

## **Decisions in projects using machine translation and post-editing – an interview study**

**Jean Nitzke**, University of Agder, Kristiansand

**Carmen Canfora**, University of Mainz, Germersheim

**Silvia Hansen-Schirra**, University of Mainz, Germersheim

**Dimitrios Kapnas**, University of Mainz, Germersheim

### **ABSTRACT**

Machine translation (MT) and post-editing (PE) have become increasingly important in the professional language industry in recent years. However, not every translation job is suitable for MT and there are many options for carrying out translation/post-editing projects, e.g. no PE, light PE, full PE, full PE plus revision or translation without MT assistance. In 2019, we published a decision tree for post-editing projects (Nitzke et al. 2019) that aimed to take all considerations into account and guide the stakeholders in charge of deciding whether a job is suitable for MT and PE and, if so, what kind of quality assurance might lead to fit-for-purpose translations.

To empirically test our decision tree model now, we developed a semi-structured interview with 21 questions and a scoring task addressing stakeholders who work with MT projects and have to make the decisions which are essential to our model. The interview was carried out with 19 interview partners. In the article, we discuss the interviews' findings against the background of our model. Further, we present qualitative findings on strategic decisions, risk considerations, as well as the value of translation, working conditions and job profiles. Finally, we present our revised model motivated by the empirical findings.

### **KEYWORDS**

machine translation, post-editing, language industry, interview study, best practice, risk management, quality assurance, decision tree model, strategic decisions, job profiles.

## **1. Introduction**

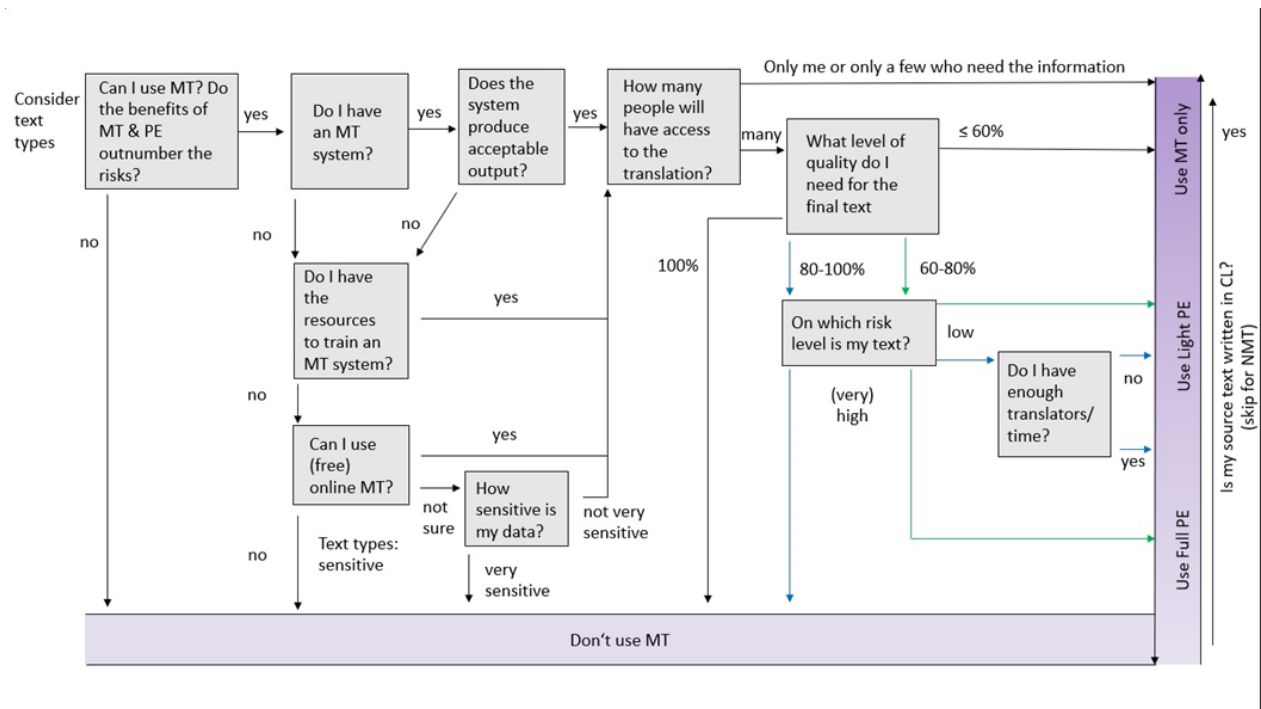
The increased usage of machine translation (MT) and post-editing (PE) in professional translation workflows has led to extensive discussions both in academia and in the language industry (e.g. Elia et al. 2022; Olohan 2021; Slator 2019). Different aspects have to be taken into account to decide for or against MT and PE, and for some translation projects the use of MT and PE might be immediately disregarded. Further, the continuum of options for translation/PE projects is broad: no PE, light PE, full PE, full PE plus revision, or translation without MT assistance.

A precise picture of how prevalent MT and PE have become in the language industry is not evident. For example, Elia et al. note in their 2022 annual report that human translation still accounts for the majority of services in European language service provision. However, PE is rated as the second most important service with the most potential for growth. In the same year, the Nimdzi (2022) report found that MTPE services are offered by 79% of LSPs worldwide.

At the same time, the role of the translators and post-editors, as well as the suitability of MT and PE for certain texts are discussed. Sánchez-Gijón

et al. (2019) recently showed that more and more MT is being integrated into translation memory (TM) systems and hence used as an additional source of information in translations. It is argued by do Carmo and Moorkens (2022: 16) that for this approach “the core purpose of translation, which is ultimately the responsibility of the translator, does not change [...] the translator’s role remains unchanged”. TAUS (Massardo et al. 2016) predicted fully automatic useful translation for areas that were previously the domain of human translators, such as user documentation, thus shifting the boundaries between machine and human translation (see also Pym and Torres-Simón 2021). Massey and Ehrensberger-Dow (2017) define the content dimensions “intuitive”, “creative”, and “ethical” as the future domain of high-quality human translation, where the dimension “ethical” also includes the area of high-risk content (see also Massey and Wieder 2018). However, developments in the high-end segment of the language industry, where high-risk documents are processed (amongst others), have not yet been systematically investigated.

