Incorporating post-editing into a computer-assisted translation course. A study of student attitudes.

Olga Witczak
Faculty of English, Adam Mickiewicz University
Poznań, Poland
owitczak@wa.amu.edu.pl

ABSTRACT

Computer-assisted translation (CAT) has become the norm rather than exception in the translation profession today. However, machine translation (MT) as a human translation aid is still facing a negative attitude in the translators’ community. When years of training are replaced by the quick and oftentimes inaccurate MT suggestions, post-editing is regarded as dehumanising the process of translation. This paper discusses the introduction of a practical post-editing component into a CAT tools course for MA translation trainees after a series of classes devoted to general features of Wordfast. In Wordfast Anywhere the students post-edited a fragment of a vacuum cleaner manual (Exercise 1) and then either a newspaper article or a patient information leaflet (Exercise 2). Exercise 3 involved a peer-review of post-edited texts according to TAUS post-editing guidelines. Also, both before and after the practical exercises students filled in questionnaires about their translation as well as language experience, the source texts, and their attitude towards MT, which enabled them to reflect on their performance and decisions while post-editing. The aim of including a post-editing component in a CAT tools course was to familiarise translation trainees with MT. Then, they would be able to adapt to a growing demand on the market, be aware of MT’s shortcomings, and not feel threatened by a tool designed to boost translation productivity. The class environment caused limitations to data collection; however, the questionnaires and the post-editing practice organized in the spirit of action research and socio-constructivist framework, provide an insight into novices’ attitudes towards working with MT and may help in designing a standalone course in post-editing.

Keywords: post-editing, machine translation, translation pedagogy, translation studies, action research

1. Introduction

Translator competence in the age of technology is invariably intertwined with the ability to use digital translation aids. Thus, the work of a translator is inevitably, as Pym (2011: 4) put it, “(…) determined by internet searches, glossaries, spell checkers, grammar checkers, translation memory and machine-translation databases, and anything else resembling a communication technology.” While digital translation aids have been used by translators for some time, machine translation (henceforth MT) is still rather notorious among the translating practitioners. MT is understood in this paper as a translation produced by a computer (Oxford English Dictionary, 2016). Regardless of the widespread bias towards MT, systems generating automatic translations have been developed and researched over the years and the quality of the output is being constantly

1 I would like to thank Bogusława Whyatt, my PhD supervisor, for her invaluable advice and guidance throughout this project.
improved (Doherty, 2016). Furthermore, while the ALPAC objective of FAHQMT (Hutchins, 2007: 5), i.e. fully automatic high quality machine translation, is yet to be achieved, MT can be successfully utilised to aid human translators in a process called post-editing (henceforth PE). This process combines both the automated element (raw MT output) and the active human translator, who is the decision-making agent.

PE can therefore comprise “tidying up the raw output, correcting mistakes, revising entire, or, in the worst case, retranslating entire sections” (Somers, 2001: 138). This process differs substantially from translation from scratch (Carl, 2013). Modern Computer-Assisted Translation (henceforth CAT) tools integrate MT into a translation environment, thus combining PE with working with translation memory (henceforth TM) and other translation aids. This paper, focuses on full PE as a standalone task – as opposed to light PE (cf. TAUS, 2010), albeit not excluding the use of online resources – which aimed at familiarising translation trainees with MT as a translation aid and investigating their attitudes towards PE before and after a series of exercises.

Section 2 briefly discusses attitudes towards MT as well as technology in the translation community reported in research, PE and CAT tools in a translation training curriculum, and examples of research corroborating productivity gains in PE. Sections 3 and 4 describe an empirical investigation of student attitudes towards MT in a series of PE exercises, focusing on the participants, materials, methodology, and procedure. Finally, in section 5 results and practical implications are discussed and some tentative conclusions are drawn.

2. Literature Review

Machine translation still arouses controversy in the translation community, which is far from unanimously embracing it as a productivity boost. This reluctance may be caused by the fact that the quality of MT output is far from perfect and MT engines still cannot produce translations that would equal human translations. It is also the dullness of the PE process that partially explains the negative attitude towards MT. Translators’ work is “demoted to the status of a fixer of seemingly unintelligent errors” (O’Brien, 2012: 10) when they are asked to work with MT suggestions, often also being paid less than for producing a translation from scratch (Doherty, 2016; Krings, 2001). Moreover, when a human translator’s many years of training are replaced by the quick and oftentimes inaccurate computer-generated results that MT produces, PE is regarded as dehumanising the process of translation (O’Brien, 2012). O’Brien (2012) mentioned that job satisfaction is affected whenever the task is to edit machine-made mistakes rather than focus on the so-called translation proper, which also affects human creativity. Moreover, automatic translation output rarely exceeds expectations of quality, which means that translators have to suffer exposure to poorly written texts. This in turn could skew their perception of good quality; however, currently no evidence has been found to support this argument (O’Brien, 2012). At the same time, the lack of social awareness of tools boosting translation productivity is still an issue as a professional is expected to produce texts of highest quality at an increasingly
shorter notice (O’Brien, 2012). This gives a good reason to include a PE course in a translation training programme, as Doherty (2016) advocates “the need of fundamental awareness of and accessible education for translation technologies, their strengths and weaknesses, and their impact on international and intercultural communications for all stakeholders.”

