Author accepted manuscript. Please cite as: Kenny, Dorothy (2025) 'Literary machine translation: from taboo to controversy' in Stefan Baumgarten and Michael Tieber (eds) *The Routledge Handbook of Translation Technology and Society*. Routledge, 381-395. 10.4324/9781003271314-33

LITERARY MACHINE TRANSLATION: FROM TABOO TO CONTROVERSY

Dorothy Kenny

INTRODUCTION

The use of machine translation (MT) in the creation of literary texts is a relatively recent phenomenon. For much of the history of MT, literary texts were simply considered beyond the reach of the technology, with Arnold et al. (1994: 6), remarking that "[t]he criticism that MT systems cannot, and will never, produce translations of great literature of any great merit is probably correct." Whether it is their characteristic formal features that make literary texts difficult for machines to translate is a matter of debate. Rothwell et al. (2024: 2-3), for instance, reject this idea, while others home in on precisely those features of literary narratives in particular that appear to make them less amenable to automatic treatment than other kinds of text. For other commentators, the argument that literature cannot be translated by machines is based on a different premise, namely that literary translation is a uniquely human endeavour, one that demands human creativity and interpretive skills. This leads some defenders to deny machines any translation ability, recasting what computers do as something other than translation. As the European Council of Literary Translators' Associations (CEATL 2023) puts it, "Machines are not translators but 'translatoids'. They do not translate; they generate textual material." Given the current backlash against generative artificial intelligence (GAI), however, the greatest contemporary concern is not with how one defines translation, but rather with how one secures the legal, regulatory and economic conditions that might underpin safe, fair and sustainable uses of GAI in translation and other creative practices. The year 2023 saw concerted efforts among authors and translators to challenge the modus operandi of certain entities promoting GAI and to expose practices that serve to undermine translation in the publishing industry. Manifestos and statements issued by PEN America (2023), CEATL (2023) and the European Writers' Council (2023) exemplify such efforts.

In this chapter, I survey relevant research on literary MT, providing brief historical and technical context, before homing in on potential consequences of increased automation in the translation of literature, and reflecting on how the above-mentioned backlash may influence future developments. My survey of literary MT focuses on sources written since the advent of

data-driven and especially neural MT (NMT, see below), drawing on a typology presented in Kenny and Winters (2024). The discussion of the consequences of literary MT is divided roughly into those that affect translated texts, the translation process, the reading experience, the translation market, and the translator, although it is acknowledged that these factors inevitably overlap.

THE EARLY DAYS OF LITERARY MT

Ironically, it was the inability of early MT to rise to the challenges presented by literary texts that first attracted some creative writers to the technology. In the days of what would be conventionally characterized as 'low-quality' rule-based MT, such writers prized MT output for its jarring, ungrammatical, nonsensical, antinarrative character (see Kenny 2020), its tendency to produce "unexpected new meanings" through "unpredictable" divergences from source texts (Lee 2011: 94-95), and its propensity to create unintended comic effects (Kenny ibid). But when data-driven MT entered the fray in the guise of statistical machine translation (SMT), things began to change: the application of MT to literature no longer served the pursuit of "avant-garde aesthetic ideals" (Taivalkoski-Shilov 2019a: 693). Rather, computational linguists began making serious efforts to develop MT systems that would generate 'accurate' and 'fluent' translations of literary texts, to cite the two signature concepts of MT evaluation (Koehn 2020: 45-47). They were motivated to do so by the inherent challenges of the task, "the aim of the translation being not just to communicate meaning but to replicate the source text's stylistic and textual effects on the reader" as Toral and Way (2015a: 243) put it, and encouraged by the undoubted quality improvements heralded first by SMT (see Voigt and Jurafsky 2012: 18) and then by NMT (e.g., Toral and Way 2018; Webster et al. 2020: 1). But there was also an element of opportunism involved. Toral and Way (2015a: 247), for example, noted that the emergence of the ebook meant that novels were now available in digital format and thus amenable to the kind of processing that would allow them to be used as training data in an SMT system. Elson et al. (2012) made similar observations about the increased supply of 'older' works in digital form thanks to initiatives such as Project Gutenberg.

Toral and Way (ibid: 241) balanced such opportunism with apparent altruism, noting that:

> human translators are *extremely* poorly paid for their work on translating literary texts […]. Accordingly, if SMT could be shown to be of use in this domain, then

> MT would be capable of helping these poorly paid literary translators to become more efficient, and make more money. (emphasis in original)

As we shall see below, it is not self-evident that human translators become more efficient when using MT to create first drafts of their literary translations, or that publishers would pass on any cost savings attributable to MT use to human translators in the first place, but this passage from Toral and Way (2015a) nonetheless shows that, from the outset, computer scientists were aware that their work on MT could have consequences for translation markets. Other computer scientists who have speculated about possible economic, social and cultural benefits of literary MT include Matusov (2019: 18), who suggests that the technology could "help to make largely undiscovered foreign language books instantly available online to readers worldwide" and assist publishers in their selection of material for professional translation "thus promoting an increased circulation of high-quality work among different languages and cultures" (ibid.), and also Oliver et al. (2022), who propose the use of literary MT to create bilingual e-books as a support for learners of the source language.

Alongside the utopian aspirations of much of the work of the pioneers in literary MT, their writings are also peppered with potentially taboo-breaking tropes : literary MT is presented as daring (Matusov 2019: 10), an apparent affront to logic or "a near-contradiction in terms" (Voigt and Jurafsky 2012: 18), and a "challenge" to received wisdom (Toral and Way 2015a, 2015b, 2018). There is a strong sense that the status quo is about to be disturbed, and if literary texts are indeed "the last bastion of human translation" (Toral and Way 2015b: 123), then human translators already occupy a defensive position.

THE TECHNICAL CONTEXT: NEURAL MT AND LARGE LANGUAGE MODELS

From a technical point of view, Toral and Way's (2015b) experiments were important because they served as proof of concept, showing that an SMT engine trained partly on literary data could produce better output – as measured by automatic metrics and human rankers – than a generic SMT engine. MT still fell short of human translation quality, but a path to improvement emerged with NMT While a detailed technical discussion of how NMT works is beyond the scope of this chapter (instead see Pérez-Ortiz et al. 2022), it is worth mentioning features of NMT that have particular social and ethical consequences, and that are shared with other technologies of significance to contemporary society.

