correcting the MT errors. In case of doubts, translators referred to external help materials. After correcting mistakes, they confirmed the segment. After editing the segments, using reference material when needed, and confirming all the segments, they performed the final check. This means, the second, third, and fourth steps occurred iteratively in the process of PE. In contrast to the approach of confirming segments after editing them, one translator confirmed all segments before starting to post-edit and two translators started PE without confirming the segments after being post-edited. One of these two abandoned the strategy of moving forward with PE without confirming because the segments were not stored in the TM this way. It became clear from the interview that for the translator who abandoned editing without confirming strategy, time was an issue. The same translator also mentioned the need to establish a systematic approach to PE as a lesson learned during the task. The translator who confirmed all the segments at the end of the task explicitly stated time not being of a concern, because they took the necessary time to be thorough and were in no hurry. In the context of this study, this points to the need for the translators to choose an optimal strategy for PE, because time tends to be limited.

6.1.2 How does post-editing compare to translation and editing? (RQ1.2)

Most of the translators taking part of this study did not have much previous experience with PEMT (refer to Table 1). For most of them this study task was the first PE task received from a translation agency. As seen from the findings (refer to 4.1.2), they often compared PE processes and the needed skills to the ones of translating and editing and did not see PE as that much of a task on its own but rather as translating or editing or partly both. It was pointed out that translation with the help of TM and fuzzy matches is effectively editing. During this method of translation combining TM and FMs there usually is still a proportion of no matches to translate from scratch, while MT entirely eliminates translating from scratch. And the wider use of MT amplifies the translation industry's already existing tendency of moving towards editing. When comparing working with FMs and MT, three translators mentioned that editing FMs requires more effort than editing MT. Thus, it seems that for the technical translators taking part in this study, the previous use of TMs and the need to edit FMs in CAT-tools has made the further MT-related shift towards editing more gradual.

In comparison to editing human translation, the translators expressed opinions that machine errors are more unexpected, different from human errors, and easier to spot. An important common denominator appeared to be rather low expectations for MT quality compared to the quality of a professional translator. This made the participants expect the worst and thus to be predisposed to spot errors more easily. Half of the translators compared the MT of this task to the work done by an inexperienced or substandard translator which leads to the conclusion the MT output in this study was not too poor when it is measured against a certain level of human quality. The comparisons of PE to translating and editing the translators mentioned in

this study seem to depend on the quality of the given MT (or TM): a low-quality MT means more effort while a high-quality MT could produce perfect matches with no need to edit.

Although technical translators are accustomed to editing the work of other translators and FMs based on translation memories, the translators taking part in this study would still prefer to stick to translating instead of starting to post-edit. Rico Pérez and Torrejón emphasize their presumption that "PE is a task to be performed by translators" (Rico Pérez and Torrejón, 2012: 167), while some translators in this study expressed their opinion of PE not being a task for a translator but rather for an editor. This view is in accordance with the revision and proofreading skill as the most important skill required from post-editor candidates mentioned by Ginovart Cid (2020). The findings of this study suggest that the participating translators are still shaping their views of PE and whether it is a more suitable task for a translator or an editor.

Based on what subtasks translators completed while doing the study task, they addressed skills and knowledge needed for the successful completion of these subtasks. Ginovart Cid (2020) mentions the need to decide whether to edit MT or translate from scratch and Nitzke et al. (2019) mention the need to decide between choosing full and light PE. The findings of this study suggest decision making as an integral part of being able to finish the PE task of this study: the post-editor is seen as the one who needs to ensure the final quality of the job. Some participating translators seemed to find comfort in the fact that there is an editor correcting their possible mistakes and they do not think of themselves as "editors". This seems to refer to the translator's self perception that Nitzke et al. (2019) pointed out as being important for PE to succeed. Now the translator in the role of a post-editor is responsible for deciding whether the final quality is good enough and if all the translation mistakes were corrected as needed. Two translators who took part in this study also had extensive editing experience and they did not mention the role of being the final decision maker to be an issue. In the context of current study, this seems to point towards the need for translators to acquire an editor's skill of decision making to become successful at PE.

The capacity to post-edit up to human quality and the capacity to identify MT errors was mentioned as two of the top three skills in PE by Ginovart Cid (2020: 178). It is in line with this study, where translators pointed out the post-editor's need to achieve high quality of PE text and take responsibility for correcting MT errors. The third skill in this top three (Ginovart Cid, 2020: 178), the capacity of PE according to PE guidelines, was not explicitly mentioned by any of the participating translators as a requirement of post-editor. However, the importance of instructions accompanying the PE task was addressed in the interviews and it is discussed under section 6.2. Out of three skills that LSPs seek in a post-editor candidate, according to Ginovart Cid (2020), the revision and proofreading skills, and subject field knowledge or specialization (Ginovart Cid, 2020: 179) are represented in the findings of this study. The third skill, CAT-tool knowledge (Ginovart Cid, 2020: 179), was not explicitly mentioned by the participating translators in the context of necessary PE skills, although it was addressed when talking about CAT-tools and it is discussed under section 6.2. Also Rico Pérez and Torrejón (2012) and Nitzke et al. (2019) indicate the importance of instrumental

competence/CAT-tool/MT skills. It can be argued that such instrumental skills grow over years of using the CAT-tools. In this study there appeared to be a relationship between the participating translators' memoQ experience and their editing time: the ones with more memoQ experience post-edited faster. This leads to the potential hypothesis that helping train translators to have a better command of the program could significantly help reduce their PE time. On the other hand, the correlation between translation experience and editing time was weak and pointed in the opposite direction: those having less translation experience used less time for the task in this study.