Reluctance towards PE may be also explained through the lens of a more general phenomenon. Pym (2011: 4) noted that “[r]esistance to technological change is usually a defence of old accrued power, dressed in the guise of quality”. The statement only emphasises the dire need to raise translators’ awareness of MT possibilities as well as its limitations. The still prevalent negative attitude as well as prejudice is surprising as empirical studies corroborate the productivity-boosting capacity of PE. For instance, O’Brien (2007), Carl et al. (2011), Läubli et al. (2013) reported shorter task time for PE when compared with translation from scratch. Also, Plitt and Masselot’s (2010) study showed an increase in productivity for post-editors when compared with translation times.

Thus, one way to raise awareness of MT capabilities among translators is to include it into translation training curricula. While translation technology is already incorporated into those curricula and the need for technological literacy is undisputable, PE is still not a priority. Şahin and Dungan (2014) point out that technological competence is one of the six competences enumerated by the European Master’s in Translation (henceforth EMT) programmes. Thus, the EMT’s strategy for the year 2012 and beyond features the objective “[to] acknowledge, recognise, and respond to changes in translation training imposed by technological and market developments” (European Commission, Directorate-General for Translation, 2012: 2). EMT programmes include teaching technological skills that are necessary for modern translators in their course objectives, e.g. using search engines, corpora, term extraction, CAT tools as well as raising awareness of MT limitations and applications (Bowker, 2014: 90 after EMT expert group, 2009: 6–7).

Along general technological competence connected with translation, the need to teach MT PE has been argued since 2002, when Sharon O’Brien published a course proposal for PE (2002: 100). She reported an increasing demand for translation along with the demand for its fast delivery as reasons to include PE in translation training courses, which would make the graduates of such courses comfortable with PE. In a conference talk, more than a decade after publishing the paper, O’Brien (2014) confirmed that post-editors are indeed needed because of the current demands of the market. In the same talk (O’Brien, 2014) it was also stressed that PE needs extra skills on top of those required of translators, with which Doherty (2016) also agrees. In a different paper (O’Brien, 2012: 22) pointed out that a translator nowadays ought to take the “opportunity to expand skill sets and take on new roles.”

This is echoing Kiraly’s socio-constructivist approach towards translator competence:

---

1 According to Doherty (2016: 949), the language services market displayed an annual growth rate of 5.13% in the years 2009 to 2013.
Translator competence does not primarily refer to knowing the correct translation for words, sentences or even texts. It does entail being able to use tools and information to create communicatively successful texts that are accepted as good translations within the community concerned. (…) With the changes in the translation profession in mind, it is time to reconsider the viability of conventional approaches for educating translators, which date back almost half a century, when the translation profession was something altogether different from what it is today.

(Kiraly, 2000: 13-14)

Kiraly’s words from the beginning of this century seem to be still valid as not only in-service translators need to adapt to the ever-changing market, but also trainees need to be ready to use new tools and technologies. Awareness of this necessity needs to be raised during training.

An example of an experimental course in MT PE is Mehmet Şahin’s (2014) course, described in his presentation during the Tralogy II conference. Şahin explored students’ attitudes towards MT and aimed at increasing their translation competence. The course also had the purpose of acquainting the students with PE and the conclusions presented show that MT PE could be an important and informative part of a translation teaching curriculum. According to Şahin (2014), PE also allows to help students monitor their translation decisions if process logs in the form of questions that the participants answered during and after the PE task, are kept and analysed. Şahin explained that “[a]lthough these questions do not reflect the whole of the translation process, they attempted to channel participants’ attention to the ‘success’ of the MT so that they could identify which parts of the MT output they should focus on” (Şahin, 2014: unpaginated). Thus, not only incorporating PE into training curricula may help develop translator competence but also analysing and reflecting on the process of PE may prove beneficial. Both aspects were included in the experiment, on which this paper reports.