First, NMT, like SMT, learns a translation model from a training corpus of existing source texts and their translations. NMT is thus heavily reliant on existing training data, and such

training data are not available in equal quantities for all languages. Languages like English are generally very well served, but there is a long tail of low-resource languages for which getting enough data to train an MT system remains a challenge. Even if minority languages could benefit most from literary MT, as suggested by Toral and Way (2015a) – on the basis that translations for such languages "have heretofore been rather limited" (ibid: 243) – they remain the very ones for which training data is sparse. Data-centric approaches to address this issue include the use of 'synthetic' (that is, machine translated) data and data that have been translated using a pivot language (see Ramírez-Sánchez 2022: 175).

Second, the models that NMT learns (and here it departs from SMT) are based on purely numerical representations of the running words (or subwords) in the training corpus. These representations take the form of sequences of numbers ('vectors') known as 'word embeddings'. The artificial neural networks that the technology is named after consist of layers of such word embeddings, where each layer distils information from the preceding layer using techniques developed in the field of deep learning. When a trained NMT model is called upon to translate a new input sentence, it first 'encodes' that sentence in a series of word embeddings. The encoded sentence is then passed on to a 'decoder', which outputs the target language sentence. The decoder proceeds by predicting the next word in the target-language sentence, given the final layer of the multi-layered representation created for the input source sentence *and* the words that it has already output in the target language.

In all the above aspects, NMT is closely related to large language models (LLMs) such as those behind the highly touted GPT family (see, e.g., Zhao et al. 2023): LLMs are also trained on existing corpora, they also use word embeddings to represent words as they occur in natural language texts, and they share the basic transformer architecture used by state-of-the-art NMT (Vaswani et al. 2017). They differ in that NMT models are designed for a specific task (automatic translation in a given language pair) while LLMs are general-purpose models that can be applied to a wide range of tasks, including text summarization, classification, copywriting, script writing, computer coding, etc.

Despite the fact that LLMs are trained overwhelmingly on text in a single language (principally English), their training data as crawled from the web typically also contain text in other languages (Brown 2020), much of which is likely to have come from parallel data. This gives such LLMs multilingual capacities, and it did not take long for them to be applied to

automatic translation in general and, given its cachet, literary translation in particular (see, for example, Karpinska and Iyyer 2023).

Most non-technical users access LLMs through conversational agents like ChatGPT or Microsoft's Copilot. These agents allow users to input a 'prompt', which may contain a natural language instruction like "Translate into French" followed by the text the user would like translated. The system responds to the prompt with a 'completion', which in this case should be a translation into French. Alternatively, the prompt can contain one or several examples of the task, in this case a source sentence or paragraph in English followed by its translation into French. The provision of such examples enables a kind of 'finetuning' of the LLM. If these examples are then followed by a source sentence or paragraph with no translation, the system will provide a completion in the form of a translation. Users can also give explicit instructions or implicit guidance (in the prompt) on the style they would like the completion to be written in, a particularly interesting feature in the case of literary translation.

LITERARY MT: A TYPOLOGY

Using NMT and LLMs to translate literary texts

A key interest in the field of literary MT is how existing resources can be used directly for the purposes of literary translation. Some research thus investigates how 'generic' MT, that is, "MT based on engines trained to cover a wide range of topics, styles and genres, and not specialized in any particular domain" (Ramírez-Sánchez 2022: 165), can be applied to literary text. Authors who take this approach (e.g., Webster et al. 2020; Şahin and Gürses 2021; Thai et al. 2022; Kollektive Intelligenz 2023) frequently use free, online commercial NMT systems such as DeepL, Bing Translator and Google Translate.

Other research focuses on how generic MT engines might be adapted for literary translation, following seminal work by Besacier (2014) and Toral and Way (2015b). The most common approach is to customize (or 'finetune') generic NMT engines upstream (that is, at the training stage), by adding large quantities of literary data, usually novels and their translations. Such customization constitutes a kind of 'genre adaptation' (see Figure 1). The experiments by Antonio Toral and his colleagues are perhaps the most extensive studies in upstream genre adaptation to date (e.g., Toral and Way 2018; Toral et al. 2018; Guerberof-Arenas and Toral 2020; Toral et al. 2020; Guerberof-Arenas and Toral 2022; Toral et al. 2024).

Author accepted manuscript. Please cite as: Kenny, Dorothy (2025) 'Literary machine translation: from taboo to controversy' in Stefan Baumgarten and Michael Tieber (eds) *The Routledge Handbook of Translation Technology and Society*. Routledge, 381-395. 10.4324/9781003271314-33

