2

Translating the unknown: a case study on the usefulness of machine translation in comparative literature research

Sergio Portelli

Introduction

Despite the uneasiness with which the discipline of comparative literature (CL) long regarded translation, at least until the 'cultural turn' in translation studies (TS) in the last decades of the twentieth century (Bassnett 1993; Lefevere 1995; Apter 2006; D'hulst 2007; Tee 2012; Ning and Domínguez 2016), new areas of research in the latter field of study have provided CL with interesting tools that are being increasingly tested, evaluated and used in contemporary scholarship. Apart from translation theory, which, since Walter Benjamin, has increased its relevance within CL (see Apter 2006), the sub-discipline of computer-aided literary translation (CALT), together with the emergence of the digital humanities, has introduced technology as a potentially useful tool in comparative literary research.

CL researchers who deal with non-canonical works may find themselves in a position where their studies are limited by language barriers. Some minor texts, which may be of interest within a wider research topic, may never have been translated, thus rendering it impossible for the researcher to access them without human or technological assistance. Since a fully fledged, publishable translation by a competent literary translator may be expensive and possibly unnecessary for the scope of a research project, machine translation (MT) can prove to be a useful tool. This chapter investigates the extent to which MT may

come to a CL researcher's aid in accessing untranslated literary works for specific types of analysis. It limits itself to the genre of historical drama and derives the data for its conclusions from a case study conducted on the German nineteenth-century tragedy *Marino Faliero* by Heinrich Kruse (1876). The text was machine translated, post-edited by the researcher and reviewed by a professional German linguist to obtain an adequate translation in English for CL research purposes.

Comparative literature and translation

D'hulst (2007, 96) states that 'for comparatists, translations are an important source of information to understand the *rapports de faits* between literatures'. However, earlier generations of CL scholars were reluctant to use translations and regarded them as unreliable sources due to the intervention of the translator in the source text (Ning and Domínguez 2016). The only alternative was that comparatists had to be polyglots with a profound knowledge not just of various languages, but also of their respective cultures. In Bassnett's (1993, 139) words, 'a comparatist [...] would read original texts in the original languages, an infinitely superior form of reading than any which involved translation'. Such competence is not very common, especially when it involves languages that are not widely known in the West; this constituted a serious obstacle for prospective researchers interested in the field (Bernheimer 1995). An ever-increasing attention to the literatures of non-Western countries rendered direct access to texts more problematic (Apter 2006). Consequently, the issue of translation could not be ignored much longer (Tee 2012).

The emergence of TS as an academic discipline in the 1970s was a significant step towards the legitimisation of translation in comparative literary studies. The contribution of scholars such as Itamar Even-Zohar on the need to study translations as an important aspect of national literature cannot be underestimated. He expressed his conviction in no uncertain terms when he wrote, 'I conceive of translated literature not only as an integral system within any literary polysystem, but as a most active system within it' (Even-Zohar 1990, 46). The importance of translated literature as literature proper, as well as the need to overcome the focus on the issue of equivalence between source and target texts, was taken further by the scholars of the Manipulation School, who advocated 'a descriptive method [that] takes the translated text as it is and tries to determine the various factors that may account for its particular nature' (Hermans 1985, 13). They focused on the reasons why texts are translated into a

specific language, the conditions under which a translation is produced and the intention(s) behind the process and the product itself. In the early 1980s, Lefevere introduced the concepts of 'refraction' and 'rewriting' in the manipulation of literature. He defined refraction as 'the adaptation of a work of literature to a different audience, with the intention of influencing the way in which that audience reads the work' (Lefevere 1982, 4), whereas in his later elaboration of the concept of rewriting he emphasised the role of power dynamics, where patronage, ideology and poetics all play a part in the manipulation of literature, also by means of translation (Lefevere 1992). Together with Bassnett, he overturned the generally accepted hierarchy between CL and TS by stating that:

