Teaching machine translation and post-editing

Dorst, A.G.; Sterkenburg, P. van; Bonth, R. de; Heylen, K.

Teaching Machine Translation and Post-Editing

Aletta G. Dorst (Leiden University Centre for Linguistics)

This article reflects on the current training in machine translation and post-editing that students enrolled in the Master Translation at Leiden University Centre for Linguistics are receiving by discussing the findings of three recent studies. Especially at the master's level, machine translation and post-editing are becoming increasingly important in translator training curricula with an eye to changing workflows in the language industry. Yet one question that remains is which specific skills, competences and knowledge future translators need, and to what degree an understanding of the computational side of machine translation is required.

Introduction

Society is becoming increasingly globalized and digitized. For most people, the use of machine translation websites and mobile phone apps has become a normal routine (Bojar et al. 2016). There's also no denying that the language industry has fully embraced machine translation and post-editing as a fast and cheap way to boost productivity, especially since the launch of Neural Machine Translation (NMT) in 2015 and subsequent claims about NMT achieving human parity (Hassan et al. 2018). While there is currently no evidence that human translators will soon all be replaced entirely by machines (Way 2019), there are growing concerns about the ethical implications of the widespread adoption of machine translation as well as about unrealistic expectations and common misconceptions (e.g. Vieira 2020; Moorkens 2022). In addition, scholars have been reflecting on how to integrate machine translation and post-editing into translator training curricula (Bowker 2002; Doherty and Moorkens 2013; Doherty and Kenny 2014; Guerberof Arenas and Moorkens 2019; O'Brien 2002; Pym 2014). Knowing how to use machine translation effectively is recognized as an essential competence for future translators (EMT Competence Framework 2009, 2017, 2022; Roth-

well 2019), and students more generally (Bowker 2020; Dorst, Valdez and Bouman 2022; Loock, Léchauguette and Holt 2022).

Machine Translation and Post-Editing in the MA Translation at Leiden University

Students admitted to the programme typically have a bachelor's degree in languages and cultures (especially English, Korean and Japanese), with a completed minor in translation, including Dutch to English and English to Dutch translation. Most of them have little to no knowledge of computational linguistics or computer science. Machine translation and post-editing are part of the obligatory 5-EC technology course *The Translator's Tools* in the one-year (60 EC) full-time master's degree in Translation at Leiden University. The course is taught in the first semester, alongside an obligatory 5-EC theory course (*Translation Studies*), an obligatory 5-EC practical course in English to Dutch and Dutch to English translation (*Advanced Translation*), a 10-EC specialisation course in either Legal, Literary, Medical or Multimodal Translation, and an obligatory 5-EC thesis seminar.

In *The Translator's Tools* students are introduced to a variety of tools and technologies, including Translation Memories (Trados and MemoQ), Termbases (Multiterm), corpus tools (AntConc and SketchEngine), and Machine Translation (especially Google Translate, DeepL, and e-Translation). The course takes the form of 13 practical sessions of 90 minutes – 45 minutes lecture and/or instruction followed by 45 minutes practice – in the Translation Lab, which has dedicated laptops and software as well as related hardware (e.g. external keyboard, mouse, headphone). Assessment for the course involves two practical in-class assignments (50% of course grade) with CAT-tools (one in Trados, one in MemoQ) and a research paper (50% of course grade) on a machine translation post-editing research project.

At present, post-editing is still done in Microsoft Word, using 'track changes' and comments, though current professional practice is rapidly moving towards the use of MT-plugins in CAT-tools and using interactive and adaptive forms of machine translation. However, it is our experience that the students are struggling to understand all of the different tools and technologies and what each does, and using integrated suites blurs these lines even more (i.e. the difference between matches from the Translation Memory and 'matches' from the Machine Translation engine) and obscures the translator's role and agency in the process. We are also not currently training students to do quality assessment

using metrics, since they lack the necessary background in programming and computational linguistics. In our view, one central question is whether all translators actually need to understand the computational side, and to be able to customize engines and run metrics, or whether for most translators it will be sufficient to develop competences in doing post-editing and using machine translation without acquiring a full understanding of how machine translation works. The three studies discussed in the remainder of this article will demonstrate areas where we have found students need more machine translation and post-editing training, especially in terms of comprehension.