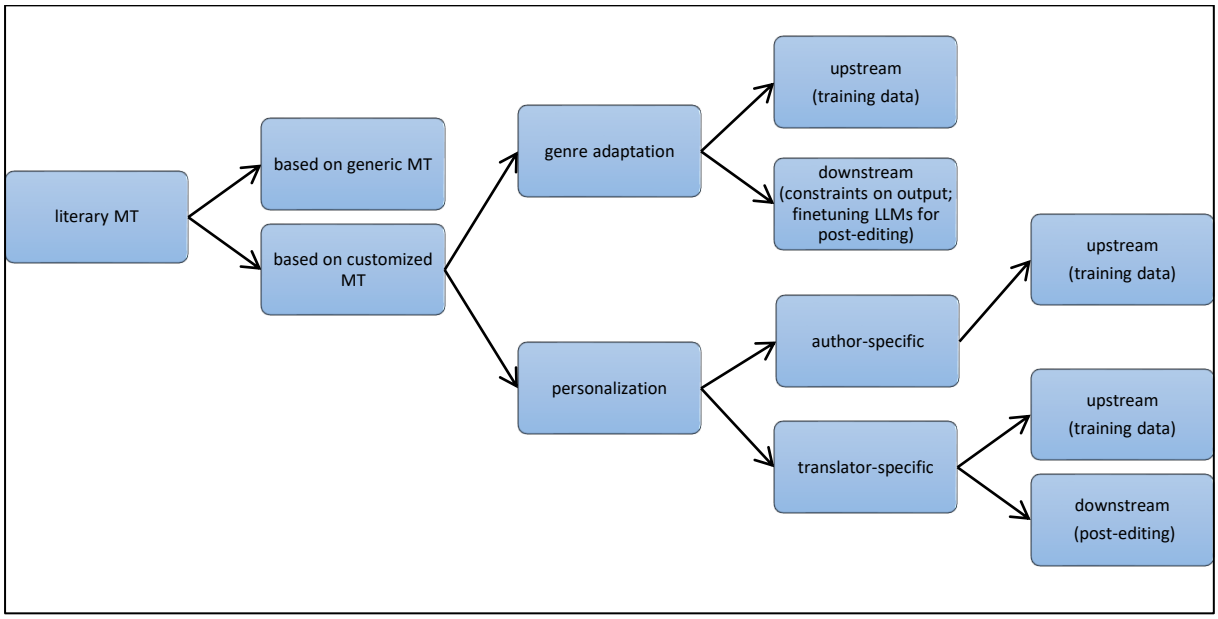


Figure 1: Types of engines used in literary MT

Van de Cruys (2024), whose interest is in automatic translation of poetry, takes a downstream approach to genre adaptation. He trains NMT engines using generic non-literary prose (pairs of sentences in the source and target language) and then imposes constraints at the output (or 'decoding') stage, enforcing a given rhyme scheme. A second kind of downstream genre adaptation is at play in Thai et al.'s (2022) use of GPT-3 to automatically post-edit NMT output. They effectively finetune the LLM so that it can be applied to the intralingual post-editing of literary texts translated by a generic NMT engine.

A second broad type of customization involves personalization. In 'author-specific personalization' training data from texts by a single author are added to the generic training data. Oliver (2024), who refers to "author-tailored" MT, exemplifies this approach. Given that the texts in question are usually taken from a particular genre, namely novels, such personalization could be characterised as genre-specific, but the focus here is on trying to capture something of a single author's style. Translator-specific personalization, meanwhile, involves adding translations completed by a single translator to the training data, this time in the interest of capturing that particular translator's style (e.g., Yirmibeşoğlu et al. 2023). It is possible to combine upstream author-specific and translator-specific personalization by finetuning existing engines using source texts by a single author and their translations by a single translator, as shown by Kuzman et al. (2019) and Hansen and Esperança-Rodier (2022). A second type of translator-specific personalization can be attempted downstream: in this case rather than finetuning an engine using training data from the translator's oeuvre,

researchers work with a translator who post-edits the output of a generic MT engine, so that the target text simulates their normal output (see Kenny and Winters 2020, 2024; Winters and Kenny 2024). In principle, such post-editing could also be attempted by an LLM, following Thai et al.'s (2022) lead, if the language model has been finetuned using data from a single translator.

Studies that focus on upstream customization frequently compare the output of literary-adapted systems with that of a baseline 'general domain' research system and/or with commercial systems. On some occasions (Oliver 2024; Toral et al. 2024), finetuned research systems are found to score better than baseline research systems on metrics used to measure performance, but both types of research system may be outperformed by the commercial systems on at least some of the metrics used. The superior performance of the commercial systems in these cases is likely due to the greater quantities of data on which they are trained (measured in word or subword tokens) and the size of the models they build (measured in the number of 'parameters' or 'weights' the neural networks learn during training).

The performance gap between commercial systems and some (even personalized) research systems provides motivation for the use of the former in literary MT research. An even stronger motivation is the fact that they are free to use and easy to access. Experiments in literary MT by practising literary translators (as opposed to computer scientists) thus unsurprisingly often use only free online commercial systems like Google Translate, Bing Translator and DeepL (see, for example, Constantine 2019; King 2019; Oeser 2020; Kollektive Intelligenz 2023). There are, however, limits on the amount of text that can be translated using free online MT, and practising translators may choose to pay for more extensive MT services from these commercial providers. Not much is known about how literary publishing houses use paid-for commercial MT, although it is likely that this is happening to some extent, given the findings of surveys conducted by professional literary translators' associations (see below).

Relevant features of source texts

Researchers in literary MT are also interested in features of source texts that may influence the performance of the technology. Voigt and Jurafsky (2012), for example, stress the rich referential cohesion of English-language literary texts, a feature that was captured less well in translation into Chinese by Google Translate (when it was still using SMT) than by human translators. Taivalkoski-Shilov's more speculative work (2019b) considers why free indirect

discourse is likely to be problematic for MT. Poncharal's (2021) empirical study, in turn, shows how free indirect discourse, English simple past tense forms and expressions of subjective perception present what appear to be insurmountable obstacles for state-of-the-art MT into French. Toral et al. (2024), meanwhile, consider the relationship between the perceived literariness of texts and the performance of literary-adapted NMT engines for English into Dutch, where literariness is treated in the first instance as a deliberately vague attribute imputed to texts by readers. Guerberof-Arenas and Toral (2020, 2022), for their part, home in on 'units of creative potential' (UCPs) in the source texts used in their experiments, where such UCPs include features like figurative language and phonaesthetic elements. These features turn out to be handled more creatively in human translation (HT) than in other modes.

Features of machine-translated and post-edited literary texts

The above studies are complemented by research into the features of machine translated and post-edited texts. Generic MT output has been found to be less lexically and structurally diverse than comparable HT in a number of studies (e.g., Toral 2019; Vanmassenhove et al. 2019). In what might be the most extensive work on the matter to date, involving three language pairs and thousands of different morphosyntactic constructions, Luo et al. (2024: 1) find that "MT is consistently more conservative than HT, with less morphosyntactic diversity, more convergent patterns, and more one-to-one alignments." Among the studies that have made similar findings specifically for literary MT are Webster et al. (2020), Lee (2021) and Castilho and Resende (2022). Webster et al. (2020) further note that NMT-generated texts are less cohesive than their corresponding human translations. Reduced diversity in machine translated texts is considered a problem for literary translation in particular, as it may suggest homogenization of style. More generally, it is associated with the amplification of unwanted bias in machine-translated text (Vanmassenhove et al. 2019; and see below).

Toral (2019) also looks at post-edited MT (PEMT), finding that PEMT falls somewhere between HT and MT on measures of lexical and structural diversity. Kenny and Winters (2020) likewise focus on post-edited texts in a study of literary translator Hans Christian Oeser's style, finding that Oeser's style is still detectable although somewhat diminished in his post-edited work (see also Winters and Kenny 2024).