The task of PEMT is still new for most of the translators in this study and although it contains similar traits to translating and editing human translation, there still are aspects about it that evoke hesitation in participating translators. One such aspect is the higher responsibility that comes with providing final quality. While the PE skills mentioned by the participating translators are in line with previous research, some of the skills from the respective research were not mentioned in the same context, such as CAT-tool skills and the ability to follow instructions. While CAT-tool skills seem to be very important in achieving good PEMT results, they need more attention.

6.2 What are the challenges translators encounter when post-editing machine translation? (RQ2)

The findings of this study propose there was a wide range of challenges that the participating translators faced while completing the study task. The fact that most of the translators are new to PE might have contributed to the numerous difficulties they faced. Thus, these findings could be a useful source of information for preparing translators for the tasks of a post-editor.

Contemporary translation is made up of an increasingly large proportion of interaction with CAT-tools, a tendency that is increasing with the wider use of MT. Good command of CAT-tools is also seen as one of the crucial skills in the skill set of a post-editor (Ginovart Cid, 2020; Rico Pérez and Torrejón, 2012; Nitzke et al., 2019). Therefore, it is important to take a closer look at the problems translators have with CAT-tools both in direct connection to this study task and on a wider scale. In keeping with the results of O'Brien et al. (2017), the findings of this study confirm the complexity and segmentation of the user interface to be problematic for some translators. Removing the automatic capital letters, changing text color, and using subscript and superscript were mentioned either as missing or problematic, although these functionalities all exist in memoQ. One translator explicitly said they had problems with finding the necessary settings. Contributing to the complexity, it was mentioned that QA in memoQ is particularly inconvenient to conduct, because it is not taking the multitude of grammatical cases in the Estonian language into consideration and is thus giving several false errors. It seems that some of the reported issues with memoQ in this study were related to a lack of knowledge concerning some functions and settings. The problem of the text being harder to follow/segmentation can be linked to how the participants

perceived memoQ's layout and to possibly yet another important aspect that revealed itself in three interviews: the preferences of participating translators are often determined by what they are used to.

A large portion of the findings in the challenges section 4.2 are categorized as cognitive challenges. In a similar way to Bundgaard et al. (2016), entrapment by the suggestions of MT was mentioned by translators as MT interfering with their train of thought, causing a mental block that resulted with them sticking to the MT instead of trying to form a translation themselves, and MT not happening by means of the free thought. Such interference points to the very core of these translators' thought process: in the case of PEMT, the way they shape the translation in their mind gets filtered or even hindered by MT. This could, in a way, alienate them from the results of their thought process and cause them to lose their agency. According to Ehresberger-Dow and Massey (2017) the aspect of agency is more decisive in the successful socio-technical deployment of MT than the technological aspects.

Some translators found it difficult to balance between personal preferences and what really needed editing in the MT as well as not having the option to discuss translation choices with a colleague in the traditional, reciprocal translator-editor role. These insights are related to the decision-making skill which was discussed in section 6.1.2. It is interesting that according to the possible relationship between perceived cognitive effort and editing time in this study, more cognitive effort does not appear to indicate higher temporal effort but the other way around: the translators who perceived a higher rate of cognitive effort took less time to complete the task. This could be connected to the time pressure that several translators mentioned which did not allow them to take as much time for the task as they would have needed for less-strenuous decision making. When comparing the translators' PEMT experience with their edit distance, a strong negative correlation was found, which shows a possible relationship between translators having experience with PEMT projects from LSPs and making fewer edits. From this and within the limits of this study, it appears as if the translators with PEMT experience see less need to make changes to the MT when compared to translators new to PE.

In addition to cognitive and instrumental difficulties, the challenges surrounding PE were also addressed. One general problem was the lack of time. In connection to the discussion in sections 6.1.1 and 6.1.2, the lack of time indicates the need for the participating translators to adopt the most suitable PE strategy and acquire key skills for PE. It might also make sense to inspect one's hardware, so that this would not become an obstacle in the PE process. Physical tools were not particularly addressed in this study, while it became evident during the interviews that at least half of the translators used only one screen when performing this study task. In addition, having an old computer or an older version of memoQ were mentioned. Time is on one's side only when creating suitable conditions for it.

Nitzke et al. (2019) mentioned the importance of a PE brief as one of the factors affecting success in PE. The instructions of this study task were to edit up to human quality, but did not specify how to achieve this target or what the purpose or target group of the text are, nor was

any reference to the original file added. Translators in this study would prefer more precise instructions and to have the reference material provided. The need for more precise instructions in this study also points towards decision making — when the quality expectations are made clear for the translator, it could partly relieve their burden of deciding. Better technical preparation of a PE task (appropriate settings adjusted in the CAT-tool and the document formatted in the correct way) as preferred by some participating translators would remove part of the complexity of using a CAT-tool. These improvements could streamline the PE for a post-editor.

