

Know thine enemy

Michael Farrell picks up the perennial topic
of machine translation post-editing

There is no hiding it: the original idea of machine translation (MT) was to replace human translators completely. Way back in the 1950s (yes, MT is that old – older actually), when the first computer translation systems came into being, some of the researchers working on them were predicting that translators – or at least technical translators – would be gone within a matter of years.

Predictions and reality for the machines

However, other experts also realised that there were problems with bringing the MT output produced at the time up to human-quality levels. In 1951 John Edwin Holmström (incidentally the grandfather of the present *Bulletin* editor) produced a highly sceptical report for UNESCO, declaring that MT output ‘would be atrocious and fuller of “howlers” and false values than the worst that any human translator produces’. The following year, at the first International Conference on Mechanical Translation, Erwin Reifler proposed both a pre-editing and post-editing stage in the process to solve the problem. Conference organiser Yehoshua Bar-Hillel went even further and said that post-editing would always be necessary because fully automatic high-quality translation (FAHQIT) was unattainable ‘not only in the near future but altogether’, since machines don’t have real-world knowledge.

In an appendix to his report, he gave a demonstration of its non-feasibility. (Every year, I encourage my students to launch a petition to ask the rector of the university where I lecture to name one of the computer labs after our *hero*, but they never do: I think they think I’m only joking.)

At the time, it looked as if the predominant translation process was destined to become MT plus post-editing, possibly preceded by a pre-editing stage to eliminate some of the ambiguities in the original text, particularly if a document needed translating into several languages. Indeed, some good source language pre-editing upstream can save time during the several target language MT post-editing (MTPE) passes downstream.

So why wasn’t the market of that period suddenly flooded with MTPE assignments? Technology aside – even when PCs became widespread, MTPE jobs were still few and far between – the real reason was that, until recently, raw MT output was just not good

enough. The increase in productivity from MTPE has to be large enough to compensate for the time and money needed to train and look after an MT engine and produce raw output. If the productivity gain is too small, it is just not worthwhile.

In reality, MTPE was adopted early in some limited domains where it was possible to train custom MT engines. One very good example is the much-cited case of the METEO system in Canada, developed in the 1970s and used by Environment Canada to translate weather forecasts between English and French from 1982 to 2001. Some commentators state that the system achieved a 95 per cent quality level, clearly making MTPE a feasible approach in this case.

However, for a few years now, things have been gradually changing, and an increasing number of post-editing assignments are surfacing on the market. What happened?

The answer is neural networks. We have to be honest: the quality of raw MT output *has* got better and – at least in some cases – the productivity gain from MTPE justifies the method.

Present practice using machines

Professional translators are sometimes accused of being late technology adopters: we tend to *endure* changes in working methods rather than pioneer them. Early CAT tools were not particularly user-friendly from the translator’s point of view, and – to some extent – some still aren’t even today. They (rightly) often contain a lot of project management functions, which are of limited use or even an encumbrance to individual freelance translators and which some feel are a downright impediment. The figures suggest that something similar is happening with MTPE.

I conducted a recent survey among professional translators and found that just over 53 per cent of translators stated they sometimes or often accept MTPE assignments from language service providers (LSPs) and clients. However, only just under 19 per cent of the nearly 70 per cent who responded that they use MT in some way said that they adopt the MTPE workflow model when doing their own translations. The vast majority prefer to use MT in a wide variety of other ways.

Yet we don’t need to implement any new technology to transform a translation assignment into an MTPE job. Most translators today use CAT tools (just over 86 per cent, according to the previously mentioned survey). So all we need to do is pre-translate the text with MT turned on, and Bob’s your uncle.

Why do so many
translators not treat MTPE
as the best approach?



It is a known fact that MT does a better job with precisely the same kinds of documents that lend themselves to processing with CAT tools

So why don't most of us do that? After all, we evolved as a tool-using species (it's one of the qualities that differentiates us); and most translators today see MT not as a competitor but as a tool which can be used not just to increase productivity – perhaps not even *primarily* to do that – but also as a source of *inspiration* or simply as a kind of concordancer or sophisticated dictionary. Some might argue that the real reason we don't do MTPE is our lack of familiarity with using our tools in this way – but then there is that significant percentage of translators who also accept MTPE assignments, and who have

PHOTO © PEKELS/TARA WINSTEAD

MT output tends to limit the vocabulary choices to only a subset of the various solutions that human translators are capable of finding

presumably done some specific training (one would hope, at least). So, again, why do so many translators not treat MTPE as the best approach?

There's no specific research into this, but as someone who lectures in this field I can make some informed guesses. It may depend on the types of texts they translate. It is a known fact that MT does a better job with precisely the same kinds of documents that lend themselves to processing with CAT tools: boilerplate-style manuals, sales contracts and anything that needs consistent terminology and sentences written to a formula. It does not do so well



with texts that need to be engaging and intellectually stimulating, in fields such as marketing, advertising, literature, journalism, education, entertainment, humour and creative writing in general. Humour, in particular, thrives on ambiguity, and ambiguity is MT's worst enemy (after all, it is particularly challenging for human translators too).