3. Experiment

The PE experiment was a part of a general CAT tools course for translation trainees during the first year of their postgraduate studies and it took place during 2 classes (90 min each) plus a summing-up discussion during the 3rd class. Before the PE exercises there were 4 classes focused on Wordfast Classic3 (henceforth WFC) and Wordfast Anywhere4 (henceforth WFA) exercises (e.g. creating and editing glossaries, TMs, as well as various functions of both WFC and WFA). A screenshot of WFA user interface is visible in Figure 1.

3 An offline CAT tool, which is a MS Word-based macro (Retrieved from: http://www.wordfast.com/products_wordfast) (date of access: 29 April 2016).
4 An online and cloud-based TM service (Retrieved from: https://www.freetm.com/) (date of access: 29 April 2016).
Thus, PE was introduced once the students were familiar with the CAT environment of WFA. However, the main objective of the incorporation of MT PE into a translation training course was to acquaint the students with this way of performing a translation and to let them decide for themselves whether they thought MT was useful or not. This was based on Kiraly’s (2012: 84–85, 89) approach, namely “to empower [students] to take responsibility for their own learning, their own sense-making and their own futures” as well as “to contribute to the emergence of independent thinkers, competent heuristic problem-solvers and knowledgeable translators.” The classes were designed using the action research framework, which is discussed in section 3.3. The data gathered from the exercises is both qualitative and quantitative.

3.1. Research questions

The experiment primarily attempted to answer the following research questions:

1) What are students’ attitudes towards machine translation?
2) Do their attitudes differ for different text types (newspaper article vs. manual)?
3) Do students’ attitudes change after a PE exercise?

3.2. Participants

The participants were translation trainees in their master’s programme (1 MA) at Adam Mickiewicz University (henceforth AMU). Thus, they were beginner translators who also admitted to little or no experience with PE. The first group was from a written translation specialisation at the Faculty of English (L1: Polish, L2: English, N=12). The second group was from a Master’s Programme in Specialized and Professional Translation, run by the Faculty of Modern Languages jointly with Faculty of English at AMU (L1: Polish, L2: English, L3: German/French, N=9).
3.3. Methodology, materials, procedure

In order to supplement Kiraly’s socio-constructivist approach towards translation pedagogy, action research was adopted as a theoretical framework for conducting the experiments in a naturalistic, classroom setting. As can be seen in the questionnaires (cf. Appendix), the participants signed them and their post-edited texts with their names so they took full responsibility for their opinions and translations. However, for the purposes of this paper all the names have been coded and anonymised. The crucial aspect of the action research framework used in the study was the emphasis on reflection. As Cravo and Neves (2007) point out, “it is through the reflection of both the researcher and all the other participants in the process that problems are addressed, new decisions are made and new plans are devised.” Additionally, reflection encourages self-criticism (Hatim, 2001), helps understand the reality of the profession and incites, among other things, dialogue, empowerment, social justice, as well as improvement (Cravo & Neves, 2007).

The first class started with an introduction into PE. The teacher prepared a presentation, which explained, among others things, the differences between MT and CAT, how statistical machine translation works, and provided a very brief history of MT. The presentation also included information concerning the applications and limitations of MT. The applications mentioned to the students included gisting, instant messaging/e-mails, social media and localisation (Fiederer & O’Brien, 2009; Qun & Xiaojun, 2015; Somers, 2003). The limitations that were presented to the students included features of natural language that MT may have problems parsing: ambiguity, untranslatability, lexical/structural mismatches and idioms/collocations (Arnold et al., 2001; Baker, 1992; Lewicz, 2013). The students also got to know what PE is and how it differs from translation and human translation revision as well as reasons for employing PE to create a translation.

After the introductory presentation, the participants filled in a pre-task questionnaire in Google forms concerning both their language proficiency and the text they were about to post-edit. This questionnaire and all the ones that followed were based on closed and open questions. In other words, the students provided answers on a five-point Likert scale but also provided longer answers for some of the questions. Before commencing the PE task, the students received a list of PE guidelines based on those available on the TAUS website5 (TAUS, 2010). Then, they proceeded to the PE (Exercise 1) and after finishing it, they filled in another Google forms questionnaire, this time documenting their impressions after the task. The source text for PE was a fragment (137 words) of vacuum cleaner safety instruction in Polish, which the participants post-edited into English.