Based on empirical evidence proving the efficiency of PE compared to translation from scratch (e.g., Nitze and Oster 2016) and considerations involving risk analysis and risk management (Canfora and Ottmann 2018), we published a decision tree for PE jobs (Nitzke et al. 2019) that aimed to guide stakeholders in charge of deciding whether a project is suitable for PE and, if so, what kind of quality assurance might lead to fit-for-purpose translations (see Figure 1). The criteria in our model included the availability of suitable MT systems, the availability of resources (suitable translators/post-editors/language engineers, suitable tools, time, budget), data confidentiality, the number of recipients of the translation, as well as the text type and the desired quality of the target text. This model is in line with Way, stating that the “degree of human involvement required – or warranted – in a particular translation scenario will depend on the purpose, value and shelf-life of the content” (Way 2013: 2). Another criterion was the specific risks when using MT due to MT errors, liability risks or risks for information security (Canfora and Ottmann 2020).



**Figure 1: The initial decision tree model for PE tasks (Nitzke et al. 2019: 246)**

The decision tree model asks questions in the order suitable for PE projects. Depending on how project managers or the respective decision makers answer these questions, they will end up with a recommendation, i.e., “Use raw MT”, “Use Light PE”, “Use Full PE” or “Don’t use MT at all”.

The decision tree model emerged from an academic approach to a practical matter. The next step, which is presented in this publication, is to validate our considerations with practicing decision makers. Our main research questions are:

- How are PE tasks handled in reality (tools, prices, time budget, etc.) and to what extent do the criteria from our initial model (Figure 1) play a role for project managers and decision makers?
- Are decision makers opting for or against MT based on strategic considerations (e.g. by excluding certain text types categorically from the PE process) or on a project-by-project basis (e.g. due to budget constraints and/or time pressure in individual projects)? Do risk considerations play a role in their decision-making process?
- How should we adapt our model to real-world circumstances?

By addressing these questions, we aim to show how translation automation has been implemented in real-world workflows and whether they can be considered as best practice for the language industry. We furthermore address social sustainability, by investigating the role of the human factor, the status and well-being of translators within said workflows.

In order to do so, we developed a semi-structured interview which we will introduce in Section 2, addressing project managers and other decision

makers from language service providers (LSPs) or other organisations (companies and institutions) that work with MT projects. In Section 3, we introduce our participant group and describe the data collection process. We present our results in Section 4. First, we show the results from the scoring task (Section 4.1) as primarily quantitative data and then the answers from participants to our interview questions (Section 4.2). As the interviews were semi-structured, they opened the floor for the interviewees to reflect on other aspects relating to technological changes in the professional market. Therefore, we present qualitative findings on strategic decisions, risk considerations, as well as the value of translation, working conditions and job profiles in Section 4.3. Finally, we present our revised model, critically discuss the study and give an outlook on further research (Section 5).

## 2. The semi-structured interviews

According to the research questions mentioned above, we outlined a semi-structured interview with 21 open questions and a scoring task to combine both qualitative and quantitative analyses. We piloted the questions in a trial interview and implemented the necessary changes accordingly.

The interviews were structured as follows. First, we asked demographic/metadata questions followed by the first block of questions (Q1 questions in the list below). We then continued with the scoring task and the second block of questions (Q2 questions). Sometimes the order of the questions was adapted to the flow of the interview. Some of the questions were (partly) repetitive, depending on the answers from participants and to check for consistency of answers. We finished the interview by revisiting the scoring task to see if the assessment of the scoring items changed after further discussions.

The interview guideline consisted of the following questions, here ordered in thematic categories. The information in brackets indicates when the question was asked with the first number indicating the question block and the second the question within the block, e.g. question Q1.5 was the fifth question in the first block of questions:

- metadata (6)
  - Where do you work?
  - What role do you have in the LSP/organization?
  - Which domains/text types does your LSP/organization specialise on?
  - Which languages does your LSP/organization work with?
  - Does your LSP/organization work with in-house or external translation providers?
  - What amount does your LSP/organization translate per year?
- decision for/against MT & PE (6)
  - Do you use MT and PE for your translation projects? (Q1.1)

- Are certain projects predetermined for or against MT and PE use? Why? (Q1.2)
- Are certain text types excluded from MT and PE use? (Q1.3)
- Do you debate if MT and PE are suitable for the project before you start the project? Which factors do you take into consideration? (Q1.6)
- Do you sometimes decide actively against MT and PE? (Q2.10)
- When do you full or light post-edit? When do you choose not to use MT? When do you conduct further quality assuring measurements? (Q2.11)
- TM use (2)
  - Do you use MT output as an additional source of information (e.g. additionally to a TM database) or are the texts completely translated by an MT system? (Q1.4)
  - Do you use MT output text-based or segment-based within a TM system? (Q1.5)
- economic factors (3)
  - How do you bill your PE projects (hours, words, etc.)? (Q1.7)
  - How do you calculate the necessary resources for your PE project (time, people, etc.)? (Q1.8)
  - Do you sometimes decide against MT and PE if, e.g., there is enough time or money for a translation without MT assistance? (Q2.12)
- MT systems and training (4)
  - Which MT systems do you use for your projects? Do you use different MT systems for different projects? (Q2.1)
  - Do you host the MT environment yourself or do you use a cloud-based system? (Q2.2)
  - Do you provide training data yourself, do you use pre-trained engines (trained by MT provider), or do your clients provide you with MT? (Q2.3)
  - How acceptable is the MT output? Are there language-specific differences? (Q2.4)
- guidelines and quality (4)
  - Do you conduct any further quality assuring measurements? If so, which? (Q2.5)
  - Do you use the features within a TM system to assure quality? Do you use other tools? (Q2.6)
  - Do you use guidelines or style guides for your projects? If so, which? (Q2.9)
  - Do several post-editors sometimes work on one project? If so, do you conduct extra quality assurance measures? (Q2.13)
- risks (2)
  - Do you use MT also for sensitive texts? (Q2.7)
  - Do you think about risks that can arise from post-editing? (Q2.8)

In the scoring task, the interviewees rated the importance of the following factors in their decision for or against MT in translation projects on a Likert scale, with 0 being “not important” and 4 being “very important”:

- availability of translators (are enough qualified translators available),
- intended publication or purpose of the target text (will the target text be published or is it for internal communication; how many people will read the target text),
- time pressure (is enough time available for the project),
- availability of suitable MT systems,
- sensitivity of the source text content,
- risks in the source text (do the benefits of using MT and PE outnumber the risks, on which risk level is the source text, Canfora and Ottmann 2018), and
- required quality of the target text.