But in addition, it would seem that some translators just plain don't like MTPE. In the survey mentioned above, one respondent defined it as *soul-destroying*. One reason for this might be – again I speculate – that unrealistic, unsustainable requests are being made of post-editors (see Jenny Toal's feature in the November-December 2022 *Bulletin*). Can a highly educated professional linguist be asked to transform themselves into a *correction robot*?

There are plenty of studies that show that MTPE leads to higher productivity. These studies generally take a text and compare the time needed to translate it with how long it takes to post-edit the same text after MT. What they don't do, as far as I know, is measure whether this higher productivity is sustainable in the long term. If job satisfaction is virtually zero, the rates unrewarding and MTPE mind-numbingly boring, surely our post-editors won't last a week.

MTPE without translation knowledge

Another interesting question we can ask is, 'Can someone without the training and linguistic knowledge of a professional translator be asked to post-edit?' Some research has been done into monolingual post-editing. The idea is that a suitably educated and trained native speaker of the target language can correct the errors in raw MT output without needing to understand the original source text. It has even been posited that this way of working may have some advantages, since the post-editors in this case could be experts in the subject the translation is about. As a counter-argument, it would obviously be an advantage if they were bilingual experts in their field – possibly with some experience of translation – and did bilingual post-editing. But this would of course cost more. And let's be honest: at the end of the day, post-editing is all about cutting costs.

In any case, neural MT (NMT) has put paid to the monolingual post-editing dream, and large language models (LLMs), such as ChatGPT, have put a further nail in its coffin. With the older technologies, like rule-based or statistical MT, when a translation was wrong, it was pretty obviously wrong, sometimes making virtually no sense at all. In this case, a non-linguist could most likely identify the mistakes, although they might not always have enough information to put them right without being able to read the source language.

NMT output is more fluent, often producing perfectly grammatical sentences with no apparent errors...except for the fact that they sometimes do not actually mean exactly the same thing as the original. There is no sure-fire way of working out what is wrong without understanding the source language and, what is more, it is advisable to read the original text first before the raw MT output in order not to be influenced by the MT version and perhaps miss the possibly only subtle differences. It has been noted that LLMs are sometimes capable of *hallucinating* and writing perfectly plausible things which are actually utterly false. So, if generative artificial intelligence (GenAI) is used to machine-translate, a bilingual human will still be needed to check

the translation against the source text.

Another phenomenon which there has been plenty of research into is *post-editeuse*; that is characteristics of post-edited MT output that distinguish it from human-translated text. One of these features, namely lexical impoverishment, stems from the fact that MT output tends to limit the vocabulary choices to only a subset of the various solutions that human translators are capable of finding. The post-editors are then inevitably primed by the raw output and tend only to partially restore the variety. They may even be given post-editing guidelines that explicitly instruct them to use as much of the MT output as possible, rather than bring the terminological diversity up to human level.

It is important to realise that a lack of diversity in the choice of words is not always a bad thing. If a control on the front panel of a washing machine is called a *button* on the first page, a *switch* on the second page and a *selector* on the third, the poor washing machine user would be pardoned for not being particularly amused.

But it is easy to imagine how boring a piece of creative writing would be with virtually no synonyms. Sticking closely to the MT output is also not necessary for productivity. I have personally carried out an experiment that showed that productivity increases with or without this guideline, which means it is not only inadvisable but also unnecessary (especially, of course, when lexical variety is a quality factor).

Having said that, there is some evidence that this phenomenon is attenuating and that it may not be an important problem in the future since raw MT output is becoming more lexically varied from the outset. Moreover, it is possible to prompt LLMs to automatically post-edit the raw output from other MT engines and increase the range of lexical solutions found in the translated text.

So, what is the conclusion? Are translators doomed? Will MTPE become the predominant approach to translation in all but just a very limited number of genres, particularly with the advent of GenAI? In my opinion, the jury is out. It is very hard to predict how things will go. However, it is important to keep our eyes open and be ready to find ways of using all new technology to our advantage, as a tool, and to keep the human in the loop.

It has been noted that LLMs are sometimes capable of hallucinating and writing perfectly plausible things which are actually utterly false

ITI has a new 'Artificial Intelligence' page on the website where we talk about the actions that ITI is taking in response to evolving technology.



Michael Farrell is primarily a freelance translator and transcreator. Over the years he has acquired experience in the cultural tourism field and in transcreating advertising copy and press releases, chiefly for the promotion of technology products. Being a keen amateur cook, he has also translated texts on Italian cuisine. He is an untenured lecturer in post-editing, machine translation and computer tools for translators at IULM University, Milan, Italy; the developer of the terminology search tool IntelliWebSearch; a qualified member of the Italian Association of Translators and Interpreters (AITI); a member of Mediterranean Editors and Translators (MET) Council; and the author of *A guide to machine translation for today's professional translator* (ISBN: 979-8375879758).