There was a very similar procedure for Exercise 2: an online questionnaire before (Exercise 2.1) and after the task (Exercise 2.2), this time referring only to the text. The students were assigned one of the two English texts, which they then proceeded to post-edit into Polish. One of the texts was an excerpt (195 words)

---

5 TAUS stands for the Translation Automation User Society. The guidelines were those for achieving quality similar or equal to human translation on the TAUS website.
from a newspaper article from the New York Times\(^6\) (henceforth *NY Times article*) and the other one was a fragment (184 words) of a Paracetamol leaflet (henceforth *Paracetamol leaflet*). The reason why such different texts were selected for PE is that in the final task, Exercise 3, the students would see the contrast between the two text types in terms of MT quality. Doherty (2016: 953) refers to text type as an important characteristic of MT. Namely, “it typically works best with simplistic and repetitive linguistic features and within the same genres, domains, and texts types.”

Exercise 3 was a questionnaire-based review of a post-edited text but not the same text type that the reviewer post-edited themselves. That is, if a student post-edited the *Paracetamol leaflet*, they would get a post-edited version of the *NY Times article* to rate on a 5-point scale (1: strongly disagree – 5: strongly agree) with reference to statements based on TAUS guidelines (2010).

The first two exercises were completed in WFA. The provider of MT output was WorldLingo and Microsoft Translator, however, the students were instructed to use Microsoft MT suggestions exclusively via the Force MT function in WFA due to the poor quality of formatting performed by World Lingo, which would have consumed a much larger amount of time for post-editing. During Exercise 1 the participants uploaded the MT suggestions on their own by means of the above-mentioned function, however, shortly before Exercise 2 WFA underwent an upgrade and removed WorldLingo for Polish, which disabled free MT for the Polish-English language pair. Therefore, in this unfortunate and unforeseeable situation the students were instructed to copy and paste raw MT output from the Microsoft Translator website, while the other group – which had the class a week later – was supplied with the source texts and MT output already in WFA via the Share function. The experiment was finalised with a discussion on all the three exercises and a summary of the general impressions of PE.

4. Results

4.1. Exercise 1.1

21 students took part in this exercise and in the pre-task questionnaire more than half (12) admitted to never having post-edited MT before, as Figure 2 illustrates:

![Figure 2](chart.png)

Figure 2. Have you ever post-edited MT output before? \(N = 21\).

In the same questionnaire all the participants assessed their language and translation experience and the level of difficulty of the source text\(^3\) on a five-point Likert scale (cf. Figure 3) from 1 (very experienced/very easy/very familiar) to 5 (very inexperienced/very difficult/very unfamiliar). The source text was a fragment of a Polish vacuum cleaner manual and on average the students admitted to being relatively familiar with the text type and that the text was of average difficulty. They were quite confident about their L2 proficiency, but not so much about their translation experience, which is understandable considering that at the time they were in their first semester of the MA programme.

Figure 3. Exercise 1.1: before post-editing. \(N=21\). Error bars represent standard deviations.

The participants also provided a longer justification of their attitude towards PE, i.e. whether they considered it a meaningful way of performing translation or not. 11 out of 21 students were explicitly in favour of PE, while 9 were undecided, and only one gave a directly negative answer. They admitted that PE accelerates work, saves time, and is easy and quick (“it makes the translator’s work much easier and quicker when translator uses first machine translation and then post-edits the text” (ST201\(^8\)), however, it is best used for simpler texts with little specialised language (ST204). There were conflicting opinions as well, e.g. that PE is only meaningful for short sentences (ST107), whereas another student (ST106) thought the exact opposite. Concerns were voiced with reference to the overall coherence and quality of the final translation, not only the MT output. The following quotation illustrates that they were aware that MT may not be as good for every text type and that they had to, above all, be able to assess whether the editing would exceed the time of manual translation or not: “If the translator assesses the source text as one that can be machine translated and post-edited in as efficient [sic] (or more efficient) [sic] way as human translation, it may be the way to proceed” (ST202). This appeared in a few answers, e.g. another student admitted that “[i]f post-editing requires small changes, then it is great. If it requires me to change a lot, then it is double work” (ST207) and another said that MT allows translating more with less work (ST208). An interesting answer partially in favour of PE concentrated on the issue of choice (ST211): “[Post-editing] can ‘enforce’ one way of translating on the editor (i.e.

\(^3\) Statement 3 and question 4 were based on the background log question in Şahin (2014).

\(^8\) All direct quotations in this and subsequent sections are cited verbatim from questionnaires and marked with the anonymised code of the student, e.g. ST201, unless stated otherwise.
there can be so many things to change it’d either force the translator to deal with an awkward passage or start from the [sic] scratch.” Certainly, there is less freedom of choice in PE if the goal is to save time. However, as yet another student (ST103) pointed out, “[i]t may be useful when there is little time” and the already-mentioned ST106 admitted that it helps with the tediousness of translating common words.