The evaluation of literary MT

Literary MT research, like most MT research, is very concerned with evaluation. The purview of translation evaluation was already large (see, e.g., Moorkens et al. 2018), but engagement with literary MT has spurred researchers into embracing metrics and methods not formally considered germane in MT evaluation. Researchers in literary MT thus measure: the aforementioned textual features of lexical diversity, cohesion and syntactic divergence from source texts (e.g., Tezcan et al. 2019; Toral 2019; Webster et al. 2020); the creativity of texts produced in MT, PEMT and HT modes (Guerberof-Arenas and Toral 2022); and reader responses to machine translated literature (Guerberof-Arenas and Toral 2020, 2022, forthcoming), to name just a few. New or adapted error classification schemes have been proposed specifically for literary MT (e.g., Matusov 2019; Tezcan et al. 2019) and while automatic string-based evaluation metrics like BLEU (Koehn 2022: 53-55) persist, increasing doubt is being cast over their relevance for literary MT (e.g., Matusov 2019: 13).

Literary translators' reactions to and interactions with MT

Recent work on literary MT has involved eliciting literary translators' judgements in evaluating MT output (e.g., Moorkens et al. 2018; Guerberof-Arenas and Toral 2022). Literary translators may also be asked to post-edit MT output, in which case the number of post-edits made can be used as a proxy measure for MT quality (Moorkens et al. 2018; Guerberof-Arenas and Toral 2020). Some studies compare translators' productivity—measured in time spent on the task—in 'from scratch' HT and PEMT (Moorkens et al. 2018; Toral et al. 2018; Guerberof-Arenas and Toral 2020), reflecting one of the overarching concerns of general MT research. In keeping with previous translation process research, these sources tend to also capture technical and cognitive effort, measured through keystrokes and pauses respectively. Kolb (2024), utilising both keylogging and think aloud protocols, provides a particularly rich analysis of the processes involved as five different literary translators post-edit a short story translated into German by DeepL.

An interesting subcategory of research on literary translators' interactions with MT comes in the form of self-reports. Oeser (2020), for example, describes his use of DeepL to create a first draft translation of an entire novel in German; members of the German literary collective *Kollektive Intelligenz* (2023) outline their use of various functions offered by DeepL in translating an excerpt from a romantic novel; and Rothwell (2024) relates how he integrates MT (also from DeepL) into his translation environment when re-translating Proust's *La*

Prisonnière into English. Concrete findings from these sources complement the higher-level insights available from a recent survey of French literary translators who were asked about their interactions with publishers regarding MT use (ATLF 2023).

WHERE ARE WE NOW?

Any analysis that attempts to generalize on the basis of the empirical research conducted so far into literary MT is faced with a number of challenges. First, despite growing interest in the field, the absolute number of studies that investigate state-of-the-art literary MT is still small. Second, it is difficult to compare results from the studies that do exist, given the vast number of variables involved. These include the language pairs investigated: the target languages in the studies addressed in this chapter, for example, are Catalan, Chinese, Czech, Dutch, English, French, German, Japanese, Korean, Polish, Russian, Slovene and Turkish. English is, however, by far the most commonly encountered source language, perhaps unsurprisingly, given its importance in global publishing. Other variables include those related to the training and testing environment, for example, whether researchers are testing generic or custom MT. In the former case, studies are often based on 'black-box' analyses of commercial MT output, where researchers have limited insight into the training data and techniques used and thus are forced to speculate on the causes of patterns they observe. In the latter case, studies can diverge widely not just in the training data used, but also in the high-level settings (or 'hyper-parameters') selected during system training. Where LLMs are used, the output will also depend on the nature of the prompt given. And the computing resources available to large corporations and small research labs will also differ to a great extent, with possible consequences for the quality of the systems they train. Further variables include the subgenres investigated (e.g., classics in Webster et al. 2020; fantasy literature in Hansen et al. 2022), the authors covered (see below), and the expertise level of participants in various experiments (professional translators, student translators, language learners, etc.). Third, even when researchers purport to study a common concept, the operationalization of that concept can vary: 'translator style', for instance, is accessed through n-gram distribution and imputed to individual MT systems, on the one hand, and human translations as an undifferentiated mass, on the other, in Lee (2021). In Winters and Kenny (2024), meanwhile, it is accessed through keyword analysis and imputed to an individual human translator. Likewise, 'creativity' is operationalized differently by different authors, with Guerberof-Arenas and Toral (2020) going so far as to introduce a new metric to quantify creativity in machine- and

human-translated texts. Even error typologies differ from study to study. These caveats aside, in the following paragraphs I attempt to draw some tentative conclusions about the actual and potential consequences of literary MT, drawing on the above-cited sources.

THE CONSEQUENCES OF LITERARY MACHINE TRANSLATION

Translated texts

The theoretical risk of MT-induced harm to translated texts and by extension the source-text authors' reputations has been addressed by Taivalkoski-Shilov (2019a). While empirical research has barely touched upon factors related to source-text authors, there is, by now, considerable evidence that machine-translated literary texts are in many ways inferior to human-translated literary texts. Human translations are often considered 'better' than machine translations in studies that involve ranking (e.g., Toral 2019; Thai et al. 2022) and nearly all empirical studies report on high numbers of errors in MT and LLM output. Such errors range from fairly innocuous grammatical agreement errors to meaning-deforming lexical and syntactic errors, omissions, unwarranted additions and repetitions, register violations, incorrect pronoun assignment, and inconsistent translations, among others (see, e.g., Matusov 2019; Webster et al. 2020; Karpinska and Iyyer 2023). The presence of such errors means that post-editing remains a vital part of literary MT workflows, but there is no guarantee that even professional literary translators will find and eliminate all errors in the MT output. As Şahin and Gürses (2021: 193) put it in their study of English-to-Turkish literary MT, "we observed that a number of serious errors were either unnoticed and therefore passed on from the MT output, or else misinterpreted in the human intervention". Guerberof-Arenas and Toral (2022: 197) make a similar observation, and Herland, one of the contributors to the Kollektive Intelligenz (2023) experiments, expresses lingering anxiety about the possible persistence of errors even in her carefully post-edited work.