> translation has been a major shaping force in the development of world culture, and no study of comparative literature can take place without regard to translation. We have both suggested on occasions, with a deliberate intention of subverting the status quo and drawing attention to the importance of Translation Studies, that perhaps we should rethink our notions of Comparative Literature and redefine it as a sub-category of Translation Studies instead of vice versa. (Bassnett and Lefevere 1990, 12)

Bassnett (1993, 161) reaffirmed this position by declaring that CL was on its way out and that it should be considered as a 'subsidiary subject area' with respect to TS. The affirmation of the cultural turn in TS in the 1990s brought the two disciplines closer together. At the beginning of the new century, translation became an important feature in CL at a time when world literature was gaining prominence (Ning and Domínguez 2016). The potential for collaboration between the two disciplines with reference to world literature was highlighted by Bermann (2009, 432, 443), who, in her 2009 American Comparative Literature Association (ACLA) presidential address, envisaged the creation of an '*and* zone', a space of interaction between TS and CL that 'becomes the site for articulating an interpretation, or a theoretical meditation on the qualities or limitations of the text, or a critical reflection on literature more generally'.

The success of academic programmes in TS worldwide in recent decades has put the discipline on a firm standing and affirmed its interdisciplinarity. In light of this development, CL can only gain from the potential insights provided by TS. D'hulst (2007, 103) recognised this when he rhetorically asked, 'why should comparative literature be reluctant when it comes to recognizing the conceptual apparatus and methodology that has been developed for the study of translation?'

Computer-aided literary translation

Apart from literary translation and translation theory, which are the two sub-fields of TS that are most readily considered to be close to CL, another area that has recently been gaining more attention is CALT. It may be classified within the emerging research area of the digital humanities, where digital technology is applied to the various fields in the humanities. The application of computers to the study of literary translation began in the 1960s, when machines became capable of processing elements of texts in electronic format (Zanettin 2017). As digital technology improved, the possibility of obtaining data from text processing opened new avenues for research in various areas of literary studies. Zanettin (2017) has identified two main areas to which computer-assisted research has more readily been applied, namely the creation of searchable electronic archives of texts in digital format and textual analysis within the wider context of stylistics. Searches at a lexical and syntactic level may shed light on the linguistic and stylistic characteristics of one or more translations of a particular author and pave the way for broader comparative analyses. The creation of corpora of literary translations allows for a vast range of studies based on the analysis of statistical data, such as the visibility of translators, the latter's lexical and stylistic choices and differences between translations of a particular text across time.

Another important aspect of CALT concerns the types of translation technology applied to literary texts, especially MT. In recent years, the quality of MT has improved greatly (Toral and Way 2015). The transition from rule-based and statistical methods to neural networks has resulted in the production of higher-quality raw output that requires less human post-editing (see Castilho et al. 2017). Whereas MT is generally used for non-literary texts, in recent years studies have been carried out to evaluate the extent to which such tools may be used in the actual production of literary translations. Genzel et al. (2010, 163) used a custom statistical phrase-based MT tool to produce translations of poetry with meter and rhyme, reaching the conclusion that 'it seems that at the present state of MT, one does indeed have to choose between getting either the form or the meaning right'.

Greene et al. (2010) used statistical machine tools to analyse and generate love poetry in English and then to translate rhythmic poetry from Italian to English. What is of interest here is their attempt to translate Dante's rhymed hendecasyllables of the *Divine Comedy* into English rhymed iambic pentameters. The quality of the translations, bound as they were by prosodic and rhyme constraints, apart from the need to

render the meaning and effect of the source text, was unsatisfactory. This led the researchers to suggest that rather than seeking such an objective, a text generation and translation program may be used 'for inspiration' rather than for the production of poetic texts (Greene et al. 2010, 532).