### **3. Data collection and analysis**

The semi-structured interview and scoring task were conducted in the German-speaking market with 19 interviewees from Germany (n=15), Switzerland (n=3) and France (n=1). 17 completed the full interview. All interviews were conducted in German and were on average 41.7 minutes (SD= 14.7; 24-80 min) long. Our 19 interviewees are translation experts from the high-end sector, being active in a variety of domains and offering multiple language combinations. They represented either in-house language departments (n=7) of internationally operating companies or institutions, or independent LSPs (n=12). The seven participants from companies and institutions with in-house language departments have the following functions within their companies and institutions: head of department (n=4), translation manager (n=2), and language data expert (n=1). They operate in the fields of aerospace, heating technology, air conditioning technology, power engineering, climate technology, inspection technology, medical technology, metrology and transportation.

The twelve participants from the independent LSPs have the following functions: language data expert/PE expert/MT manager (n=6), managing director (n=3), translation manager/quality manager (n=2), head of department for strategy and innovation (n=1), and process manager (n=1). They provide language services in a wide range of domains (technical, legal, financial, legal, medical, etc.), but mainly for technical documentation. All of the LSPs could be categorised as “boutique” (USD 1-8m revenues) or “challenger” (revenues USD 8-25m; Slator 2022). As both interview groups belong to the high-end segment, they can give us valuable insights on how the boundary between translation with or without MT assistance shifts and if and how MT and PE are used responsibly for critical text types.

All interviews were conducted online and all but one audio recorded. We used automatic speech-to-text conversion (Adobe Premiere Pro 22.1), which was then manually edited. Our internal editing guidelines were not

highly detailed. The main aim of the transcription was that the reader should be able to follow the conversation without having heard the recording before. During or after editing, we tagged which parts of the text belonged to which question.

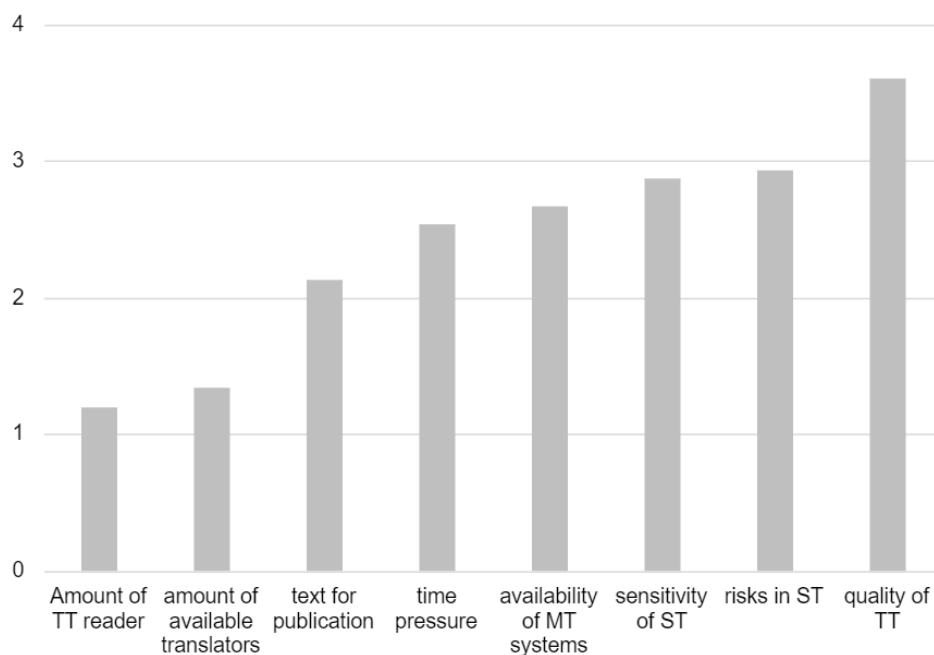
The scoring task was analysed quantitatively (see Section 4.1) and the answers to the actual questions were evaluated quantitatively and qualitatively (see Section 4.2). Some topics were discussed in a broader scope or even beyond the scope than we initiated with our questions. Hence, we analysed these topics qualitatively (see Section 4.3).

## 4. Results

In this section, we will present our quantitative results (Section 4.1) first, followed by the quantitative-qualitative analysis (Section 4.2), and the qualitative evaluation (Section 4.3) of the following topics: risk awareness and risk attitude, strategic decisions and quality considerations.

### 4.1. Quantitative data – the scoring task

To get a quantitative understanding of the importance of certain factors in their decision for or against MT in translation projects, we asked participants to rate the factors included in our decision tree (Nitzke et al. 2019) on a Likert scale, with 0 being “not important” and 4 being “very important”.



**Figure 2: Results of the scoring task**

Figure 2 shows the arithmetic mean of all responses to our scoring task in order from the lowest to highest scoring. The most important factor for our interviewees when deciding for or against MT is the required “quality of the target text” ( $\bar{x}$  = 3.6,  $SD$  = 0.51). The discussion surrounding this factor

showed that the final quality of the target text is an essential factor and that our participants were generally striving for high quality of the translation product (see Section 4.3). However, many of them would use MT and PE for certain projects even when aiming for high quality. So, the use of MT and PE and high quality are not seen as mutually exclusive.

“Risks in the source text” ( $\bar{x}$ = 2.93, SD= 1.33) and “sensitivity of the source text” ( $\bar{x}$ = 2.87, SD= 1.64) follow as second and third place with roughly equal results. “Availability of MT systems” ( $\bar{x}$ = 2.67, SD= 1.35) and “time pressure” ( $\bar{x}$ = 2.53, SD= 1.3) were ranked as medium important factors, while “text for publication” ( $\bar{x}$ = 2.13, SD= 1.46), the question whether “enough human translators” ( $\bar{x}$ = 1.33, SD= 1.18) are available for the translation project, and the amount of “target text readers” ( $\bar{x}$ = 1.2, SD= 1.15) are ranked as least important in descending order. The results of the scoring task have been incorporated into the revision of the decision tree (Section 5).