Other issues, while they do not constitute errors as such, may also have a deleterious effect on machine-translated text. Taivalkoski-Shilov (2019a: 697), for example, has suggested that if MT engines that are trained on data from a single source are used to translate texts by a variety of source text authors, this might end up homogenizing those authors' styles. The corollary would also hold, in that an MT engine trained on data from a wide range of sources could heterogenize a single source text author's style. These hypotheses have not been tested empirically, to my knowledge. As already mentioned, it has been found, however, that

compared against a benchmark of professional human translations, MT output is less lexically varied and more syntactically homogeneous (e.g., Toral 2019; Webster et al. 2020; Lee 2021).

As indicated above, diversity is also related to unwanted bias, another well-known feature of machine-translated texts that is likely to adversely affect literary MT output. Bias takes various forms: it may be political, racial or gender-related, for example, where the latter is most famously associated with preference for male over female forms in target languages that mark gender. Bias is related not just to patterns inherent in the training data, but also to the machine learning algorithms used in NMT and LLMs, which tend to further amplify these already frequent patterns (see, e.g., Savoldi et al. 2021). Again, bias in literary MT has not been specifically addressed by researchers to my knowledge.

Finally, and as alluded to above, a number of studies by Guerberof-Arenas and Toral (2020, 2022, forthcoming), among others, have indicated that machine-translated texts are less creative than their human-translated counterparts.

The translation process

Creativity is not just a feature of translation products; it is also a phenomenon that is experienced by translators as they complete their work. When creativity is felt to be constrained, which commonly happens when literary translators are asked to post-edit MT output (Guerberof and Toral 2020, 2022), this adversely affects the translation experience. Interestingly, constraints on creativity appear to be felt as acutely by translators working on factual texts as on literary texts (Reissig and Ruppel, both in Kollektive Intelligenz 2023). Reduced creativity, as reported in a number of sources (Kenny and Winters 2020; Şahin and Gürses 2021; ATLAS and ATLF 2023: 4; Meyer in Kollektive Intelligenz 2023; Kolb 2024), appears to be one consequence of the priming effects of MT output, whereby the MT proposal sets the translator/post-editor on a linguistic course they might not have chosen if left to their own devices, and of which they might not even be fully aware (Oeser 2020: 22).

To make matters worse, reduced creativity, unwanted priming effects and residual-error induced malaise do not seem to be offset by productivity gains for many literary translators working in post-editing mode. Despite the findings of pioneers in literary MT (Moorkens et

al. 2018; Toral et al. 2018), whose studies were conducted under experimental conditions, commentators working in environments that better approximate normal professional conditions report that while there is significant variation in post-editing speeds among research participants (Kolb 2024), post-editing literary MT is generally perceived as saving little or no time (ATLF 2023; Kollektive Intelligenz 2023). Finally, even if the post-editing mode is faster on some occasions, it is not self-evident that translators want to pursue higher speed at any cost (Daems 2022: 60), given the deleterious effects of acceleration on translators' health and quality of life (ATLAS and ATLF 2023: 7).

Self-reports from literary translators point to other ways in which they have to adjust normal workflows to accommodate the default use of MT output. Herland (in Kollektive Intelligenz 2023), for example, does not normally read the full text through before she starts translating or editing, so that she can maintain the same state of knowledge at any particular point in the text as putative readers of the translation. She finds, however, that she has to abandon this approach in the face of MT output that is so riddled with errors that she cannot concentrate on the text's content. Another common complaint is that working with MT output prevents translators from getting into 'flow' states (e.g., Ruppel in Kollektive Intelligenz 2023). Others (e.g., Oeser 2020) point to the increased cognitive load that comes with having to deal with two 'source' texts (the original language version and the machine-translated version) as well as the incipient target text. Some translators may find inspiration in isolated MT suggestions (Rothwell 2024), but the overriding impression elsewhere is of degradation of the translation process when literary translators work in post-editing mode (ATLAS and AFTL 2023; Kollektive Intelligenz 2023).

The reading experience

Ana Guerberof-Arenas and Antonio Toral spearheaded empirical research into readers' reactions to texts produced through HT, MT and PE. In their broadest study to date (Guerberof-Arenas and Toral forthcoming), they find that engagement, enjoyment and reception scores are higher for HT than for MT or PE among Catalan readers, but that PE comes out on top for Dutch readers. What's more, Dutch readers score the original English version of the text used in the experiment higher than any of the translations. The study suggests that readers' experience of a text depends not just on how it was translated, but also on their own reading patterns, reading language, and perhaps even the status of the languages involved (ibid.).

Author accepted manuscript. Please cite as: Kenny, Dorothy (2025) 'Literary machine translation: from taboo to controversy' in Stefan Baumgarten and Michael Tieber (eds) *The Routledge Handbook of Translation Technology and Society*. Routledge, 381-395. 10.4324/9781003271314-33

The translation market

This chapter began by discussing the aspirations of the pioneers of literary MT to support literary translators' earnings and increasing the circulation of "largely undiscovered" works in translation. From the research conducted so far, we can conclude that neither aspiration appears to be realisable in the short term at least. The ATLF's (2023) survey of French literary translators indicates that any post-editing induced savings do not get passed on to translators/post-editors; rather they are offered lower rates for post-editing, which are not offset by time savings. Overall, the survey finds a disturbing lack of transparency on the publishers' side when it comes to MT, with publishers presenting post-editing jobs as 'revision' jobs in 51% of cases, and MT not mentioned at all on any final translated product. This somewhat antagonistic environment has led to calls for an end to the public financing of initiatives that support such publication practices and for greater transparency when it comes to publishers' dealings with authors whose work they will have translated (ATLAS and ATLF 2023). Furthermore, and as may already be clear from the research literature surveyed above, lesser-used languages and minoritized writers are not the ones who are suddenly given attention in MT workflows. Rather, research has mostly focused on older, classic works that are out of copyright and often already available in human translation in the target language in question (even if the passage of time means that retranslations may be welcome). The dominant source language in the studies reviewed is English, and the writers whose work is translated make for a rollcall of canonical and commercially successful authors (Jane Austen, Agatha Christie, Charles Dickens, Arthur Conan Doyle, Ernest Hemingway, etc.), few of whom could be described as "largely undiscovered". The ATLF (2023) survey confirms English as the dominant source language (featuring in 42% of cases) in literary MT workflows in the contemporary French professional context. It can certainly be argued that lessons learned from the proof-of-concept work done using classics might be fruitfully applied in due course to the work of lesser-known authors in lesser-used languages, but there is no consensus that this is a worthy goal, with CEATL (2023) in particular disputing the idea that AI can be a tool for "easier access to minoritized languages." The same source is also dismissive of the idea that "some books are better suited than others to undergoing generative AI processing" as it "fosters a dangerous distinction between high literature and every other text, impeding societal development through reading." Meanwhile, the suggestion that machine-translated literature could be used to promote reading in the *original* language, has been tested to a limited extent in Oliver et al. (2022). In this source, a small number of

Catalan readers (including primary-school pupils) confirmed the usefulness of having access to a machine translation into their mother tongue when reading an English-language short story.