Findings about challenges in PE suggest the need for a broad-based approach to MT-related problems. It appears that a high-quality MT engine will be of little help if the translator editing its output feels they are unable to decide on what to correct. Even worse, the translator might feel a mental block and thereby lose their agency.

6.3 What are translators' attitudes towards machine translation? (RQ3)

According to Cadwell et al. (2018), the rise of NMT will make the wider use of MT inevitable. The same thought was expressed by the translators who took part in this study. An important attitude-shaping aspect is how translators feel affected by this wider proliferation of MT, which was asked in the interviews. This tendency was seen as a rather positive one in the context of personal MT use, however, the translators had a more sceptical stance towards professional MT use. Some of them mentioned the pressure to decrease translation prices as a problem and some were concerned by translators losing a portion of their work or fewer translators being needed in the future. Others were confident that MT is not able to substitute humans and good translators will not run out of work as a result of MT. In the context of comparing human work to MT output, the translators mentioned to have more trust towards human translation than MT. This is contrary to the findings of Scansani et al. (2019) who concluded no difference in trust towards human translation and MT.

Some of the limitations of MT mentioned by translators in this study were similar to the reasons that translators gave to Cadwell et al. (2018) for not using MT. For example, translators in this study also saw MT as being ineffective for some text types (literature, poetry, legal, marketing texts) and problems based on the language pair (MT sentence structure sounding foreign when source language is Germanic and target Finno-Ugric). While translators in this study mentioned MT to be constantly developing and getting better in some areas (sentence structure and terminology was mentioned here), the machine was said to remain limited in understanding context. Related to this limitation, certain creative text types with a particular richness of subtext, metaphors, and cultural background are seen to persist for humans to translate. This also gives translators confidence in their trade remaining in the hands of human translators. At the same time, translators in this study saw technical texts, short sentences, and unambiguous terminology as something that the machine handles rather

well. According to previously mentioned perceptions, MT affects technical translators more than legal translators, translators of fiction, or translators of marketing texts.

Google Translate was extensively discussed in the interviews and one translator expressed their opinion of how GT shapes the opinion about MT in general, and not in a positive way. Given GT really appears to be the most notable MT based on the interviews of this study. As such, it has the potential to create general prejudice about MT, including custom MT solutions like the one used for this study. However, translators mentioned that custom MT solutions perform better in the professional context and GT was seen as more useful for private purposes. Translators' attitudes were notably positive about using MT for private purposes, while the professional use caused scepticism. To the best of the knowledge of the author, this difference between translators' opinions on professional and private use of MT has not been addressed in previous research. The willingness to use MT for work was expressed, as long as the translator would know more about how it works. This relates to what O'Brien (2012) said about MT being perceived as a black box out of a translator's reach to know or affect and thus creating distrust.

The answers given about the usefulness of MT show that translators in this study are quite willing and open to using MT when it comes to repetitive tasks like producing the same standard sentences in a safety manual. In line with Cadwell et al. (2018), translators mentioned MT to increase productivity while doing more repetitive tasks. This leaves translators an opportunity to focus on the real difficulties in the translation which is in line with Yngve (1954: 21). As an analogy, the machine could be seen as a decorator here, as the translator is an artist. Some translators saw the usefulness of MT on a global scale as an aid in research or overcoming the challenges of the COVID-19 crisis. These insights reflect a positive attitude towards MT, although in a disconnected way — translators who mentioned these perspectives for MT would personally still prefer translation over PE. In contrast to these optimistic opinions, one translator pointed out a wider problem: MT multiplies and amplifies language of poorer quality which in turn will negatively influence language.

Translators' opinions and insights about MT expressed in this study touched a wide variety of topics:

A. limitations of MT were perceived to be first and foremost connected to the inability of understanding context (which also could give a feeling of job security for translators who expressed this thought);
B. seeing changes in the translation market in an unfavorable direction for translators;
C. MT performs fine on technical texts; a custom MT solution is better than a generic one;
D. using MT for personal purposes was seen positively;
E. reservations regarding the professional use of MT, concerning the quality of MT and how it might affect translator professionals' working conditions.

In the context of this study, translators' attitudes about MT reveal concerns that are not limited to MT performance. These shared insights suggest the need for a well-planned socio-technical deployment of MT.

6.4 Implications for Research

As this study is exploratory, it should give ideas and possible hypotheses for future research about PEMT. Followingly, there are a few ideas that might be worth taking from this thesis to be further studied.

This research only brushed the surface of which devices translators use. Some translators suspected that their equipment might be holding them back and many used just one monitor during the study task. It would be interesting to test out the hypothesis of better equipment resulting in less PE effort and if so, then how significant the role of equipment is.

The findings of this research showed a moderate correlation between more memoQ experience resulting in shorter editing time. The same relationship did not reveal itself in comparison of translation experience and editing time. It would be interesting to see the results of testing out a hypothesis of translation software knowledge affecting editing time more than translation knowledge in a more experimental setting with more participants.