Study 1 A survey on how, when and why students use machine translation

Our first investigation (Dorst et al. 2022) concerned a survey examining how, when and why students at Leiden University's Faculty of Humanities use machine translation. Using an online questionnaire (conducted between 26 January and 28 February 2021 via Qualtrics), we determined which websites and apps students use and for what purposes. We also gauged their awareness of issues concerning privacy, academic integrity and plagiarism. In total, 283 Humanities students filled out the questionnaire (average age 22.3), most of whom were enrolled in a bachelor's programme (86.2%).

In line with other surveys on student usage of machine translation (e.g. Loock, Léchauguette & Holt 2022), we found that students are highly familiar with and have a clear preference for Google Translate: while 96.8% of the respondents indicated being very or extremely familiar with Google Translate, they indicated that they were not at all familiar with Google Translate's main competitors, DeepL (82.3%) and Bing/Microsoft (72.8%). Since the output for each engine will be different, this is one reason why we have students use different engines and compare the output for the technology course in the master. In terms of machine translation literacy, one thing students have to learn is that it does in fact matter which engine you use. CAT-tools like Trados and MemoQ work the same way and it does not matter which one you use to align your texts, create the translation memory and translate the file – the tool itself does not affect the translation. But your translation will be different depending on which machine translation plugin, website or app you choose, and the quality may be considerably worse for particular language combinations and domain-specific texts.

Also in line with previous research, we found that students use machine translation websites and apps primarily as an alternative to a dictionary to look up isolated words and phrases. Most of the respondents reported that they used

machine translation like a bilingual dictionary to translate single words, idioms or expressions (mentioned 157 times), followed by using it like a thesaurus to find synonyms (mentioned 25 times), and using it like a monolingual dictionary to understand the meaning of an unknown word or to check the spelling (mentioned 10 times). With regard to the usefulness of free machine translation websites and apps for educational purposes, some of the reasons students mentioned included: to look up single words (mentioned 86 times), to understand the main idea behind academic texts (mentioned 40 times) and to improve their own produced content (mentioned 18 times). A closer look at when and how students used machine translation for their studies showed that students primarily used it to read or understand articles (mentioned 103 times) and to write papers or do homework (mentioned 78 times). Taken together, this shows that students are not making optimal use of machine translation for either gisting purposes or production purposes, since looking up isolated words and phrases deprives the engine of the context it needs to determine the right translation.

These results demonstrate a clear need for more machine translation literacy for students in general, not only those enrolled in translation of language degrees. Students don't necessarily need training in *how* to use MT, but there is room for improvement in terms of *when* and *why* they use it. The results of this study were used to develop a series of infographics and animated videos explaining central topics in machine translation as part of an NRO Comenius Senior Fellow project (The value of machine translation in the multilingual academic community, grant number 405.20865.446). These materials are available on an open access resource website for Machine Translation Literacy: universiteitleiden.nl/machinetranslationliteracy.

Study 2 A survey on how students conceptualize machine translation

In a second survey (Salmi et al. 2023), carried out in collaboration with two Finnish universities (Turku and Eastern Finland) who are also members of the EMT Network, we investigated how students conceptualize the processes involved in machine translation, and how they understand the similarities and differences between human and machine translation. The survey (available in Finnish, Dutch and English) was distributed to two groups of bachelor's students (one Finnish, one Dutch) and two groups of master's students (one Finnish, one Dutch) who were taking a course on translation. In total, 58 students took part in the study, 25 from Finland and 33 from the Netherlands. Students filled out the questionnaire during tutorial, after they had received a brief introduction to

the history and basics of machine translation, including an overview of the three main types of machine translation (rule-based, statistical and neural). They were asked to reflect on their understanding of how machine translation engines work and how humans translate and answer the following questions: Do humans translate in the same way machines do? If yes, what is similar about translating? If not, in what way is a human translator different from a machine?

The results show that students often describe both human translation and machine translation in similar terms, which suggests that they do not sufficiently distinguish between them and use human descriptors to describe machine processes, either because they do not fully understand how machine translation works or because they lack the terminology to describe machine translation. While 38 out of 58 students thought human and machine translation are different, and an additional 13 students thought human and machine translation have both differences and similarities, the students found it hard to explain how they are different, and some of their explanations revealed misconceptions. A closer examination of the characteristics students mentioned to explain their similarity judgments indicate that students understand the differences between human and machine translation predominantly through the characteristics that are typical of human translators. They show a very limited understanding of machine characteristics, and of the four characteristics that were clearly associated with machines – *Uses probabilities*, *Translates directly*, *Always translates the same way*, and *Uses logic* – the characteristic *Always translates the same way* is technically not true for neural machine translation, and *Uses logic* is ambiguous in terms of whether they mean mathematics or common sense.