Jones and Irvine (2013) analysed samples of machine-translated French literature into English in order to gauge the extent to which a machine can convey the 'experience' of reading a text, since literature elicits a special type of involvement by the reader that other text types do not. More specifically, they analysed the raw output in terms of textual fidelity and the ethical aspect of the translator's (in)visibility, as discussed by Venuti (2008). They used both a custom phrase-based tool and Google Translate for comparison in order to see how MT deals with domestication and foreignisation. The results of their study showed that statistical MT cannot adopt translation strategies considering the issue of ethnocentricity. The issue of the translator's 'voice' in CALT has more recently been discussed by Taivalkoski-Shilov (2019) and Kenny and Winters (2020), from an ethical perspective.

In a pilot study, Besacier (2014) used MT to translate a previously untranslated text by American writer Richard Powers into French. The raw output was then post-edited by non-professional translators and subsequently evaluated by a professional translator, who noted that the final version had certain linguistic and stylistic shortcomings, namely syntactic calques, the presence of anglicisms and the incorrect rendering of culture-specific elements (Besacier 2014, 393). Besacier concluded that such a low-cost process may have some benefits, such as the possibility for authors to have their books quickly translated into many languages. However, he wondered if sacrificing quality for a wider diffusion would be worthwhile for a writer. From the readers' perspective, potential benefits would be faster access to translations of their favourite author's works, and having the translation as an aide for non-native readers, who may struggle to understand certain passages in the original text (Besacier 2014, 394).

Like Besacier's study, that of Toral and Way (2015) also focused on the quality of machine-translated literary texts. They built custom statistical MT systems to translate novels from Spanish into Catalan and from Spanish into English. In order to investigate whether statistical MT may be used successfully to translate literary texts, they compared the degree of freedom of the translation and the narrowness of parallel corpora in the domain of literature to other non-literary domains. They also focused on two specific parameters: the closeness of the languages and the literary genre. Their most interesting claims, based on their results, are that the closeness of

related languages is a primary factor in translation freedom and that MT 'can be useful to assist with the translation of novels between closely related languages' (Toral and Way 2015, 130).

Recently, a stronger focus has been placed on the post-editing of literary texts, in which professional literary translators have been asked to post-edit statistical and/or neural machine-translated texts, thus providing analysable data (Moorkens et al. 2018; Toral and Way 2018; Toral et al. 2018; Kenny and Winters 2020; Castilho and Resende 2022). Toral and Way (2018) trained both a phrase-based statistical and a neural MT engine on large amounts of literary texts (over 100 million words) and then used the engines to translate 12 widely known modern and contemporary novels. The Bilingual Evaluation Understudy (BLEU) automatic evaluation metric showed that neural MT performed better than phrase-based MT, whereas an additional human evaluation of three of the novels showed that in some cases the automatic translations were deemed of similar quality to those done by human translators.

In another study, Toral et al. (2018, 1) claimed that theirs was 'the first experiment in the literature in which a novel is translated automatically and then post-edited by professional literary translators'. As a case study, they translated a chapter of an English-language fantasy novel into Catalan, using both phrase-based statistical and neural MT. The raw output from both tools was then post-edited by six experienced professional translators, who also translated the text from scratch for comparative purposes. In all three cases, the researchers recorded all the keystrokes, the time taken to translate each sentence, the number of pauses and their duration. The results showed that the use of both types of MT led to an increase in productivity, a reduction in the number of keystrokes and a reduction in pauses, though the pauses tended to be longer. In all cases, neural MT performed better than its statistical counterpart. The data collected for this study was used by Moorkens et al. (2018), who collected feedback from the participants before and after the post-editing and translation. The researchers found that the participants unanimously preferred to translate from scratch because the option allowed them greater freedom, but those with less experience appreciated the usefulness of the raw output from MT.

Another study on literary translators' perspectives regarding post-editing was conducted by Kenny and Winters (2020). They sought to investigate to what extent the raw output produced by neural MT influences the *textual voice* of the literary translator who post-edits it. A professional translator from English into German was asked to post-edit an excerpt from a novel he had previously translated and to comment

on his task. The researchers concluded that post-editing affected the translator's textual voice, which was deemed less perceivable than in the version he had previously translated from scratch.