Finally, the participants were asked to add further factors that influence their decision process and score them. Budget considerations were mentioned quite often ( $n=6$ ), however, mainly regarding if the effort of implementing MT was worth the savings PE would create ( $n=5$ ), while the costs for the customer were only mentioned once. Three MT-related factors were mentioned by two participants (location of MT server, customised vs. default engine, time for training a customised engine). Other factors that were mentioned related to the source text, concerning the text type ( $n=3$ ), domain ( $n=1$ ), language pair ( $n=1$ ), formatting ( $n=1$ ), and quality of the source text ( $n=1$ ).

## **4.2. Quantitative-qualitative data – answers to questions**

In this section, we look at the participants’ answers to our questions. Therefore, we use a mix of qualitative and quantitative analysis. We analyse the questions along the question blocks introduced in Section 2.

The first content block of questions is concerned with the general decision for or against MT and PE. The answers to the first question (Q1.1) brought one of the most surprising results of the study. From our actual interview partners 14 out of 19 use MT and PE “not at all” ( $n=1$ ), “not yet” ( $n=2$ ) or in less than 20% of their projects ( $n= 11$ ). The remaining participants use MT and PE under certain circumstances, e.g. for certain texts or language combinations, however, they were more vague or precautious about a concrete percentage ( $n=4$ ) or they mentioned that they do both ( $n=1$ ) without giving any limitations. Initially, we reached out to 27 potential participants, however seven of them declined their participation right away, because they did not use MT at all.

The other questions in this category were concerned with different aspects of the decision-making process. Q1.2 asked if there were projects where it is predetermined whether to use MT and PE, e.g. by the client. Our participants reported different scenarios and different reasons that factor



into the decision-making process but what was mentioned most often (n=6) was that the clients decide. Other scenarios include internal guidelines that specify when to use MT, such as time or money considerations, and/or project volume.

In Q1.3, we asked if there were text types that were generally excluded from MT and PE. Only four of our participants answered with “no”. The others reported restrictions that varied a lot, however creative texts (n=4) and advertising/marketing texts (n=5) were mentioned most. Further, text types with little context (n=2) or in software environments (n=1), and texts with a high terminology density (n=2) or a lot of abbreviations (n=1) were mentioned. Also, more confidential text types (n=3) like business reports (n=1) or financial statements (n=1) were indicated. Ten participants stated that they sometimes decided actively against MT and PE (Q2.10).

Q1.3 gave the first indications that our participants decided rather strategically rather than evaluating the usefulness of MT and PE for each project anew. This was elaborated in Q1.6. Four participants answered that they decide for or against MT and PE not on a project-by-project basis, but rather that it was a more general decision (see Section 4.3 for the qualitative analysis of strategic decisions). The remaining participants named various factors, the most commonly mentioned of which were clients’ requests (n=5), suitability of the source text (n=5), text type (n=4), and volume of the project (n=4). However, other factors were also mentioned, such as quality of MT output (n=2), languages involved (n=2), domains (n=2), existence of style guides (n=2), cost-benefit considerations (n=2), and price considerations (n=2). Here, we have to keep the potential interplay of the mentioned factors in mind, e.g. languages and domains probably also contribute to the quality of the MT output or a domain is restricted because the content is sensitive.

Seven of our participants reported (Q2.11) that they always use full PE, four use light PE very seldom, and the rest only use light PE when the clients request it (n=6, two of them mentioned that this seldom happens). Neither disregarding the MT output nor extra quality assurance (QA) measures were mentioned. However, the latter was discussed in other questions.

The next block of questions was concerned with the use of TM systems and gave very clear results as almost all of our participants post-edit in TM environments and use segment-based MT as additional input when they get no results from the TM instead of carrying out complete PE projects, with very few exceptions.

Regarding economic factors, the interview contained three questions. Most of them (n=11, only 13 answered this part of question Q1.7) stated that they calculated prices (predominantly) on a word basis. Eleven participants reported an estimation how much PE effects prices compared to translation without MT assistance. On average, they reduce prices 25% (sd: 8.17; min: 10%, max: 50%). Most of them reported a range of reduction according to the project which could be as extensive as 20–50%. When we asked how

they calculated the necessary resources like time or translators for PE projects (Q1.8), we mainly got feedback on the time aspect, which is either estimated as equal or less than human translation. Four participants reported that PE takes more or less the same time as human translation, but the processes differ: the PE process itself, e.g., might be quicker, but the overall process is not (n=2); which is also discussed in do Carmo (2020). One participant reported that they estimate 20% less time for PE compared to human translation, but they assume that translators can only post-edit five hours a day. Our second-to-last question was whether our participants sometimes decided against MT and PE if, e.g., there was enough time or money for a human translation (Q2.12). Most of them said “no” (n=5), two specified that it was the client’s decision, two said “yes”, and one stated that it could happen.

Our next question block was concerned with MT systems and training. Roughly half reported that they use one MT system (n=9) and the rest that they use more than one (n=8; Q2.1). They usually use pro versions of commercial systems that they sometimes train with their own data (see Q2.3). Most participants reported that they use cloud-based systems (n=9) and/or externally hosted servers (n=3; Q2.2). Seven stated (Q2.3) that they used their own training data, four that they worked with closed systems, and only three that they used both closed systems and self-trained systems or only a fully self-trained engine. Finally, one participant from the group of independent LSPs reported that they also received machine translated texts for PE from their clients. The participants reported many differences in the quality of the MT output (Q2.4) and mentioned that they were mostly related to language combinations, text types, and domains.

One question cluster was concerned with guidelines and quality. Almost all participants reported that they conduct further QA measures (Q2.5), mostly through an extra revision step after translation. Further, twelve reported that they did the QA within their TM environment, three mentioned that they used an additional tool, and one said that they had their own QA tool (Q2.6). When asked if they used guidelines or style guides for their projects and if so, which they were using (Q2.9), five answered that they used guidelines but did not specify which they used and two answered “no”. Six mentioned that they use the same guidelines for PE as for human translation and three stated that they used different guidelines and style guides for different clients. Finally (Q2.13), we asked if they sometimes let several post-editors work on one project and if that required extra QA measures. Only three of our participants answered “yes”, while the rest would not do it at all or tried to avoid it. Only two stated that they would conduct extra QA measures if more than one post-editor worked on the job, one said that it would be done externally, one said that it might be done. However, the majority (n=7) stated that those projects would have the same QA standards. Finally, we had two questions specifically targeted at risk considerations (see also 4.3). One aspect we were curious about was whether they used MT for sensitive texts. This is handled very differently.