Looking at some of the explanations students proposed, there also appears to be a lack of accurate terminology to describe their understanding. For example, T12 argued that 'Humans and machines, translation memories for example, both explore their prior knowledge and try to find the correct equivalents of words in the target language' (translated from Finnish). This indicates that the student either does not grasp or does not have the right terminology to describe how the type of pre-defined knowledge that machines and humans use are fundamentally different, and there appears to be confusion between machine translation and translation memories. This may therefore reflect either a conceptual misunderstanding or a vernacular misconception based on the idea of machines and humans seeking information in a 'memory'. Similar problematic explanations were found using the word 'database', since, in a technical sense,

neither human translators nor neural machine translation engines retrieve previous translations from a database the way translation memories do.

This study illustrates that teaching students machine translation is not a straightforward task and that students may have conceptual and vernacular misconceptions in their understanding of how machine translation works. While most research is focused on how to train students in using different systems and in doing post-editing in different genres, there appears to be a lack of attention to determining how students understand the processes involved in machine translation and whether misconceptions affect either their usage of the technology or their perception of its value. If students are struggling to conceptualize the process of machine translation this may also make it harder for them to understand their own role and agency in this process.

Study 3 An experiment investigating whether training in post-editing affects creativity

In this experiment (Guerberof Arenas et al., under review), eleven students from two universities (Groningen and Leiden) translated and post-edited three literary texts distributed on the first and last days of their translation technology modules. In the source text excerpts, all taken from Hemingway's collection of short stories *In Our Time*, units of creative potential (Guerberof Arenas & Toral Ruiz 2022) – such as cultural references, idioms and metaphors – had been identified to determine whether training in post-editing had a positive impact on students' creativity during literary post-editing. After completing the assignment, the students were asked to write a short reflective essay (max. 300 words) in response to a number of questions on their opinion about the quality of the output, the nature of the changes they made, and whether they liked using machine translation.

The results showed that while students liked using machine translation and felt more confident in doing post-editing after training, there was no quantitative evidence to conclude that the training significantly affected students' creativity; that is, they were not more creative in their decisions after training than before training. They made very few changes to the machine translated output, they mislabelled the errors they corrected, and many of the changes they made were unnecessary (as in, they did not correct an actual error). However, after the training, we did observe a change in the quantitative data: the students felt more confident to try creative shifts and they felt more confident to make changes to the output. Yet we also saw that they made more errors themselves.

We found that the students were more creative when they translated themselves (= human translation) but they also made significantly more errors in human translation than in post-editing, especially before training.

This experiment taught us that students may need more focused training in applying creative shift when doing post-editing, especially for creative domains like literature. That is, when doing creative or literary post-editing, they need to be trained specifically how to resist accepting output that is indeed grammatically and semantically correct in order to do more creative rewriting rather than error correction, so as to ensure literary quality and similar stylistic effects and narrative engagement. While many students appeared to lack the required linguistic competence to identify and correct errors, there appears to also be a clear need to train post-editors in stylistic analysis and creative (re) writing skills.

Conclusion

This article has discussed three studies on machine translation and post-editing in light of the current EMT Competence Framework and technology training in Leiden University's master's in Translation. The studies show that the focus of training in machine translation and post-editing may need to be shifted more towards comprehension than technical competences, and that the skills required to do domain-specific post-editing may need to be defined more clearly. For our own training, we feel that ideally students would be offered the option to choose between post-editing courses that focus on doing post-editing for different domains and text types (e.g. literary post-editing versus medical or legal post-editing, or post-editing of automated subtitles) without going into the technical details of machine translation, and machine translation courses that focus on programming, customization, training and testing, and running metrics for those that are not afraid to tackle the computational side of translation technology.