Another very recent study on literary post-editing was carried out by Castilho and Resende (2022), who investigated the phenomenon of post-editese in literary texts. The term post-editese refers to the difference between the characteristics of texts translated by a human translator and those of the respective post-edited versions, in relation to the raw output produced by an MT tool (Castilho and Resende 2022, 4). The researchers carried out a study on two novels, Lewis Carroll's *Alice's Adventures in Wonderland* and Paula Hawkins' *The Girl on the Train*. The texts were translated from English into Brazilian Portuguese using Google Translate and analysed to identify the nature and extent of post-editing features present in them. The post-edited translations were then compared to human translations to highlight the differences. The results showed, among other things, that the Hawkins text required less post-editing than the one by Carroll, leading the researchers to conclude that 'while literary texts whose author's style is full of figurative language pose a harder challenge to the MT system, texts that emphasise action over language style are less challenging' (Castilho and Resende 2022, 19).

Scope of the study

The present study brings together both CL and CALT in that it investigates how CL researchers may resort to MT in order to overcome the not-uncommon problem of not knowing the language of a text they need to access for a specific research project. When studying groups of lesser-known and often untranslated works across languages, researchers may be faced with a potentially unsurmountable language barrier. As a possible solution to the problem, the present researcher sought to devise a methodology that could allow access to untranslated literary works for comparative studies specifically focused on plot and narrative structure. The methodology was tested on a German text, which the researcher could not read because he does not know the language. A freely available neural MT tool was used to obtain a raw output before the text was post-edited by the researcher himself to obtain a version that was comprehensible enough for him to follow. The third and final stage consisted of a revision of the post-edited version by an experienced literary translator who works out of German into English. If successful, the model could be replicated for other works and other languages, thus allowing CL scholars to access literature in a wider range of languages.

The object of the study was a German five-act tragedy written in the nineteenth century, entitled *Marino Faliero*. It was published in 1876 by Heinrich Kruse, a liberal journalist and writer who believed in a united German state led by Prussia (Kruse-Jarres 2008). His 16 historical dramas focus on issues such as justice, morality and uprightness. They were never successful, neither during his lifetime nor after, and no English translations have been traced.[1] *Marino Faliero* was never reprinted, and therefore only one version of the text exists. It consists of 2,893 verses (mostly unrhymed iambic pentameters), stage directions and a small number of footnotes.

Methodology

In the pre-translation phase, the first step was to prepare the source text for MT. The text is copyright-free and freely available on the internet (such as on Google Books). The book is printed in the Gothic Fraktur typeface, which initially made the text difficult to read. Moreover, since the PDF document is a scanned version, the text could not be digitally copied. The commonly used optical character recognition (OCR) tool for Fraktur, ABBYY FineReader, could not be used, for technical and financial reasons – the software does not run on a Mac operating system and its license is too expensive for a one-off text conversion. Moreover, any character recognition errors would have been difficult for the researcher to identify because of the language barrier. Consequently, the text was manually transcribed into an editable document.

As previously mentioned, reading the text was initially complicated because it required getting used to the typeface. In the beginning, the transcription process was considerably slow, because every letter had to be distinguished from the others, especially those that are graphically very similar; however, eventually the researcher got accustomed to the script and proceeded at a faster pace. Transcribing an unknown language requires great attention and a relatively slow pace, but once the most common words become familiar, the process becomes quicker. Typing errors and mistakes were corrected with the help of a spellchecker. There were a few cases where spelling mistakes were identified in the post-editing stage because the raw output of a specific verse was wrong or out of context. In fact, correct spelling was always checked first whenever the corresponding raw output was difficult to understand. The transcription strictly followed the page layout of the book, so that the verses could be easily located in the book for verification purposes. This was very important, since the researcher could not understand the source text and orient himself accordingly.