Five said “yes”, five “no”, five “sometimes”, and four stated that it was done when the clients requested it. Four of them stated that they take the reliability of the engine into consideration. And finally (Q2.8), most of them consider risks (n=7) for MT and PE projects, many of them (n=6) stated that they consider risks as they would with human translation. Three of them even mentioned that they have an insurance policy. One participant stated that if they make the strategic decision to use MT and PE, they don’t consider risks again; and one of them mentioned that they advise clients on MT and PE if they feel that they are not sufficiently risk-sensitive.

### 4.3. Qualitative data – discussion

Of particular interest are the insights of the discussions that did not emerge from direct answers to our questions. These insights either arose from the flow of conversation or across several questions. The three thematic areas we want to discuss here are “strategic decisions”, “risk awareness and risk attitude”, and “the value of translation, working conditions and job profiles”.

#### 4.3.1 Risk awareness and risk attitude

Although “risk” includes both positive and negative deviations from objectives (e.g., ISO 31000 2018 section 3.1), the interviewees interpreted risks almost exclusively as negative. In total, 17 participants made 52 statements about risks, which we used to draw conclusions about risk awareness in three different categories: high-risk awareness, low-risk awareness, or neutral. Statements expressing that the interviewee cannot influence the decisions regarding risks were rated as “neutral statement” (n=3).

Most interviewees had a high level of awareness about the general risks of using MT as this example shows (P14; 00:30:33 - 00:30:50):

Q: In which cases do you actively decide not to use machine translation and post-editing?

A: Yes, the two cases I mentioned, i.e., a) if the order comes from a client [...]. And b) for all confidential texts. And then c) for texts that entail an increased risk. For example, texts related to technical documentation, where a, let’s say, small mistake in translation can pose a serious risk to the person reading it.

MT is here excluded for high-risk texts, differentiating between HT and PE. Only a few interviews contained statements that suggested a relatively low awareness regarding additional risks in PE projects, e.g. P11 (00:29:59 - 00:30:15):

Q: Do you think in general about risks that might arise in post-editing? [...]

A: Well, it’s not different from human translation.

Additionally, many statements revealed a high level of awareness for specific risk types related to MT and PE projects. These statements referred to information security risks, reputational risks, risks of injury, or risks of

property damage. “Information security” (data breach or data loss) was mentioned most frequently (n=18), even if this risk type was not explicitly mentioned in the question. This is consistent with the Allianz (2022) risk barometer, which rates cyber risks as the number one risk for European companies.

The decisions concerning MT in general or regarding specific risk types also depend on the risk attitude of the associated company or organisation (see Dohmen 2011; Tversky and Kahneman 1985). Risk attitude is defined as “the subjective perception of the importance of risk” and is divided into three categories: risk neutral, risk averse and risk seeking (Heckmann et al. 2015: 127). Given identical risk awareness on a certain subject, a risk-averse entity makes a different decision than a risk-seeking entity. In PE projects, risk awareness is directly related to MT literacy (Do I know something about this risk?; Bowker 2021; Krüger 2022), while risk attitude moderates how this risk is dealt with (How important is it to me or to my organisation?; Dohmen 2011). The combination of “high risk awareness” + “risk-averse attitude” can thus lead to either MT being rejected in general or for particularly risky texts (P14; 00:27:29 - 00:28:09):

Q: “Do you also use machine translation for sensitive texts, for example patents, research documents, business reports?”

A: “No, we don’t do that.”

Q: “Although DeepL guarantees data security with the professional version?”

A: “Yes, well, on the one hand, the question is always whether you can be sure that nothing will leak out. That is somehow, well, there have been cases where we were not so sure.”

However, “high risk awareness” + “risk-seeking attitude” can also be found in our interviews. This included all statements in which the specific risks of MT projects are discussed in the organisation, but assessed as manageable because measures have been taken to minimise the risks either by selecting suitable MT providers or licences, or by adopting special QA processes (P11; 00:37:54 - 00:38:14):

Q.: “How important is data sensitivity when deciding for or against PE?”

A.: “I still give zero points because we trust that the systems are closed. Yes, that is what our agreements are. Anything else would be a breach of contract regarding our engines.”

It would be advisable for LSPs or even in-house language services to base their decisions on PE projects not only on “hard criteria” such as the level of risk or text types, but also on the risk attitude of the client or their own company, because ideally the risk attitude of a company should manifest in the decisions of all corporate divisions. It is therefore paramount that project managers and other decision makers in the language industry know how to assess and mitigate the specific risks of PE projects and that these

competencies are included in future translation curricula (see Canfora and Ottmann 2020; Nitzke and Hansen-Schirra 2021; O'Brien and Ehrensberger-Dow 2020).

### **4.3.2 Strategic decisions**

31 statements by 14 participants explicitly address strategic decisions. They commented, for instance, that they do not plan the use of MT and PE on a project-by-project basis, but that they rather establish a long-term strategy for dealing with PE for certain clients or text types and therefore no long decision-making process is necessary for individual translation projects.

As examples they mentioned business reports, user interfaces and transcreation projects, which they generally reject for MT and PE, while texts were often translated with MT for internal or gisting purposes. Some participants also stated that they generally reject MT when the source texts contain defects, which include segmentation and layout problems.

Another aspect that plays a role in strategic decisions are language combinations, as some generally deliver low-quality MT output and therefore cause more effort than human translation, regardless of the text type. Furthermore, the existence of training corpora was mentioned as a factor for a strategic decision for or against MT/PE. Very interestingly, some of our interviewees reported that they use MT to be regarded as cutting-edge and to offer state-of-the-art technologies, regardless of whether they achieve a higher margin in the projects as a result.

A: "And also this, I don't know how to call it, keeping up-to-date, to embrace the development and to be able to offer it [MT] to the customer." (P3; 00:16:23:13 - 00:16:40:09)

### **4.3.3 The value of translation, working conditions and job profiles**

Some topics arose during the interviews without being explicitly addressed. We identified 20 statements by 13 participants of this kind in our discussion since they provide valuable, however anecdotal insights on the value of translation, working conditions and job profiles.