The translator

On top of the challenges presented by the aforementioned factors, literary translators may also have to deal with a particular type of technological anxiety, given that translators in general are so often described as "exposed" to the encroachment of AI (e.g., Eloundou et al. 2023), and given the ease with which certain commentators claim that the applicability of MT to literary translation is "the ultimate litmus test for whether machine translators will ever replace human translators" (Lee 2021: 1-2). Such an outcome would be particularly galling for those who argue that current AI-assisted translation is based on the theft of data from the very translators the technology would replace (European Writers Council 2023). As suggested above, however, there are myriad reasons why the work of literary translators remains crucial, but it is possible that AI solutionism could wear down enthusiasm for human literary translation among prospective entrants to the profession.

CONCLUSION

Literary MT is now well established as a research area, even if it remains difficult to generalize findings from the diverse range of studies in the field. The application of NMT and LLMs to literary texts remains fraught with difficulty, however, and steeped in controversy. Meanwhile, recent interventions from translation collectives and professional organizations suggest that literary translators are becoming more vocal participants in debates about literary MT, and like millions of other workers, they will demand fairer treatment in any technologically-imbued future. Against this background, researchers who continue to engage with literary MT do so in an environment that may become more hostile to their work. Why, others may ask, should researchers contribute to a field that does not deliver on its own promises, but routinely appears to misappropriate others' data, to degrade target text quality, and to threaten the already precarious livelihood of so many literary translators? There are many possible answers to this question; mine will be three-pronged: first, the empirical studies conducted by researchers in literary MT have informed discussions and strengthened arguments presented by literary translators in their recent critiques of literary applications of AI. Evidence from several of the studies referenced earlier in this chapter has found its way,

directly or indirectly, into manifestos issued, for example, by ATLAS and ATLF (2023), CEATL (2023) and the European Writers' Council (2023). Research into literary MT can thus be used both to support and to problematize the practice. Second, it is still possible that AI could be of benefit to human literary translators in their own practice. Thus far, researchers have barely broached the topic of how AI could augment rather than displace the work of literary translators. Future augmentation-focused empirical research has the potential to show how translators can use the technology in empowering ways. To do so, it will need to integrate professional literary translators as co-creators, full participants and beneficiaries, as anticipated by Kenny (2024). Third, the application of machines to literary translation is already a commercial reality (Aslanyan 2024; ATLF 2023). By shining a light on the practice, researchers can contribute to the moulding of public opinion, and ultimately to the evolution of the market for literary translations.

References

Arnold, D., L. Balkan, S. Meijer, R. L. Humphreys and L. Sadler (1994) *Machine Translation: an introductory guide*, Oxford: NCC Blackwell Ltd.

Aslanyan, A. (2024) "AI translation: how to train 'the horses of enlightenment'", *The Guardian*, 15 March 2024. https://www.theguardian.com/books/2024/mar/15/ai-translation-literature

ATLAS and ATLF (2023) "IA et traduction littéraire : les traductrices et traducteurs exigent la transparence" ATLAS, www.atlas-citl.org/wp-content/uploads/2023/03/Tribune-ATLAS-ATLF-3.pdf

ATLF (2023) "Traduction automatique et post-édition", Association de Traducteurs littéraires de France, https://atlf.org/wp-content/uploads/2023/03/ENQUETE-TRADUCTION-AUTOMATIQUE.pdf

Besacier, L. (2014) "Traduction automatisée d'une oeuvre littéraire: Une étude pilote" [Automatic translation of a literary work: A pilot study], paper presented at the 21st Traitement Automatique du Langage Naturel (TALN 2014), Marseille, http://www.aclweb.org/anthology/F14-2001

Brown, T. (2020) "Languages by word count", https://github.com/openai/gpt-3/blob/master/dataset_statistics/languages_by_word_count.csv

Castilho, S. and N. Resende (2022) "Post-Editese in Literary Translations", *Information* 2022, 13(2): 66, https://doi.org/10.3390/info13020066

CEATL (2023) *Statement on Artificial Intelligence* https://www.ceatl.eu/wp-content/uploads/2023/11/CEATL-AI-statement-EN.pdf

Constantine, P. (2019) "Google Translate Gets Voltaire: Literary Translation and the Age of Artificial Intelligence", *Contemporary French and Francophone Studies* 23(4): 471-479, DOI:10.1080/17409292.2019.1694798

Daems, J. (2022) "Dutch literary translators' use and perceived usefulness of technology. The role of awareness and attitude", in J. L. Hadley, K. Taivalkoski-Shilov, C. S. C. Teixeira and A. Toral (eds) *Using Technologies for Creative-Text Translation*, London: Routledge, 40-65.

Eloundou,T., S. Manning, P. Mishkin and R. Daniel (2023) "GPTs are GPTs: An Early Look at the Labor Market Impact Potential of Large Language Models", preprint, https://arxiv.org/pdf/2303.10130.pdf

Elson, D. K., A. Kazantseva, R. Mihalcea and S. Szpakowicz (2012) "Preface", in *Proceedings of the NAACL Workshop on Computational Linguistics for Literature*, Stroudsburg, PA: Association for Computational Linguistics, iii-iv, https://aclanthology.org/W12-25.pdf

European Writers Council (2023) Analysis: *The success of Generative AI in the book sector is based on theft*, 28th September 2023 https://europeanwriterscouncil.eu/gai-is-based-on-theft/ written originally in German for the Netzwerk Autorenrechte (Authors' Rights Network) by Nina George and André Hansen.