Bibliography

Bojar, B., Chatterjee, R., Federmann, C., Graham, Y., Haddow, B., Huck, M., Jimeno Yepes, A., Koehn, P., Logacheva, V., Monz, C., Negri, M., Névéol, A., Neves, M., Popel, M., Post, M., Rubino, R., Scarton, A.S., Specia, L. & Turchi, M. et al. (2016), 'Findings of the 2016 Conference on Machine Translation', in: *Proceedings of the First Conference on Machine Translation: Volume 2, Shared Task Papers*. Berlin, Germany: Association for Computational Linguistics, 131–198. <doi.org/10.18653/v1/W16–2301>

Bowker, L. (2002), *Computer-Aided Translation Technology: A Practical Introduction*. University of Ottawa Press.
Bowker, L. (2020), 'Machine Translation Literacy Instruction for International Business Students and Business English Instructors', in: *Journal of Business & Finance Librarianship*, 25(1–2), 25–43. <doi.org/10.1080/0896.1794739>.
Doherty, S. & Kenny, D. (2014), 'The Design and Evaluation of a Statistical Machine Translation Syllabus for Translation Students', in: *The Interpreter and Translator Trainer,* 8(2), 295–315. <doi.org/10.1080/1750399X.2014.937571>
Doherty, S. & Moorkens, J. (2013), 'Investigating the Experience of Translation Technology Labs: Pedagogical Implications', in: *Journal of Specialised Translation,* 19, 122–36.
Dorst, A.G., Valdez, S. & Bouman, H. (2022), 'Machine Translation in the Multilingual Classroom: How, When and Why Do Humanities Students at a Dutch University Use Machine Translation?', in: *Translation and Translanguaging in Multilingual Contexts*, 8(1), 49–66. <doi.org/10.1075/ttmc.00080.dor>.
EMT Expert Group (2009), *Competences for Professional Translators, Experts in Multilingual and Multimedia Communication*.
EMT Expert Group (2017), *European Master's in Translation—EMT Competence Framework 2017*.
EMT Expert Group (2009), *Competences for Professional Translators, Experts in Multilingual and Multimedia Communication*.
Guerberof Arenas, A. & Moorkens, J. (2019), 'Machine Translation and Post-Editing Training as Part of a Master's Programme', in: *Jostrans The Journal of Specialized Translation,* 31, 217–38.
Guerberof Arenas, A. & Toral Ruiz, A. (2022), 'Creativity in translation: machine translation as a constraint for literary texts', in: *Translation Spaces* (Online First). <doi.org/10.1075/ts.21025>
Guerberof Arenas, A., Valdez, S. & Dorst, A.G. (under review). 'Does training in post-editing affect creativity?' Submitted to *JoSTrans41*: Special Issue on Translation Automation and Sustainability.
Hassan, H., Aue, A., Chen, C., Chowdhary, V., Clark, J., Federmann, C., Huang, X., Junczys-Dowmunt, M., Lewis, W., Li, M., Liu, S., Liu, T.Y., Luo, R., Menezes, A., Qin, T., Seide, F., Tan, X., Tian, F., Wu, L., Wu, S., Xia, Y., Zhang, D., Zhang, Z. & Zhou, M. (2018), 'Achieving Human Parity on Automatic Chinese to English News Translation', in: *Computing Research Repository*. <arxiv.org/abs/1803.05567>
Loock, R., Léchauguette, S. & Holt, B. (2022), 'Dealing with the 'Elephant in the Classroom': Developing Language Students' Machine Translation Literacy', in: *Australian Journal of Applied Linguistics*, 5(3), 118–134. <10.29140/ajal.v5n3.53si2>
Moorkens, J. (2022), 'Ethics and machine translation', in: *Machine translation for everyone: Empowering users in the age of artificial intelligence*, Translation and Multilingual Natural Language Processing, 121–140. Language Science Press, Berlin.
O'Brien, S. (2002), Teaching Post-Editing: A Proposal for Course Content', in: *6th EAMT Workshop 'Teaching Machine Translation' 14–15 November 2002*, 99–106.

Pym, A. (2014), 'Translation Skill-Sets in a Machine-Translation Age', in: *Meta*, 58(3), 487–503. <doi.org/10.7202/1025047ar>
Rothwell, A. (2019), 'Tracking Translator Training in Tools and Technologies: Findings of the EMT Survey 2017', in: *Journal of Specialised Translation*, 32, 26. <jostrans.org/issue32/art_rothwell.pdf>
Salmi, L., Dorst, A.G, Koponen, M. & Zeven, K. (2023), 'Do Humans Translate like Machines? Students' Conceptualisations of Human and Machine Translation', in: *Proceedings of the EAMT conference 2023*, Tampere (Finland).
Vieira, L.N. (2020), 'Machine translation in the news: A framing analysis of the written press', in: *Translation Spaces*, 9(1), 98–122.
Way, A. (2019), *Should Human Translators Fear the Rise of Machine Translation?* <adapt-centre.ie/news/should-human-translators-fear-the-rise-of-machine-translation>