Once the students expressed their general opinions concerning PE prior to the actual task, they also provided their quality expectations regarding the source text presented to them. The quality was anticipated to be: poor (ST102, ST109), poorer than human translation (ST101), low (ST106, ST210) or medium/average (ST104, ST201), not high (ST108, ST203), readable (ST204), mediocre (ST209). However, some of the participants thought the MT output will be of good quality or even acceptable (ST103, ST207, ST208, ST211): “A rather good quality of the output. This type of text is quite formulaic.” They predicted grammatical and syntactic mistakes, i.e. Polish word order (ST109) and problems connected with the distance between the two languages (“English and Polish do not belong to the same family of languages which may cause additional mistakes.” ST210). Also, on the one hand, they thought MT would help with terminology ST202. On the other hand, there were sceptical voices, i.e. “The terms in the text may be a problem for the machine translation” (ST209). Generally their answers meant that the students were not ready to take all that MT produced at face value and therefore they would double-check any terminological suggestions. The students were aware of limitations of the MT output and were expecting “(…) a few general guidelines concerned with the text in question” (ST203) and a general draft, rather than a publishable translation (ST212). Lack of metaphorical language or ambiguity was also stressed as a factor contributing to a better MT output.

4.2. Exercise 1.2

After the PE exercise, 20 participants filled in a post-task questionnaire in which they thought that MT was helpful in general as well as for other text of this type. They also expressed their preference either for PE or translation from scratch (cf. Figure 4).

![Figure 4](http://www.testsjournal.org)

**Figure 4.** Exercise 1.2: after post-editing. N=20. Error bars represent standard deviations. Statement 3 was based on chart 3 from Şahin (2014).

As the Figure illustrates, the students were enthusiastic about MT after this particular exercise, as on average they appeared to have preferred PE to manual translation (2.15 out of 5). The respective average scores for statements 1 and 2 only corroborate their enthusiasm. Furthermore, apart from providing general 1 – 5
scores, they also justified their preference. Those justifications confirm that they were mostly pleasantly surprised with the MT output quality, for instance:

Translating even this small piece from scratch would be very time-consuming. I didn't expect MT to be precise to such an extent as it was. I was prepared for a lot of editing and looking for terminology on my own. Surprisingly and fortunately to me, MT did that part for me and I only had to challenge a few of its choices. (ST206)

The answers also showed a positive attitude towards MT – e.g. that post-editing is easier and less time-consuming (ST105). Thus, once the students knew what to expect from MT, they were not unnecessarily disillusioned if the system performed badly. They found MT helpful (ST107), especially with terminology (which was a starting point for research according to ST102) and grammar was not difficult to correct (ST109), the mistakes were easy to notice (ST210). They were particularly happy with how fast they completed the task (ST207, ST206, ST106, ST209, ST210) and one student (ST106) emphasised that they were “quite surprised with the overall speed of translation” and concluded that MT proved more useful than they thought.

Those more sceptical, however, insisted that they “would rather translate text from scratch, because translations via machine can sometimes be misleading. Though MT output can come in handy on some stages of translation” (ST104). Another such opinion (ST108) maintained that manual translation is a better choice than correcting poor quality MT.

All in all, the participants were not hostile towards MT and PE both before and after the exercise and the mistakes they encountered did not bias them negatively towards MT. After the introduction into PE and the first exercise, the students proceeded with the second PE task during the next class.

4.3. Exercise 2.1

19 students participated in this part of the experiment. Before the task itself, students rated their familiarity with the text topic and the difficulty level of the text. This time there were two different texts to post-edit. 10 people post-edited the Paracetamol leaflet and the remaining 9 post-edited the NY Times article. Their scores were relatively similar, although the difference is not statistically significant (cf. Figure 5).
Figure 5: Exercise 2.1: before post-editing. N=19. Error bars represent standard deviations. Two-tailed t-test, assuming unequal variances. 1. $t=2.11$ with $p>.05$; 2. $t=2.14$ with $p>.05$. Both 1. And 2. were based on the background log question in Şahin (2014).

Apart from the general scores, the students also provided a longer answer concerning their MT quality expectations. Those about to post-edit the Paracetamol leaflet gave similar answers to the ones in the previous exercise, i.e. concentrating on the straightforwardness and simple language of the text (ST103), its formulaic nature (ST208), and the fact that MT provided a general translation outline (ST203). Their expectations oscillated between good quality (ST208, ST212, ST103), rather high (ST211), and worse than that produced by a human (ST101) with concerns about the syntax of the translation.