One aspect addressed in RQ3 was trust. A comment made by one of the translators about trusting a MT engine with quality guaranteed by some authority institution gave an idea for a research design to see how big of a role this would play in trusting MT. Partly this design is inspired by Scansani et al. (2019) who addressed trust towards MT when the same translation was given to two groups to edit; one group was told it is MT and another one that it is a human translation. They did not find there to be any difference in trust based on edit rate or editing speed (Scansani et al., 2019). The suggestion would be to give the same MT output to two groups for post-editing where one receives information about this MT engine fulfilling all quality requirements evaluated by some respective institution, while the other group will receive just the MT output. In addition to editing speed and edit rate, some opinions could also get collected. This design would also help to see how big of a role prejudice plays in assessing MT quality. In the present study, the findings pointed to there being prejudice towards MT because of unfavorable opinions on GT, whereas the opinion of custom MT solutions is much better. Now, if two groups were given the same MT output and one group was told it was the result of custom MT complying to high quality standards and another group was told it is the product of a generic MT, would there be a difference in how the translators treat it?

6.5. Implications for Practice

The findings of this case suggest that accurately describing the job of a post-editor is important, so that the participating individuals are able to preserve their agency. In addition to

the job description, precise task instructions seem to be relevant to help translators overcome the cognitive challenges of not being able to decide between their own translation preferences and those suggested by MT, or resisting the urge to make corrections just for the sake of it. Providing post-editors training on MT could help with some of the possible cognitive difficulties, such as challenges related to the responsibility of providing final quality, not being used to the task, or not being able to determine MT errors very easily. In addition to cognitive difficulties, several challenges faced during the task were due to not being fully aware of the functionalities of the CAT-tool. Supportive evidence of comparing editing time with memoQ experience revealed a relationship of more experience resulting in decrease in edit time. Thus, taking part in training about the use of CAT-tools could also be helpful. Among other measures, a fair remuneration policy for a successful socio-technical deployment of MT would be necessary.

This study did not go into details in analyzing the content of edits that translators made in MT output. It would be useful to study them to be able to find out how translators exactly interpret what constitutes a necessary change, and see these results in a wider framework of client requirements and editors' opinions.

Based on the results of comparing technical equipment used in translation (suggested for research in 6.4), the practitioners might see how acquiring a new laptop or a second screen is beneficial for overall work results.

6.6 Limitations

The goal of this research was to explore how translators approach the task of PE, what challenges they face during PE, and what their attitude towards MT is. In order to accomplish this, the case study method was chosen. Some of the limitations of the design of this study were already addressed earlier in the study, such as the challenge of finding suitable participants and limitations of direct observation due to COVID-19. This section includes these and other limitations of this study known to the researcher.

Firstly, it is possible this study is unrepresentative due to the small number of participating translators and no generalizations about all translators can be made based on it. The limitation due to the small number of participants was mitigated by involving translators with varying rates of professional experience (refer to 3.2 for more details) and using data triangulation to achieve data saturation.

Secondly, there was no direct observation of participants while they performed the PE task. Due to this limitation, it was not possible to collect first-hand data about the work process. To get around this constraint, the interview questions addressed the subtasks in a detailed manner.

Thirdly, there was an interval of up to two weeks between the time when translators performed the PE task and gave the interviews. Such an arrangement might have allowed participants to forget about the details of the task. This problem was addressed by making the task materials available to translators before the interview and for the duration of the interview, so they could refer back to what they did.

Fourthly, translators needed to set up editing time reporting in their memoQ. This was anticipated by the researcher to possibly cause difficulties for the translators, since this might not be a usual task for translators. To mitigate the risk of failing to measure time spent on the task, two measures were taken: instructions with the necessary steps were handed out and for redundancy the translators were told to keep track of their starting and ending times using a timepiece. The second measure proved to be necessary with two translators.

Fifthly, the coding scheme used in thematic analysis was not validated by a second researcher, which can cause researcher bias. There was an attempt to mitigate this risk by making careful considerations about the process of thematic analysis and discussing the coding process with one of the supervisors of this study.

Sixthly, it was difficult to compare data from different sources. This can be too subjective and lead to misinterpretation of the results. To mitigate this risk, the author followed the methods used in the literature about measuring relevant experience in years or binary options of either an existing or non-existing experience.

Lastly, two translators holding strongly negative opinions about MT refused to take part in the study. This might show the translators' attitude in the context of this study to be more positive as they would have been with these additional participants. However, it is not a generalizable study and it would probably not prove to be effective to conduct interviews with individuals who do not wish to contribute to the advancement of MT research.

7 Conclusion

While MT has proven to be a useful tool for translating technical texts, careful consideration of certain aspects of its implementation seems to be necessary. One significant obstacle to the implementation of MT is the continuing resistance from translators to post-edit MT output. The author conducted an exploratory case study in order to provide insight for future research (section 6.4) and translation agencies who wish to implement MT in their workflow (section 6.5). The purpose of the study was to attempt to find out more about how translators approach PE, challenges in PE, and their attitudes about MT. RQ1.1 addressed the main subtasks and workflow within the PE task assigned to translators for this study and was answered in section 6.1.1. RQ1.2 asked how PE compares to translation and editing and was discussed in section 6.1.2. RQ2 was asked to try to find out more about challenges the translators of this study had with PE. The analysis of the challenging aspects of the PE task were presented in section 6.2. Finally, RQ3 asked about translators' attitudes on MT and was addressed in section 6.3.