Concerning the value of translation, we came across negative and neutral statements. Translation itself was reported as still not being recognised as an integral part of the value chain that not only raises costs but also saves money. Regarding MT and PE, it was reported that some translators first perceived it as a downgrading of the profession, but that they saw the benefits after a while and understood that human input was essential to create high quality. One participant felt that two sectors have evolved in the translation business, including a high-end business ("luxurious segment", P17, 00:31:32) for translators and the rest for PE.

When it comes to working conditions, different opinions can be detected. One participant perceived a neutral effect from MT and PE on working conditions as MT occasionally delivers very good inspiration, but on the

other hand, it is not objectively proven whether PE of MT always saves time. For some translators, better working conditions have evolved, shifting the former production pressure to other responsibilities such as training and supporting external translators, developing terminology resources, and improving the MT output. In contrast, increasing pressure is perceived by another interviewee to get more done in less time and for less money. Another aspect that was discussed by some participants was an adaptation and development phase to address the challenges for translators. Some introduced payment by the hour for PE tasks because it is hard to calculate how long a PE task will take translators, even though the LSP generally prefers to pay by word. Fair working conditions would allow the translators to get used to the new working mode.

Finally, there has been a lot of discussion on how the field and job profiles have been changing (Gambier 2014; Nitzke and Hansen-Schirra 2021), which P10 (00:24:30 - 00:24:50) also mentioned in his interview: “[...] the job description of a translator will, I think, massively change in the next years. Therefore, I think it is incredibly interesting to observe how the whole set-up is changing”. They argue that there are already scenarios in internal technical communication where pure MT output is sufficient to convey the information. Some statements indicated that PE would need some additional training. One participant exemplified how they, at a former employer, expected translators to translate up to 3,000 words a day, while they expected post-editors to post-edit up to 10,000 source text words a day depending on the expected quality. They also discussed a potential ignorance of content mistakes in PE of neural MT, proposing this as one of the issues that needs to be addressed in PE training.

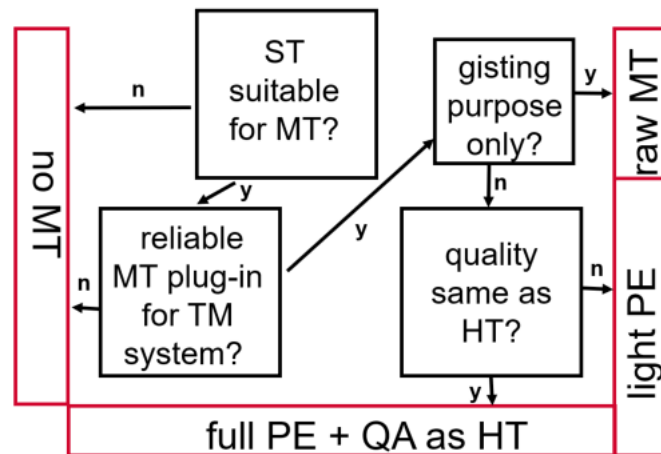
Many of the interviews suggest that consultation and expectation management of clients play a big role in the language industry at the moment. Further, one participant discussed how they would offer MT and PE to their clients as an option for very urgent projects. The clients need to be consulted about the different factors to be taken into consideration in the decision process. Consulting services need to be offered project-wise but also at regular intervals as a strategic consultation.

## **5. Discussion and the updated model**

Our starting point for this article was to compare the academic model we developed for decisions in PE projects with the decisions made in practice. Our first question was, hence, how PE tasks are handled in reality, and the interviews brought some unexpected insights. First, we were surprised to see that far less MT and PE are used for translation projects than we expected in the high-end segment of the language industry. The decisions for or against MT and PE are often rather strategic and not taken for each individual project. When MT is integrated in the process, our participants do not often differentiate between handling translation projects with or without MT assistance, especially in regard to use of MT in TM systems, QA measurements, and expected final quality. Their use of MT and PE indicates

that the translators can maintain a certain degree of control as to when to use MT output and thus attain a higher degree of self-efficacy and autonomy. This can contribute to translators' acceptance of the technology and is consistent with the findings of Vieira and Alonso (2020: 14) and Krüger (2019). Revision and QA measures were more often included in the PE process than we expected. However, this is in line with the use of MT in TM environments and the high-quality expectations. Hence, MT output is primarily just an additional source of information in the translation process and the translation processes with and without MT assistance are almost identical for most projects.

The second research question was concerned with the role of strategic decisions and risk considerations. Overall, our participants decide to use MT thoughtfully and cautiously when projects fulfil certain characteristics. These characteristics are determined strategically and not on a project level. Most participants have a high level of risk awareness and make sure to use MT only when the risks can be managed. Further, the quality of the target texts is generally their main concern and the margins are not regarded as their highest priority, although attention is certainly paid to translation costs. Our results show that aspects such as shelf-life or outreach of the final text, are not weighted as much as we initially expected from our 2019 model. Above we argue that our initial model was in line with Way (2013), however our results show that in high-end market segments, where light PE is rarely used, decision criteria such as shelf-life etc. are not taken into account. We believe that this new model nevertheless does not contradict Way (2013), as he also refers to use cases where raw MT plays a role, e.g. for texts with no shelf-life. Our interviewees do not offer translation services involving raw MT. With the results of our study, we revised our model from 2019. We reduced the model to four main decisions (see Figure 3) concerning the suitability of the source text as a starting point, the reliability of the MT (system and output), the purpose of the final text, and the quality requirements for the final text. The final outcomes of the decisions are similar to the original model, but they are weighted differently. As we have seen when we have contacted potential participants and in the answers to Q1.1, the decision not to use MT and PE is still quite prevalent, at least in the segment represented by our participants. Q2.11 and the questions concerned with guidelines and quality showed that full PE plus QA measurement is the most common procedure for MT and PE. The additional QA step was not represented in our initial model and was hence added to the decision for full PE. In our original model, full and light PE as well as raw MT took an equal amount of space on the continuum. This was adjusted according to our participants' answers.



**Figure 3: The Updated Model**

Two of our initial questions were excluded for the revised model as they play little or no role, namely "How many people will have access to the translation?" and "Do I have enough translators/time?" In the scoring task, "amount of target text readers" and "text for publication" (concerning the first question) as well as "number of available translators" and "time pressure" (concerning the later question) were scored the lowest and the answers to the questions from the content block "economy" reinforce this decision.