The group post-editing the NY Times article were divided in their quality expectations. That is, some insisted that the text is simple (ST210) and some claimed that metaphors could be problematic for MT (ST109). Others were convinced this text type is not suitable for MT:

> While machine translation can work quite well with instructions, etc., the journalistic article will be of poor quality. First of all, I expect that Polish inflectional endings will be incorrect. Second of all, machine translation will not translate metaphors or cultural references such texts are full of.

(ST207)

The students expected the quality to be lower than that of the previous text (ST107). Their predictions concentrated on syntactic problems, maintaining that the quality could be understandable/readable, but word-for-word translation would make “the text painful to read for a Polish native speaker” (ST204).

4.4. Exercise 1.2 vs. 2.2

After the second exercise, participants were presented with the same set of questions they answered previously after the task. Thus, Figure 6 provides the calculated mean of differences between scores from 1.2 and 2.2. As visible in the chart, the most drastic difference concerns the NY Times article group. The two-tailed paired samples $t$-test for the before and after scores for each text show that the difference is statistically significant for all statements concerning the NY Times article but for neither of the statements referring to the Paracetamol leaflet (cf. Figure 6).
Again, the students provided justification for their preference (PE or manual translation). When asked “What made you change/not change your mind about MT?” the two groups provided a range of reasons. The Paracetamol leaflet post-editors were again divided when it came to terminology. It was pointed out that “[t]he task was much faster thanks to the MT output as there was no need to research vocabulary or terminology” (ST102). Another student (ST103) explained the need to consult target language leaflets to become familiar with the register as MT did not help in this matter. Also, there were a few participants dissatisfied with the MT quality. They did not expect so many serious mistakes (ST208), thought that the MT output made the whole task longer than in Exercise 1 (ST206) and that some parts of MT were unreadable (ST211). However, most of the post-editors of the Paracetamol leaflet were happy with the MT output, thought it was helpful and sped up the translation (ST212, ST101, ST102, ST105, ST203, ST202).

As for the NY Times article, the post-editors thought that MT was helpful, although required many changes (ST104) as idioms as well as metaphors were mistranslated (ST106) and sentence structure was difficult for the system (ST107). Interestingly, despite admitting to having introduced numerous changes to the MT output, the students insisted that certain ideas were helpful (ST108, ST205). They also tried to approach the MT output for this text without bias, as one of the students explained that they thought this text would be easy for MT but multiple segments required amendments (ST209).

The last two statements that the students were asked to score on the Likert scale (1: strongly disagree – 5: strongly agree) were the ones featured in Figure 6. Despite the fact that the post-editors of the NY Times article were less enthusiastic about MT as a major translation help, the difference of mean scores between the two groups (the Paracetamol leaflet vs. the NY Times article) for both statements evaluating the usefulness of PE skills is not statistically significant (cf. Figure 6). In any case, it is very heart-warming that students recognise the need to become acquainted with PE apart from other translation activities.
4.5. Exercise 3

The last part of the PE exercise series was the simplified version of peer review as it did not involve a detailed error correction but was based on a general evaluation on the scale from 1 to 5. It was conducted via a Google questionnaire. The procedure is explained in section 3.3. The bars in Figure 8 below labelled The Paracetamol leaflet and NY Times article mean that the person evaluated a translation of this text post-edited by another classmate. The Figure shows that, although statistically insignificant, there is a slight difference in means in favour of the Paracetamol leaflet (except for statement 3).

Figure 8. Exercise 3: mean scores for compliance with TAUS guidelines. N=20. Error bars represent standard deviations. Two-sample t-test assuming unequal variances: p>.05. All statements were taken from the TAUS (2010) guidelines for full post-editing.
Similarly to previous questionnaires, the students were asked to explain why they thought the text they were evaluating was suitable or unsuitable for PE. Interestingly enough, there were voices advocating the newspaper article as suitable for PE, despite admitting better quality MT for the Paracetamol leaflet (ST206). Mostly, however, students from the group evaluating the NY Times article translation insisted that the literary style may be mishandled by MT (ST203) and even that “translating from scratch would be, in this case, less time-consuming and certainly less traumatising” (ST208). 5 out of 9 evaluators thought that this text was not suitable for PE and manual translation would be easier (ST212) since MT may not render the ambiguity and wordplay (ST202). The idioms were what one evaluator (ST101) thought as decisive to judge the text unsuitable for PE.

10 out of 12 evaluators of the Paracetamol leaflet said that the text was suitable for PE mostly referring to its formulaic language (ST102, ST207, ST204) and lack of metaphors (ST108, ST207). Two students dissatisfied with the MT quality and final translation insisted that the text would have been suitable, had the MT been trained with a large database of medical texts (ST106) and that the newspaper article was more suitable due to the lack of specialised vocabulary (ST205). One of the evaluators, however, was impressed with the syntax and terminology of the MT output (ST209) and another insisted that the MT required only some corrections (ST204).