Through the assignment of a PE task to translators in the CAT-tool memoQ and analyzing the subsequent interviews conducted as a part of this study, it appears that most of the participants used a particular sequence to complete the PE subtasks and the data indicated that certain deviations from this workflow could have potentially cost translators time. The study also showed that there was some ambiguity among the participants when it came to describing post-editor requirements: which skills are most important for a post-editor to possess and whether the role of a post-editor has more in common with the role of a translator or that of an editor. Challenges faced by translators in this study were categorized as cognitive, CAT-tool-related, and other challenges related to the PE task. While cognitive challenges could be interpreted as being related to translators' agency and decision making, CAT-tool-related challenges seemed to have a relationship with a lack of knowledge about the CAT-tool.

Based on the findings of this study, the translators' opinions about MT use for private purposes appeared to be positive in comparison to a more sceptical stance towards its use in a professional setting. The limitations of MT being unable to accurately translate within the necessary context or translate certain types of text to match human quality, mentioned by the participants, do not seem to diminish how these translators view it as being more capable as it applies to technical translation. In order for the participating translators to be willing to use MT in their professional work, necessary MT quality is crucial. MT is constantly being developed and therefore it is perceived by the participants of this study to be moving towards the necessary quality for professional use. While the quality of MT is perceived to be improving in general, custom MT solutions are seen to have a better quality than generic ones, which are viewed as being more suitable for personal use. As a relevant reason for scepticism towards the professional use of MT, PE pricing as it relates to translator remuneration was mentioned.

Based on the data collected in this study, suggestions were made in section 6.5 that could be helpful with the organizational change that coincides with the implementation of MT. These suggestions include providing a precise job description for post-editors, detailed instructions of the PE task, in addition to training translators on the use of CAT-tools, on working with MT output, and on the theory behind how MT works. The implications for future research in section 6.4 suggested to conduct research on translators' technical equipment, relevant experience in the use of CAT-tools in relation to PE time, and a possible research design for addressing the issue of trust in PEMT. The author of this research contends that both practitioners and researchers in the field can benefit from these suggestions.

References

Braun, V., & Clarke, V. (2006). Using thematic analysis in psychology. *Qualitative research in psychology, 3*(2), 77-101.

Bundgaard, K., Christensen, T. P., & Schjoldager, A. (2016). Translator-computer interaction in action: An observational process study of computer-aided translation. *Journal of Specialised Translation*, (25), 106-130.

Cadwell, P., o'Brien, S., & Teixeira, C. S. (2018). Resistance and accommodation: factors for the (non-) adoption of machine translation among professional translators. *Perspectives, 26*(3), 301-321.

Carl, M., Dragsted, B., Elming, J., Hardt, D. and Jakobsen, A.L., 2011. The process of post-editing: A pilot study. *Copenhagen Studies in Language*, *41*, pp.131-142.

Castilho, S., Moorkens, J., Gaspari, F., Sennrich, R., Sosoni, V., Georgakopoulou, P., ... & Gialama, M. (2017). A comparative quality evaluation of PBSMT and NMT using professional translators.

Chandra, Y., & Shang, L. (2019). Inductive coding. In *Qualitative research using R: A systematic approach* (pp. 91-106). Springer, Singapore.

Cid, C. G., Colominas, C., & Oliver, A. (2020). Language industry views on the profile of the post-editor. *Translation Spaces.*

De Almeida, G. (2013). *Translating the post-editor: an investigation of post-editing changes and correlations with professional experience across two Romance languages* (Doctoral dissertation, Dublin City University).

DeepAI: The front page of A.I. (2019). https://deepai.org/machine-learning-glossary-and-terms/neural-machine-translation. *Last accessed 12.05.2021.*

Dix A. (2009) Human-Computer Interaction. In: LIU L., ÖZSU M.T. (eds) *Encyclopedia of Database Systems. Springer, Boston, MA*. https://doi.org/10.1007/978-0-387-39940-9_192. *Last accessed 12.05.2021.*

Ehrensberger-Dow, M., & Massey, G. (2017). Socio-technical issues in professional translation practice. *Translation Spaces, 6*(1), 104-121.

Fusch, P. I., & Ness, L. R. (2015). Are we there yet? Data saturation in qualitative research. *The qualitative report*, *20*(9), 1408.

Gambier, Y. (2009). Competences for professional translators, experts in multilingual and multimedia communication. *Mode of access*: https://ec.europa.eu/info/sites/default/files/emt_competences_translators_en.pdf. *Last accessed 13.05.2021.*

Garabık, R. (2006, November). Slovak morphology analyzer based on Levenshtein edit operations. *In 1 st Workshop on Intelligent and Knowledge oriented Technologies* (p. 2).

Guerberof Arenas, A. (2008). Productivity and quality in the post-editing of outputs from translation memories and machine translation. *Localisation Focus The International Journal of Localisation, 7*(1), 11-21.

Guerberof Arenas, A. (2013). What do professional translators think about post-editing. *JoSTrans The journal of specialised translation*, *19*, 75-95.

Hutchins, W. J. (1995). Machine translation: A brief history. In *Concise history of the language sciences* (pp. 431-445). Pergamon.

Karamanis, N., Luz, S., & Doherty, G. (2011). Translation practice in the workplace: contextual analysis and implications for machine translation. *Machine Translation, 25*(1), 35-52.

Koehn, P., & Knowles, R. (2017). Six challenges for neural machine translation. *arXiv preprint arXiv:1706.03872.*

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

Lazar, J., Feng, J. H., & Hochheiser, H. (2017). *Research methods in human-computer interaction*. Morgan Kaufmann.