Guerberof-Arenas, A. and A. Toral (2020) "The Impact of Post-Editing and Machine Translation on Creativity and Reading Experience", *Translation Spaces* 9(2): 255-282, https://doi.org/10.1075/ts.20035.gue

Guerberof-Arenas, A. and A. Toral (2022) "Creativity in translation: Machine translation as a constraint for literary texts" *Translation Spaces* 11(2): 184-212.

Guerberof-Arenas, A. and A. Toral (forthcoming) "To be or not to be: a translation reception study of a literary text translated into Dutch and Catalan using machine translation", *Target*.

Hansen, D. and E. Esperança-Rodier (2022) "Human-Adapted MT for Literary Texts: Reality or Fantasy?", NeTTT 2022, July 2022, Rhodes, Greece, 178-190, https://hal.science/hal-04038025

Karpinska, M. and M. Iyyer (2023) "Large language models effectively leverage document-level context for literary translation, but critical errors persist", preprint, https://doi.org/10.48550/arXiv.2304.03245

Kenny, D. (2020) "Machine Translation", in M. Baker and G. Saldanha (eds) *The Routledge Encyclopedia of Translation Studies*, 3rd edition, 305-310, London: Routledge.

Kenny, D. and M. Winters (2020) "Machine translation, ethics and the literary translator's voice", *Translation Spaces* 9(1): 123-49, https://doi.org/10.1075/ts.00024.ken

Kenny, D. (2024) "Do it in style! Human and automatic translation revisited", (keynote) *Language of the Third Millennium XIII*, Jagiellonian University, Kraków, 13-15 March 2024.

Kenny, D. and M. Winters (2024) "Customization, personalization and style in literary machine translation", in M. Winters, S. Deane-Cox and U. Böser (eds) *Translation, Interpreting and Technological Change - Innovations in Research, Practice and Training,* 59-79, London: Bloomsbury.

King, K. M. (2019) "Can Google Translate be taught to translate literature? A case for humanists to collaborate in the future of machine translation", *Translation Review* 105(1): 76-92, https://doi.org/10.1080/07374836.2019.1673268

Koehn, P. (2020) *Neural Machine Translation*, Cambridge: Cambridge University Press.

Kolb, W. (2024) "I Am a Bit Surprised': Literary Translation and Post-Editing Processes Compared", in A. Rothwell, A. Way and R. Youdale (eds) *Computer-Assisted Literary Translation*, 53-68, London: Routledge.

Kollektive Intelligenz (2023) *Kann KI Literatur?* https://kollektive-intelligenz.de/originals/kollektive-intelligenz-kann-ki-literatur/

Kuzman, T., Š. Vintar and M. Arčan (2019) "Neural Machine Translation of Literary Texts from English to Slovene", in *Proceedings of the Qualities of Literary Machine Translation*, 1-9, Dublin, Ireland, European Association for Machine Translation, https://aclanthology.org/W19-7301.pdf

Lee, T. K. (2011) "The death of the translator in machine translation. A bilingual poetry project", *Target* 23(1): 92-112, https://doi.org/10.1075/target.23.1.06lee

Lee, C. (2021) 'How do machine translators measure up to human literary translators in stylometric tests?', *Digital Scholarship in the Humanities* 37(3): 813–829. https://doi.org/10.1093/llc/fqab091

Luo, J., C. Cherry and G. Foster (2024) "To Diverge or Not to Diverge: A Morphosyntactic Perspective on Machine Translation vs Human Translation", https://arxiv.org/pdf/2401.01419.pdf

Matusov. E. (2019) "The Challenges of Using Neural Machine Translation for Literature", in J. Hadley, M. Popović, H. Afli and A. Way (eds) in *Proceedings of the Qualities of Literary Machine Translation*, 10-19, Dublin, Ireland, European Association for Machine Translation, https://www.aclweb.org/anthology/W19-7302.pdf

Moorkens, J., A. Toral, S. Castilho and A. Way (2018) "Translators' perceptions of literary post-editing using statistical and neural machine translation" *Translation Spaces* 7 (2): 240-262. https://doi.org/10.1075/ts.18014.moo

Oeser, H.-C. (2020) "Duel with DeepL. Literary Translator Hans-Christian Oeser on machine translation and the translator's voice." *Counterpoint* 4: 19-23, https://www.ceatl.eu/wp-content/uploads/2020/12/Counterpoint_2020_04.pdf

Oliver, A. (2024) "Author-Tailored Neural Machine Translation Systems for Literary Works", in A. Rothwell, A. Way and R. Youdale (eds) *Computer-Assisted Literary Translation*, 126-141, London: Routledge.

Oliver, A., A. Toral and A. Guerberof-Arenas (2022) "Bilingual e- books via neural machine translation and their reception", in J. Hadley, K. Taivalkoski-Shilov, C. Teixeira and A. Toral (eds) *Using Technologies for Creative-Text Translation*, 89-115, London: Routledge.

Pérez-Ortiz, J. A., M. L. Forcada and F. Sánchez-Martínez (2022) "How neural machine translation works", in D. Kenny (ed.) *Machine translation for everyone*: *Empowering users in the age of artificial intelligence*, 141-164, Berlin: Language Science Press, https://doi.org/10.5281/zenodo.6760020

Poncharal, B. (2021) "La TA à l'épreuve du texte littéraire : d'une (im)possible restitution de l'expérience de lecture ? ", *La main de Thôt* [En ligne] 9, mis en ligne le 11 décembre 2021, http://interfas.univ-tlse2.fr/lamaindethot/983

Ramírez-Sánchez. G. (2022) "Custom machine translation", in D. Kenny (ed.) *Machine translation for everyone: Empowering users in the age of artificial intelligence*, 165–186. Berlin: Language Science Press. DOI:10.5281/zenodo.6760022

Rothwell, A. (2024) "Retranslating Proust Using CAT, MT and Other Tools", in A. Rothwell, A. Way and R. Youdale (eds) *Computer-Assisted Literary Translation*, 106-125, London: Routledge.