5. Discussion and practical implications

The previous sections presented both the quantitative results of mean scores and a qualitative analysis of answers provided by the students which allowed to draw some conclusions with reference to the three research questions listed in section 3.1. When it comes to question 1 concerning the opinions before any exercises, i.e. “What are students’ attitudes towards machine translation?”, it is possible to conclude that they consider it a useful tool, at least to some extent. Outright negativity towards MT was in a minority. As far as question 2 is concerned, i.e. whether the attitude differed for the two text types (the Paracetamol leaflet vs. the NY Times article), it appears that the students were less enthusiastic about MT with the journalistic text. Finally, answers to question 3 concerning a change in the students’ attitude after a PE exercise can be summarised as follows:

- After Exercise 1: positive surprise
- After Exercise 2: some disillusionment
- After Exercise 3: reinforced feeling that formulaic texts are more suitable for PE (with exceptions)

Furthermore, what was particularly interesting, some students complained about the lack of choice that PE introduced. This is in line with Pym (2011), who described the main effects various computer aids have on translation: memory externalisation, the projection of the paradigmatic onto the syntagmatic, and, as a result, the disruption of linearity. In other words, various external tools may speed up the process; however, a wide range of aids may decrease efficiency as
well as curb intuitive choices. All this may explain the scepticism among some students. What was particularly interesting as well was the point that one of the students made during the summarising discussion. Namely, when asked whether they will continue to experiment with MT as a translation aid, they answered that using MT while still being a translation trainee would be cheating because they would not learn to translate this way.

In terms of the action research methodology, the naturalistic setting chosen for the experiment caused quite a few weaknesses of the study. Firstly, students are not always dependable. The exercises needed to be done in a certain sequence because otherwise the impact of different text types on MT was missing, e.g. when someone skipped the pre-task questionnaire and only filled in the post-task one or skipped one of the exercises. Secondly, technical difficulties are always a possibility during a class that depends solely on a cloud-based service (WFA).

As far as the practical implications of the socio-constructivist approach are concerned, learning through experience (Kiraly, 2012) was an effective choice because the students could determine their attitudes in the process of exploring them on their own. A tentative conclusion is that training (teaching technical CAT skills and guidelines) combined with educating and self-reflection (emphasising and developing the translator’s individual decision-making process) is effective. Allowing students to discover the contrast between different text types processed by MT had a powerful effect on them: empowered them to take responsibility for their own choices (cf. Kiraly, 2000) and enabled them to develop their own attitudes rather than follow blindly the instructions and opinions provided in the form of a lecture. Most importantly, the reflection on the source texts and predicting MT quality was an added value as students also had to employ their metalinguistic knowledge when they were analysing the texts before post-editing.

The experimental classes started with an informative presentation that eased the students into the topic but ultimately it was them who had to decide how they felt about using MT and for what purposes they would continue to do so. Thus, technical and formulaic language texts may be good to demonstrate when MT does not usually fail, as emphasised by Doherty (2016), but for the sake of contrast and sometimes even shock treatment, it is also good to show a machine-generated translation of a newspaper article or even a piece of literary text. Although a case study on PE of literary texts shows that statistical MT has the potential to help in literary translation if properly trained and applied in a suitable workflow (Toral & Way, 2015), using general-purpose MT for literary texts may make the output rather painful to edit, let alone read. All in all, incorporating PE into a CAT tools course was an interesting experience for the students.

Ideally, any future research involving language confidence scores would supplement the self-assessment of language proficiency with a more objective measure, e.g. a LexTALE test score (Lemhöfer & Broersma, 2012). Furthermore, a viable future research question concerns how the integration of PE into CAT tool education should proceed: either combined with TMs and other tools utilised in a CAT programme, such as SDL Trados, or as a standalone PE course which would allow to diversify the exercises and make students more comfortable with MT.
Appendix I: Questionnaires

1. Pre-task questionnaire (Exercise 1.1)

BEFORE YOU START POST-EDITING...

MT = machine translation
Please fill in this questionnaire after you read the text but before you start post-editing.
* Required

1) Your first and last name:* ..........................................................

2) Have you ever post-edited MT output before?*
   i.e. have you ever used machine translation in any of your translation assignments?
   * Mark only one oval.
   ☐ Yes ☐ No

3) Do you consider post-editing a meaningful way of obtaining translation?*
   Please explain why you think so ..................................................