Läubli, Samuel; Fishel, Mark; Massey, Gary; Ehrensberger-Dow, Maureen; Volk, Martin (2013). Assessing post-editing efficiency in a realistic translation environment. In: *MT Summit XIV Workshop on Post-editing Technology and Practice*, Nice, 2 September 2013 - 2 September 2013, 83-91.

Leavitt, H. J. (1965). Applied organizational change in industry, structural, technological and humanistic approaches. *Handbook of organizations, 264*.

Maguire, M., & Delahunt, B. (2017). Doing a thematic analysis: A practical, step-by-step guide for learning and teaching scholars. *All Ireland Journal of Higher Education, 9(*3).

Mumford, E. (2000). A socio-technical approach to systems design. *Requirements Engineering, 5*(2), 125-133.

memoQdocs: Create edit distance report. https://docs.memoq.com/current/en/Places/create-edit-distance-report.html. *Last accessed 12.05.2021*.

memoQdocs: Create editing time report. https://docs.memoq.com/current/en/Places/create-editing-time-report.html. *Last accessed 12.05.2021.*

Nitzke, J., Hansen-Schirra, S., & Canfora, C. (2019). Risk management and post-editing competence. *The Journal of Specialised Translation*, 31, 239-259.

O'Brien, S. (2006). Pauses as indicators of cognitive effort in post-editing machine translation output. *Across Languages and Cultures, 7*(1), 1-21.

O'Brien, S. (2012). Translation as human–computer interaction. *Translation Spaces, 1*(1), 101-122.

O'Brien, S., Ehrensberger-Dow, M., Hasler, M., & Connolly, M. (2017). Irritating CAT tool features that matter to translators. *Hermes: Journal of Language and Communication in Business, 56*, 145-162.

Olohan, M. (2011). Translators and translation technology: The dance of agency. *Translation Studies, 4*(3), 342-357.

Plitt, M., & Masselot, F. (2010). A productivity test of statistical machine translation post-editing in a typical localisation context. *The Prague bulletin of mathematical linguistics, 93*(1), 7-16.

Pym, A. (2019). How automation through neural machine translation might change the skill sets of translators. *Retrieved from usuaris. tinet. cat/apym/on-line/translation.*

Rico Pérez, C., & Torrejón, E. (2012). Skills and Profile of the New Role of the Translator as MT Post-editor. *Tradumàtica*, (10), 0166-178.

Ruokonen, M., & Koskinen, K. (2017). Dancing with technology: translators' narratives on the dance of human and machinic agency in translation work. *The Translator, 23*(3), 310-323.

Saldaña, J. (2021). *The coding manual for qualitative researchers*. sage.

Sakamoto, A. (2019). Why do many translators resist post-editing? A sociological analysis using Bourdieu's concepts. *The Journal of Specialised Translation*, 31, 201-216.

Scansani, R., Bernardini, S., Ferraresi, A. and Bentivogli, L., 2019. Do translator trainees trust machine translation? An experiment on post-editing and revision. In *Proceedings of Machine Translation Summit XVII Volume 2: Translator, Project and User Tracks* (pp. 73-79).

SDL Producthelp: SAE J2450 QA Model http://producthelp.sdl.com/SDL_TMS_2011/en/Creating_and_Maintaining_Organizations/Managing_QA_Models/SAE_J2450_QA_Model.htm. *Last accessed 12.05.2021.*

Shterionov, D., Superbo, R., Nagle, P., Casanellas, L., O'dowd, T., & Way, A. (2018). Human versus automatic quality evaluation of NMT and PBSMT. *Machine Translation, 32*(3), 217-235.

Sotiriadou, P., Brouwers, J., & Le, T. A. (2014). Choosing a qualitative data analysis tool: A comparison of NVivo and Leximancer. *Annals of Leisure Research, 17*(2), 218-234.

Sreelekha, S. (2017). Statistical vs rule based machine translation; a case study on indian language perspective. *arXiv preprint arXiv, 1708.*

Tatsumi, M. (2010). *Post-editing machine translated text in a commercial setting: Observation and statistical analysis* (Doctoral dissertation, Dublin City University).

Trist, E. L. (1981). *The evolution of socio-technical systems* (Vol. 2). Toronto: Ontario Quality of Working Life Centre.

Yin, R. K. (2003). *Case study research: Design and methods* (Vol. 5).

Yngve, V. H. (1954). The machine and the man. *Mech. Transl. Comput. Linguistics, 1*(2), 20-22.

Zaretskaya, A. (2017). Translators' requirements for translation technologies: user study on translation tools.

Appendices

Additional material related to this thesis can be found below.