Once the transcription was completed, the source text was ready for translation. It was decided to use an open-access neural MT tool, DeepL, which has been found to be the best performing engine for the German–English language pair (Savenkov 2019). The translation was carried out in batches of not more than 10 verses at a time; this allowed the researcher to make sure that no verses were skipped or only partially translated due to errors in copying and pasting from the text document to the browser and vice versa. At this stage it was important to make sure that the verses had been fully translated. This was done by checking that the raw output reflected the source text in terms of punctuation and sentence completeness, but no other validation of the translation output was carried out. The process was time consuming, but the lack of source language competence required a particularly cautious approach. It also ensured that no errors were made in the input stage and provided a first, basic check of the completeness of the output.

Subsequently, the raw output from DeepL was post-edited by the researcher to obtain a version that could be useful for an analysis of the plot and narrative structure of Kruse's drama. This was done by resorting to online dictionaries and other resources, as well as by consulting colleagues from the University of Malta's Department of German. Such a step was crucial for this specific study, because the researcher could log all the instances where problems of comprehension arose and had to be solved. The post-edited version was then submitted to an experienced literary translator for revision. This was necessary to obtain a high-quality version to use as a reference for back translation, in the absence of a previously published English translation of the play.

In order to assess the usefulness of the raw output from the MT as a CL research tool, it was necessary to focus on its adequacy for the scope of the study. In this specific case, adequacy meant allowing the researcher to follow the plot and the text with ease. The priority being comprehension, not the linguistic correctness of the target text, the evaluation criteria were necessarily different from the ones normally applied in other areas of research on MT (see, for instance, Daems et al. 2017; Popović 2018). This sort of evaluation, for research purposes, requires a specific focus on certain error types rather than on others that may be equally if not more important for studies on MT quality.

In this study, the term 'comprehensibility issues' is used to refer to the translation errors that hindered the researcher's comprehension of the raw output. This is because a distinction was made between MT errors that did not create comprehensibility problems for the researcher and those which made it necessary for the researcher to consult dictionaries

and other resources in order to follow the text. In many cases, errors that would normally require correction in the post-editing process were not an obstacle to comprehension, since the meaning could be taken from the context. Two examples where translation errors were not counted as comprehensibility issues are '*Erster Aufzug*' and '*Erster Auftritt*' (Kruse 1876, 5), which were translated by DeepL as 'first elevator' and 'first appearance' respectively. From the context, it was clear that their meanings were 'Act One' and 'Scene One', so they did not pose any obstacles to comprehension. All the other instances where *Aufzug* and *Auftritt* were translated in a similar way were thus ignored.

The peculiar nature of this study required an ad hoc taxonomy of comprehensibility issues. These were categorised as follows:

- Lexical: Where a term hindered comprehension by altering the meaning of a phrase or verse. Examples of such instances are '*Dienertroß*' (Kruse 1876, 18), translated as 'servant's horse' (post-edited as 'group of servants'), and '*Ruf*' (Kruse 1876, 34), translated as 'call' (post-edited as 'reputation').[2]
- Semantic: Where a phrase or verse was unclear and could not be inferred from the context. An example of such issues is '*Doch will ich mir die Wahrheit eingesteh'n, / Ist aller meiner Sorgen schlimmste noch, / Es könnte dennoch nicht Verleumdung sein!*' (Kruse 1876, 40), translated by DeepL as 'But I will admit the truth to myself, / Is the worst of all my worries, / Yet it could not be slander!' The verses were post-edited as 'But I must admit to myself that in truth, / The worst of all my worries / Is not the defamation'.
- Grammatical: Where there was ambiguity, mainly caused by incorrect pronoun agreement. An example of such instances is 'The Doge does it like a chestnut woman, / That makes a great noise before the judges, / Showing the naked parts of *his* head, / Where the neighbour tore out *his* hair, / But quite forgets to tell the judges, / That it was *she* who started it / And went into the neighbour's hair'. The italicised pronouns should all be feminine because they refer to the 'chestnut woman' (a female peasant). The source passage reads: '*Der Doge macht's wie ein Kastanienweib, / Das vor den Richtern großen Lärm verführt, / Die nackten Stellen auf dem Kopfe zeigend, / Wo ihm die Nachbarin die / Haare ausriß, / Doch ganz vergißt den Richtern zu erzählen, / Daß sie es war, die angefangen hat / Und in die Haare fuhr der Nachbarin*' (Kruse 1876, 18). Another example is the translation of '*Ich kann sie fragen*

– *(unwillig mit dem Fuße stampfend)*' (Kruse 1876, 42) as 'I can ask her – (reluctantly stamping *my* foot)', where the italicised pronoun should read 'his'.