The remaining questions and considerations were summarised in two questions which can be considered as relatively broad. First, a source text can be suitable for MT if the content is assessed as not sensitive, the translation risks are low, and the language, domain, and text type are fitting for MT and create an acceptable output. However, what is considered sensitive and risky will depend on the client's risk attitude and can therefore differ.

The next decisive question concerns the reliability of the MT output. The MT output should qualitatively be good enough to increase productivity compared to translation without MT assistance and the MT system should protect the data as much as needed for the text. As almost all participants stated that they use MT in a TM environment, we added that a plug-in must be available for the respective TM system. Both questions could have been first in our model. We decided to put the question concerning suitability of the source text before the question concerning the availability of a reliable MT as "availability of MT system" was only scored as the fourth most important factor. As a result, the first question addresses those parameters right at the beginning of the decision-making process that we classified as "strategic decisions" (risk attitudes and risk considerations, text types, domain and language combinations). In contrast to our original model, we thus assign more weight to the strategic decisions.



If any of these two first questions must be answered with “no”, it is not recommended to use MT for the project, which reflects the comparably low reported use of MT and PE in our data. The last two questions are directed towards the quality of the final target text. The interviews showed that light PE plays a minor role in our participants’ work. We did not directly address if raw MT was used for gisting purposes and it was rarely mentioned, however, it is a potential option. We recommend using raw MT output for gisting purposes and light PE if the quality expectations of the final target text are not too high. If the final quality is supposed to be as good as human translation, full PE and further QA measures are recommended.

Our participants can be considered as part of the high-end segment of the language industry and, to them, there is a clear difference to the scenarios Moorkens (2020a, 2020b) and Firat (2020) describe by showing the negative effects of platform economy and bulk translation on individual translators. Our interviews showed that the human factor still has its place and translators still have the opportunity to work under reasonable conditions despite the massive changes due to digitalization and automation, even though some of our interviewees expressed huge concern about the future of the translation profession. Of course, our participants may have given more socially desirable answers. Nevertheless, we can conclude that working conditions and prospects for highly qualified and technology-savvy translators in the high-end segment are good despite, but possibly also because of, the changes in the industry (cf. OECD 2021).

We would argue that the results are transferable to other environments, as our participants conduct business on international markets, in various fields and with various language combinations. They were thus able to provide us with insights into these global language industries. Furthermore, these results are also in line with Massey and Ehrensberger-Dow (2017) and Pym and Torres-Simón (2021) and support the job polarisation theory (Goos and Manning 2007, see also Vieira 2018), which has been investigated and supported for jobs exposed to artificial intelligence by Felten et al. (2019; see also OECD 2021). Future studies should be conducted with larger participant groups further extending the pool of language combinations and domains. In order to get a comprehensive picture of the industry, it would also be important to take a closer look at the other players and to examine the working conditions from the point of view of translators, revisers and editors.

## References

- **Allianz** (2022). *Allianz Risk Barometer 2022*. Munich: Allianz Global Corporate & Specialty.
- **Bowker, Lynne** (2021). “Promoting Linguistic Diversity and Inclusion: Incorporating Machine Translation Literacy into Information Literacy Instruction for Undergraduate Students.” *The International Journal of Information, Diversity & Inclusion* 5(3), 127-151.

- **Cadwell, Patrick, Sharon O'Brien and Carlos S.C. Teixeira** (2018). "Resistance and accommodation: factors for the (non-) adoption of machine translation among professional translators." *Perspectives* 26(3), 301-321.
- **Canfora, Carmen and Angelika Ottmann** (2018). "Of ostriches, pyramids, and Swiss cheese - Risks in safety-critical translations." *Translation Spaces* 7(2), 167-201.
- **Canfora, Carmen and Angelika Ottmann** (2020). "Risks in Neural Machine Translation." *Translation Spaces* 9(1), 58-77.
- **Carl, Michael, Silke Gutermuth and Silvia Hansen-Schirra** (2015). "Post-editing machine translation: Efficiency, strategies, and revision processes in professional translation settings." Aline Ferreira and John Schwieter (eds) (2015). *Psycholinguistic and Cognitive Inquiries into Translation and Interpreting* (Vol. 115). Amsterdam: John Benjamins, 145-174.
- **do Carmo, Félix** (2020). "'Time is money' and the value of translation." *Translation Spaces* 9(1), 35-57.
- **do Carmo, Félix and Joss Moorkens** (2022). "Translation's new high-tech clothes." Gary Massey, Elsa Huertas-Barros and David Katan (eds) (2022). *The Human Translator in the 2020s*. London: Routledge, 11-26.
- **ELIA, EMT, EUATC, FIT EUROPE, GALA, LIND, WOMEN IN LOCALIZATION** (2022). *European Language Industry Survey 2022. Trends, expectations and concerns of the European Language Industry*. Brussels: Elis Research.
- **Felten, Edward W., Manav Raj and Robert Seamans** (2019). "The Occupational Impact of Artificial Intelligence: Labor, Skills, and Polarization." NYU Stern School of Business.
- **Firat, Gökhan** (2020). "Uberization of translation. Impacts on working conditions." *The Journal of Internationalization and Localization* 8(1), 48-75.
- **Gambier, Yves** (2014). "Changing landscape in translation." *International Journal of Society, Culture & Language* 2(2) (Special Issue on Translation, Society and Culture), 1-12.
- **Goos, Maarten and Alan Manning** (2007). "Lousy and Lovely Jobs: The Rising Polarization of Work in Britain." *The Review of Economics and Statistics* 89(1), 118-133.
- **Heckmann, Iris, Tina Comes and Stefan Nickel** (2015). "A critical review of supply chain risk – Definition, measure and modelling." *Omega – The International Journal of Management Science* (52), 119-132.
- **Hvelplund, Kristian Tangsgaard** (2016). "Cognitive efficiency in translation." Ricardo Muñoz Martín (ed) (2016). *Reembedding Translation Process Research*. Amsterdam: John Benjamins, 149-170.
- **ISO 17100** (2017). *Translation services – Requirements for translation services – Amendment 1*. Geneva: ISO.
- **ISO 18587** (2017). *Translation services – Post-editing of machine translation output – Requirements*. Geneva: ISO.
- **ISO 31000** (2018). *Risk management – Guidelines*. Geneva: ISO.