Şahin, M. and S. Gürses (2021) "English–Turkish Literary Translation Through Human–Machine Interaction", *Revista Tradumàtica*, 19: 171-203. https://doi.org/10.5565/rev/tradumatica.284

Savoldi, B., M. Gaido, L. Bentivogli, M. Negri and M. Turchi. (2021) "Gender Bias in Machine Translation", *Transactions of the Association for Computational Linguistics* 9: 845-874, https://doi.org/10.1162/tacl_a_00401

Taivalkoski-Shilov, K. (2019a) "Ethical Issues Regarding Machine(-assisted) Translation of Literary Texts", *Perspectives* 27(5): 689-703. https://doi.org/10.1080/0907676X.2018.1520907

Taivalkoski-Shilov, K. (2019b) "Free indirect discourse: an insurmountable challenge for literary MT systems?", in *Proceedings of the Qualities of Literary Machine Translation*, 35-39, Dublin, Ireland, European Association for Machine Translation, https://aclanthology.org/W19-7305

Tezcan, A., J. Daems and L. Macken (2019) "When a 'sport' is a person and other issues for NMT of novels", in *Proceedings of the Qualities of Literary Machine Translation*, 40-49, Dublin, Ireland, European Association for Machine Translation, https://aclanthology.org/W19-7306.pdf

Thai, K., M. Karpinska, K. Krishna, W. Ray, M. Inghilleri, J. Wieting and M. Iyyer (2022) "Exploring Document-Level Literary Machine Translation with Parallel Paragraphs from World Literature", *Proceedings of the 2022 Conference on Empirical Methods in Natural Language Processing*, 9882-9902, Association for Computational Linguistics, https://aclanthology.org/2022.emnlp-main.672.pdf

Toral, A. (2019) "Post-editese: an exacerbated translationese", in *Proceedings of Machine Translation Summit XVII: Research Track*, 273-81, European Association for Machine Translation, https://aclanthology.org/W19-6627

Toral, A. and A. Way (2015a) "Machine-assisted translation of literary text", *Translation Spaces* 4(2): 240-267.

Toral, A. and A. Way (2015b) "Translating Literary Text between Related Languages using SMT", in *Proceedings of NAACL-HLT Fourth Workshop on Computational Linguistics for Literature*, 123-132, https://aclanthology.org/W15-0714.pdf

Toral, A. and A. Way (2018) "What level of quality can Neural Machine Translation attain on literary text?", in J. Moorkens, S. Castilho, F. Gaspari and S. Doherty (eds) *Translation Quality Assessment: From Principles to Practice*, Heidelberg: Springer, International Publishing, 263-287, https://link.springer.com/chapter/10.1007/978-3-319-91241-7_12

Toral, A., M. Wieling and A. Way (2018) "Post-editing effort of a novel with statistical and neural machine translation", *Frontiers in Digital Humanities* 5(Art. 9): 1-11. https://doi.org/10.3389/fdigh.2018.00009

Toral, A., A. Oliver and P. Ribas Ballestín (2020) "Machine Translation of Novels in the Age of Transformer", in J. Porsiel (ed.) *Maschinelle Übersetzung für Übersetzungsprofis*, Berlin: BDÜ Fachverlag, 276-295.

Toral, A., A. van Cranenburgh and T. Nutters (2024) "Literary-Adapted Machine Translation in a Well-Resourced Language Pair. Explorations with More Data and Wider Contexts", in A. Rothwell, A. Way and R. Youdale (eds) *Computer-Assisted Literary Translation*, 27-52, London: Routledge.

Van de Cruys, T. (2024) "Up and About, or Betwixt and Between? The Poetry of a Translation Machine", in A. Rothwell, A. Way and R. Youdale (eds) *Computer-Assisted Literary Translation*, 158-172, London: Routledge.

Vanmassenhove, E., D. Shterionov and A. Way (2019) "Lost in translation: Loss and decay of linguistic richness in machine translation", in *Proceedings of MT Summit XVII, Volume 1: Research Track*, Dublin, Ireland, 222-32, European Association for Machine Translation, https://aclanthology.org/W19-6622.pdf

Vaswani, A., N. Shazeer, N. Parmar, J. Uszkoreit, L. Jones, A. N. Gomez, Ł. Kaiser, I. Polosukhin (2017) "Attention Is All You Need", 31st Conference on Neural Information Processing Systems (NIPS 2017), Long Beach, CA, USA. https://doi.org/10.48550/arXiv.1706.03762

Voigt, R. and D. Jurafsky (2012) "Towards a literary machine translation: The role of referential cohesion", in *Proceedings of the NAACL-HLT 2012 Workshop on Computational Linguistics for Literature*, 18-25, Association for Computational Linguistics, https://aclanthology.org/W12-2503/

Webster, R., M. Fonteyne, A. Tezcan, L. Macken and J. Daems (2020) "Gutenberg Goes Neural: Comparing Features of Dutch Human Translations with Raw Neural Machine Translation Outputs in a Corpus of English Literary Classics", *Informatics*, 7(3), 32, https://doi.org/10.3390/informatics7030032

Winters, M. and D. Kenny (2024) "Mark my keywords: a translator-oriented exploration of style in literary machine translation", in A. Rothwell, A. Way and R. Youdale (eds) *Computer-Assisted Literary Translation*, 69-87, London: Routledge.

Yirmibeşoğlu, Z., Dursun, O., Dallı, H., Şahin, M., Hodzik, E., Gürses, S. and Güngör, T. (2023). "Incorporating human translator style into English-Turkish literary machine translation", arXiv preprint arXiv:2307.11457.

Zhao, W. X., K. Zhou, J. Li, T. Tang, X. Wang, Y. Hou, Y. Min, B. Zhang, J. Zhang, Z. Dong, Y. Du, C. Yang, Y. Chen, Z. Chen, J. Jiang, R. Ren, Y. Li, X. Tang, Z. Liu, P. Liu, J.-Y. Nie and J.-R. Wen (2023) *A Survey of Large Language Models*, https://doi.org/10.48550/arXiv.2303.18223

Further Reading

Karpinska, M. and M. Iyyer (2023) "Large language models effectively leverage document-level context for literary translation, but critical errors persist", preprint, https://doi.org/10.48550/arXiv.2304.03245 (one of the first sources to investigate the potential of large language models in literary MT)

Taivalkoski-Shilov, K. (2019a) "Ethical Issues Regarding Machine(-assisted) Translation of Literary Texts", *Perspectives* 27(5): 689-703. https://doi.org/10.1080/0907676X.2018.1520907 (an important reflection on how MT might cause harm in the context of literary translation)

Related Topics

Machine Translation, Creativity, Society
Post-Editing, Translation Quality and the Translation Profession
Translation Studies and Digital Humanities