4) Assess your experience as a translator*
   i.e. how confident you feel as a translator
   * Mark only one oval.
   1 2 3 4 5
   Not confident at all ☐ ☐ ☐ ☐ ☐ Very confident

5) Assess your experience as a user of English*
   i.e. how confident you feel in writing, reading, and speaking English
   * Mark only one oval.
   1 2 3 4 5
   Not confident at all ☐ ☐ ☐ ☐ ☐ Very confident

6) Assess the difficulty level of the source text*
   * Mark only one oval.
   1 2 3 4 5
   Very easy ☐ ☐ ☐ ☐ ☐ Very difficult

7) How familiar are you with the topic of the text?*
   * Mark only one oval.
   1 2 3 4 5
   Not familiar at all ☐ ☐ ☐ ☐ ☐ Very familiar

8) What is the level of quality that you expect from the machine translation for this text? Explain the reasons.*
   ..................................................................................

2. Post-task questionnaire (Exercise 1.2)

AFTER THE POST-EDITING EXERCISE:

MT = machine translation
Please fill in this questionnaire after you finished post-editing and submitted the finished translation via Moodle.
* Required
1) Your first and last name:*
................................................................................

2) MT output was helpful.*
Mark only one oval.

1 2 3 4 5
Strongly disagree Strongly agree

3) MT output can be helpful for other texts of this type*
Mark only one oval.

1 2 3 4 5
Strongly disagree Strongly agree

4) I would rather translate a text from scratch than post-edit MT output*
Mark only one oval.

1 2 3 4 5
Strongly disagree Strongly agree

5) Explain your answer to the previous question.*
.............................................................................

3. Pre-task questionnaire (Exercise 2.1)

POST-EDITING: CONTINUED (BEFORE THE EXERCISE)

Fill in after you read the text but before you start post-editing.
* Required

1) Your first and last name:*
................................................................................

2) Assess the difficulty level of the source text*
Mark only one oval.

1 2 3 4 5
Very easy Very difficult

3) How familiar are you with the topic of the text?*
Mark only one oval.

1 2 3 4 5
Not familiar at all Very familiar

4) What is the level of quality that you expect from the machine translation for this text? Explain the reasons.*
.............................................................................

4. Post-task questionnaire (Exercise 2.2)

POST-EDITING: CONTINUED (AFTER THE EXERCISE)

MT = machine translation

Fill in after you finished post-editing and submitted the finished translation.
* Required

1) Your first and last name:*
................................................................................

2) Which text did you post-edit?*
Mark only one oval.

☐ Paracetamol 500mg tablets
☐ The Plot Twist: EBook Sales Slip, and Print Is Far From Dead
3) MT output was helpful.*
Mark only one oval.

1 2 3 4 5

Strongly disagree Strongly agree
4) MT output can be helpful for other texts of this type*
Mark only one oval.

1 2 3 4 5

Strongly disagree Strongly agree
5) I would rather translate a text from scratch than post-edit an MT output*
Mark only one oval.

1 2 3 4 5

Strongly disagree Strongly agree
6) What made you change/not change your mind about MT?*
Mark only one oval.

...............................................................................

7) Machine translation can ease the burden of translators to a great extent.*
Mark only one oval.

1 2 3 4 5

Strongly disagree Strongly agree
8) Post-editing is a skill that should be included in translation education.*
Mark only one oval.

1 2 3 4 5

Strongly disagree Strongly agree

5. Peer-review (Exercise 3)

EXERCISE 3: EVALUATING A POST-EDITED TEXT

MT = machine translation
Download and read the text assigned to you and answer the following questions:
* Required

1) Your first and last name:*
................................................................................

2) Code of the text assigned to you:*
................................................................................

3) The target text is a grammatically, syntactically and semantically correct translation.*
Mark only one oval.

1 2 3 4 5

Strongly disagree Strongly agree

4) The key terminology is correctly translated.*
Mark only one oval.

1 2 3 4 5

Strongly disagree Strongly agree

5) No information has been accidentally added or omitted.*
Mark only one oval.

1 2 3 4 5

Strongly disagree Strongly agree
6) The target text does not contain any offensive, inappropriate or culturally unacceptable content.*
Mark only one oval.

Strongly disagree 1 2 3 4 5

6. How much of the raw MT output is preserved in the final post-edited text?*
Mark only one oval.

Strongly disagree 1 2 3 4 5

1) The post-edited text complies with basic rules regarding spelling, punctuation and hyphenation.*
Mark only one oval.

Strongly disagree 1 2 3 4 5

2) Is this text type suitable for post-editing?*
Explain with reference to the text you post-edited yourself in the previous exercise and your general impression of this text.

3) Do you have any comments on the text?*
Has anything struck you as particularly unusual? You may provide examples from the text.

References