I Interview Questions

Topics	Questions
1 Translator's background	*Intro: Let's start with some questions about your experience as a translator.* 1.1 How many years have you worked as a translator? Editor? (general experience) 1.2 Have you worked with technical translations in the field of the automobile industry? If, what share of your translations approximately are technical translations of the automobile industry? What do you estimate your translation speed to be (pages per hour)? (domain experience) 1.3 What is your experience with post-editing? Years, estimated no. or projects. (post-editing experience) 1.4 Do you work as a full-time translator/editor or part-time? If part-time, what is your main profession? (full- or part-time) 1.5 What is your educational background? Have you received translator/editor/post-editing training? (education and training)
2 Post-editing task in detail (environment, tools, subtasks, comparison to translating in the same environment)	*Intro: Thank you for completing the task of post-editing MT earlier today.* **2 Please describe your activities during the post-editing task.** [probes: Did you edit the text segment by segment? Did you familiarize yourself with the target and or source text before starting to make edits in segments? Did you read the text through once PE was finished? Did you leave the PE environment for term or context searches? What was different compared to translating in memoQ?] **2.1 What is your opinion on memoQ as the environment for post-editing?** [probes: Have you used other tools for post-editing? If so, how do they compare? Which functionalities were useful/lacking? How about the overview of the text/flow in post-editing?]
3 Challenges during the task (difficulties, effort, examples of problems)	*Intro: Let's focus on difficulties in post-editing now. Please continue thinking about the same task you completed for this study, but if you have examples from other PE experiences, please refer to these as well.* **3 Please share your opinion(s) on the PE task. How was it?** [probes: How would you assess the quality of given MT text? How

	were the instructions?] **3.1 How would you assess the effort it took to post-edit this text?** [probes: What makes more sense here: post-editing or translating from scratch? Why? How was it to spot the MT mistakes (compared to editing human translation)? Would you accept MT with this project as the main method of translation?] **3.2 Do you have any suggestions on how to improve post-editing experience/overcome difficulties for translators?** [probes: How to improve the tool? How to improve the instructions?]
4 Machine Translation	*Intro*: *With the rise of neural machine translation, machine translation is becoming used more widely than ever before.* **4 Please share your thoughts about the increasing usage of machine translation.** [probes: What is your opinion on translation becoming widely accessible thanks to MT platforms like Google Translate? Why do you think MT is useful/useless/…? Which are the translation tasks MT is not able to do?] **4.1 How is the usage of MT in language service providers changing the industry?** [probes: What is your opinion on this change? How do you see this change affecting you?] **4.2 Please tell me about your use of MT.** [probes: For work or private reasons? How often? Which MT engines? How does it work out for you? Would you use MT in your everyday work?] **4.3 How have you gained knowledge about MT?** [probes: If you have taken part in any training or online course: how were they? If you have just read about MT on the internet/forums: how

II Lower level codes

Followingly, the lower level codes are presented structured in subtheme tables with the reference to the higher level codes they fall under and how many translators mentioned the respective topic.

No.	**4.1.1 Subtasks and Workflow** **Code text**	Higher code	Mentioned by
1	Preferred the flowing text in the beginning	3	1
2	Preference to have MT suggestions as reference, not pretranslation	2	2
3	PEMT task took time and effort	3	6
4	Examples of issues and improvements	3	4
5	Googling during the task	2	6
6	Using web dictionaries	2	2
7	Car Manufacturer home page as reference material	2	3
8	Communicating with the client	3	1
9	Confirming segment by segment	1	6
10	Different strategies in subtasks sequence	1	5
11	Need to post-edit systematically	1	1
12	Consulting someone	3	1
13	PE segment by segment without confirming segments	1	2
14	Need for time to get to know the task first	3	5
15	Performing final check	3	8
16	A lot of correcting	3	2
17	Difficulty to name PEMT task price	6	2
18	Preference of unit price over hourly rate	6	1
19	Remuneration of PEMT should reflect comparing to source language	6	2

20	MT usage needs more regulation and standardization	6	1

No.	**4.1.2 Post-editing in Comparison to Translation and Editing** **Code text**	Higher code	Mentioned by
21	Discussion between translator and editor goes missing with MT	4	1
22	Editing MT and FM is similar	4	2
23	PEMT of this task similar to editing a translation of inexperienced and not a very good translator	4	4
24	Preference of translating to PEMT	4	5
25	Translating includes a lot of editing	4	4
26	The use of translation memory	4	4
27	PEMT more stressful than translating	4	2
28	Human errors vs machine errors	4	7
29	PEMT of this task similar to editing human translation	4	4
30	PEMT vs editing human translation - need to compare the source text	4	2
31	PEMT work process different from translation where no text on right	4	1
32	PEMT in memoQ similar to translating in memoQ	4	3
33	Post-editors need to have command of source language	5	1
34	Post-editors need to have knowledge in the subject field	5	1
35	PEMT requires both editor and translator to be very good at what they do	5	2
36	PEMT requires editor experience	5	2
37	PEMT requires translator approach	5	2

38	Translators need to learn to use MT	5	2
39	Border between translator and editor is hazy	5	1
40	Editor as the last instance	5	3
41	Post-editor needs to ensure the terminological consistency of the project	5	2
42	Post-editor needs to feel responsible for letting MT errors slip	5	3
43	Translator and personal preferences	5	2
44	Goal to reach high language quality	5	2

No.	**4.2.1 Cognitive Difficulties in Post-editing** **Code text**	**Higher code**	**Mentioned by**
45	PEMT task - medium effort, not too difficult, not too easy	7	3
46	PEMT task had a good flow and was understandable	7	1
47	PEMT is easier than translating or editing FM	7	4