- **Krüger, Ralph** (2019). "Augmented Translation – Eine Bestandsaufnahme des rechnergestützten Fachübersetzungsprozesses." *trans-kom* 12(1), 142-181.
- **Krüger, Ralph** (2022). "Integrating professional machine translation literacy and data literacy." *Lebende Sprachen* 67(2), 247-282.
- **Massardo, Isabella, Jaap van der Meer and Maxim Khalilov** (2016). *TAUS Translation Technology Landscape Report*. De Rijp: TAUS BV.
- **Massey, Gary and Maureen Ehrensberger-Dow** (2017). "Machine learning: implications for translator education." Paper presented at *CIUTI forum 2017* (Geneva, Switzerland, 12-13 January 2017).
- **Massey, Gary and Regine Wieder** (2018). "Educating translators for new roles and responsibilities: Interfacing with corporate and technical communication." Paper presented at *didTRAD 2018: 4<sup>th</sup> International Conference on Research into the Didactics of Translation* (UAB Barcelona, Spain, 20-22 June 2018).
- **Moorkens, Joss** (2020a). "A tiny cog in a large machine. Digital Taylorism in the Translation Industry." *Translation Spaces* 9(1), 12-34.
- **Moorkens, Joss** (2020b). "Translation in the neoliberal era." Esperança Bielsa and Dionysos Kapsaskis (eds) (2020). *The Routledge Handbook of Translation and Globalization*. London: Routledge, 323-336.
- **Nimdzi** (2022). The size and state of the Language Service Industry in 2022, including the ranking of the 100 largest Language Service Providers. <https://www.nimdzi.com/nimdzi-100-2022/> (consulted 29.08.2023).
- **Nitzke, Jean and Silvia Hansen-Schirra** (2021). *A short guide to post-editing*. Translation and Multilingual Natural Processing vol. 16. Berlin: Language Science Press.
- **Nitzke, Jean, Silvia Hansen-Schirra and Carmen Canfora** (2019). "Risk management and post-editing competence." *The Journal of Specialised Translation* 31, 239-259.
- **Nitzke, Jean and Katharina Oster** (2016). "Comparing translation and post-editing: An annotation schema for activity units." Carl, Michael, Srinivas Bangalore and Moritz Schaeffer (eds.) (2016). *New directions in empirical translation process research: Exploring the CRITT TPR-DB*. Heidelberg: Springer, 293-308.
- **O'Brien, Sharon and Maureen Ehrensberger-Dow** (2020). "MT Literacy – A cognitive view." *Translation, Cognition & Behavior* 3(2), 145-164.
- **OECD** (2021). "Artificial Intelligence and Employment. New Evidence from Occupations most exposed to AI". *Policy Brief on the Future of Work*. Paris: OECD Publishing.
- **Olohan, Maeve** (2021). "Post-editing: a genealogical perspective on translation practice." Mario Bisiada (ed.) (2021). *Empirical studies in translation and discourse*. Translation and Multilingual Natural Language Processing vol. 14. Berlin: Language Science Press, 1-25.
- **Pym, Anthony and Ester Torres-Simón** (2021). "Is automation changing the translation profession?" *International Journal of the Sociology of Language* 2021(270), 39-57.

- **Sánchez-Gijón, Pilar, Joss Moorkens and Andy Way** (2019). "Post-editing neural machine translation versus translation memory segments." *Machine Translation* 33, 31-59.
- **Slator** (2019). *Neural machine translation report: Deploying NMT in operation*. Zurich: Slator.
- **Slator** (2022). *The Slator 2022 Language Service Provider Index*. <https://slator.com/slator-2022-language-service-provider-index/> (consulted 29.08.2023).
- **Tversky, Amos and Daniel Kahneman** (1985). "The Framing of Decisions and the Psychology of Choice." George Wright (ed.) (1985). *Behavioral Decision Making*. Boston: Springer, 25-41.
- **Vieira, Lucas N.** (2018). "Automation anxiety and translators." *Translation Studies* 13(1), 1-21.
- **Vieira, Lucas N. and Elisa Alonso** (2020). "Translating perceptions and managing expectations: an analysis of management and production perspectives on machine translation." *Perspectives* 28(2), 163-184.
- **Way, Andy** (2013). "Traditional and emerging use-cases for machine translation." *Proceedings of Translating and the Computer* 35 (London, 28-29 November).

## Biographies

**Jean Nitzke** is an Associate Professor with a focus on translation technology at the University of Agder, Norway. She was a lecturer and researcher at the Johannes Gutenberg University Mainz, Germany from 2012-2021 and substituted a full professor position at the University of Hildesheim, Germany (2019-2020). Her main teaching and research interests are post-editing machine translation, translation technologies, domain-specific translation, cognitive translation studies, and translations into the foreign language.

[jean.nitzke@uia.no](mailto:jean.nitzke@uia.no)

ORCID-ID: 0009-0002-7409-1796



**Carmen Canfora** holds a PhD in translation studies from Johannes Gutenberg University Mainz in Gernersheim, Germany and has been a freelance translator for over 20 years. She has been a lecturer and researcher at Johannes Gutenberg University of Mainz since 2008. Her main teaching and research interests are risk management, quality management, translation process research and translator education.

[canfora@uni-mainz.de](mailto:canfora@uni-mainz.de)

ORCID-ID: 0000-0001-9469-9702



**Silvia Hansen-Schirra** is full Professor for English Linguistics and Translation Studies and Director of the Translation & Cognition (Tra&Co) Center at Johannes Gutenberg University Mainz in Germersheim. She is the co-editor of the book series “Translation and Multilingual Natural Language Processing” and “Easy – Plain – Accessible”. Her research interests include machine translation, accessible communication and translation process research.

[hansens@uni-mainz.de](mailto:hansens@uni-mainz.de)

ORCID-ID: 0000-0002-3277-7717



**Dimitrios Kapnas** holds two M.A. Diplomas, one in Translation and one in Conference Interpreting. He finished his studies at the Johannes Gutenberg University Mainz in Germersheim in 2022. He currently is a doctoral student at the Tra&Co Center. His research interests include machine translation, accessible communication, easy language as well as gender linguistics.

[dikapnas@uni-mainz.de](mailto:dikapnas@uni-mainz.de)

ORCID-ID: 0009-0007-9592-3092