No.	**4.2.2 CAT-tool Related Issues** **Code text**	**Higher code**	**Mentioned by**
48	memoQ as a good environment for PEMT	8	6
49	memoQ features	8	5
50	Memsource for PEMT	8	3
51	Preference of CAT-tools based on habit	8	3
52	Preference of key shortcuts over buttons	8	1
53	Preference of memoQ	8	2
54	Preference of Trados over memoQ	8	5
55	Markups and tags are a hassle in memoQ and other CAT-tools	8	1
56	memoQ problems and limitations	8	8
57	memoQ QA inconvenient	8	2

58	Need for IT-support	8	1

No.	**4.2.3 Problems Surrounding the Post-editing Task** **Code text**	Higher code	Mentioned by
59	Lack of time (turned into a higher-level code)	9	4
60	Locking terms	11	1
61	Lack of PEMT instructions	10	1
62	Light PEMT instructions are unclear	10	2
63	Need for original file as reference	11	4
64	No purpose of use or target group mentioned in instructions	10	3
65	Preference of written instructions over seminars	10	1
66	Tasks that translator should not do	11	2

No.	**4.3.1 Limitations of Machine Translation** **Code text**	Higher code	Mentioned by
67	Sceptical about professional use	14	5
68	Google Translate has doubtful quality	12	6
69	MT no substitute to human — mistrust and scepticism	14	3
70	MT cannot handle context	13	3
71	MT does not handle literature	13	6
72	MT does not handle long sentences well	13	1
73	MT does not handle metaphors and proverbs	13	1
74	MT does not take client terminology or terminological consistency into account	13	3
75	MT is not able to make text user friendly and marketable	13	1
76	MT needs human editing	14	3
77	MT cannot handle sentence structure	13	6

78	MT does not handle legal texts well	13	2
79	MT quality of this task is bearable	12	4
80	MT quality of this task is poor	12	1
81	MT quality usually not good enough	12	4
82	MT has no human touch or feeling	13	1
83	MT cannot handle word structure	13	1
84	Would trust human more than MT	14	2
85	MT makes cautious and raises doubts	14	5

No.	**4.3.2 Capabilities of Machine Translation** **Code text**	Higher code	Mentioned by
86	MT handles sentence structure	16	1
87	Google Translate can convey the meaning	16	2
88	Custom MT engine superior to generic MT	16	3
89	MT is constantly developing	17	4
90	MT handles grammar well	16	1
91	MT handles short sentences well	16	2
92	MT handles technical translation well	16	3
93	MT performing better than a translator	17	3
94	MT quality of this task is good (turned into a higher-level code)	15	5
95	MT can handle medical texts	16	1
96	MT good enough for internal use	17	1
97	Info sources about MT	17	7

No.	**4.4.1 Usefulness of Machine Translation** **Code text**	Higher code	Mentioned by

98	Optimistic about professional use of MT	19	4
99	Enjoying PEMT task	19	1
100	Google Translate as first choice for personal use	19	3
101	Does not feel affected by MT spreading	19	1
102	MT helps technological development and understanding of human brain	19	1
103	MT helps the globalization of the world in COVID crisis	19	1
104	MT will be used more widely	19	4
105	MT for personal use	19	5
106	MT good for eliminating repetitive legwork	19	5
107	MT helps to save time and effort	19	6
108	MT simplifies life and communication	19	6
109	MT via a third language	19	3
110	Settling with MT when busy	19	2
111	Thoughts and examples of usefulness of MT	19	8
112	MT quality is important for professional use	19	4

No.	**4.4.2 Harmfulness of Machine Translation** **Code text**	**Higher code**	**Mentioned by**
113	MT creates price pressure	18	6
114	Feels disturbed when offered PEMT tasks	18	2
115	Light PEMT creates work dissatisfaction	18	1
116	Low quality MT at the expense of the translator	18	1
117	MT affecting interpretation more than written translation	18	1
118	MT takes away work from translators	18	6
119	MT influences language	18	1

III Litsents

Lihtlitsents lõputöö reprodutseerimiseks ja üldsusele kättesaadavaks tegemiseks

Mina, **Katrin Shuyler**,

1. annan Tartu Ülikoolile tasuta loa (lihtlitsentsi) minu loodud teose
"**A Case Study on Post-editing Machine Translation: Tasks, Challenges, and Attitudes**",

mille juhendajad on **Alexander Nolte** ja **Mark Fišel**,

reprodutseerimiseks eesmärgiga seda säilitada, sealhulgas lisada digitaalarhiivi DSpace kuni autoriõiguse kehtivuse lõppemiseni.

2. Annan Tartu Ülikoolile loa teha punktis 1 nimetatud teos üldsusele kättesaadavaks Tartu Ülikooli veebikeskkonna, sealhulgas digitaalarhiivi DSpace kaudu Creative Commonsi litsentsiga CC BY NC ND 3.0, mis lubab autorile viidates teost reprodutseerida, levitada ja üldsusele suunata ning keelab luua tuletatud teost ja kasutada teost ärieesmärgil, kuni autoriõiguse kehtivuse lõppemiseni.
3. Olen teadlik, et punktides 1 ja 2 nimetatud õigused jäävad alles ka autorile.
4. Kinnitan, et lihtlitsentsi andmisega ei riku ma teiste isikute intellektuaalomandi ega isikuandmete kaitse õigusaktidest tulenevaid õigusi.

Katrin Shuyler
14.05.2021