- Word order: Where awkward word order caused ambiguity or lack of clarity. Since German syntax differs considerably from that of English, certain phrases and verses that were translated literally were not easily understood. An example of word-order comprehensibility issues is '*Dann laß ich läuten mit der großen Glocke, / Die nur auf meinen schriftlichen Befehl / Gezogen werden darf*' (Kruse 1876, 81), which was machine-translated as 'Then let me ring with the great bell, / Which may only be rung by my written command / To be pulled'. The post-edited version read: 'Then I will give orders to ring the great bell, / Which may only be rung by my written command'. Another example is '*Jawohl, an seinem Thron, / Auf welchem fremde Fürsten und Gejandte / Im Staat der doge zu empfangen pflegt*' (Kruse 1876, 5), which was translated as 'Yes, at his throne, / On which foreign princes and envoys / In the state of the doge to receive' and post-edited as 'Yes, on his throne, / Where the Duke receives foreign princes and envoys / As head of state'.
- Cultural: Where culture-specific terms, expressions or allusions remained unadapted or obscure. For example, the noun *Fuder* in '*Und Trunk'nen muß sogar ein Fuder Heu / Ausweichen*' (Kruse 1876, 9) refers to an old liquid measure, which was mainly used for wine. However, DeepL translated the verse as 'And drunkenness must even win a cartload of hay / Evade it'. Originally, *Fuder* meant the carriage load of a two-horse carriage. However, the measure was also used for many other products, such as hay, ore or coal, or as a meadow measure (the area that provided one load of hay). As a measure for wine, there were large regional differences, which varied between about 800 to 1,800 litres. Today the term refers to a type of wine barrel still in use in Germany. Since this is a pun to mean that the person addressed must avoid drinking, it was post-edited as 'And drunkards must avoid even a barrel'. The pun is lost in translation, but the meaning is thus conveyed to the target reader. Another example is the word *Flatterhans* in '*Der Flatterhans hat mich verlassen*' (Kruse 1876, 37). The word is an appellation referring to levity or inconstancy and was left untranslated in the raw output. It was then post-edited as 'shallow man' in the phrase 'That shallow man left me'.

- Additions: Terms, strings of words or phrases, mostly repetitions, which are not found in the source text. Verification was needed to make sure that these repetitions were not used for emphasis in the German version. An example is DeepL's translation of '*Als Sodom und Gomorrha / Der Sünden voll war, wollte die Gerechten / Der Herr doch scheiden von den Ungerechten*' (Kruse 1876, 82) as 'When Sodom and Gomorrah / Were full of sins, the Lord would separate / The Lord wanted to separate the righteous from the unrighteous'. Another example is '*Ihr sollet alles Volk in Kanaan / Vertilgen mit dem Schwert, verschonet Niemand*!' (Kruse 1876, 82), which contains the following italicised addition: 'You shall *destroy* all the people of Canaan / Destroy with the sword, spare none!'. Another instance of an addition is '*Heut Morgen fanden sich am Dogenstuhl / Schmähverse angeschrieben*' (Kruse 1876, 5), which was rendered by DeepL as 'This morning at the Doge's Chair / Verses of invective written on the Doge's chair' and post-edited as 'This morning / Offensive words were written on the Doge's throne'.
- Non-translations: Where untranslated words, left in German, hampered comprehension. An example is 'Up the stairs I found myself / *Grad*' on the place of execution between the two columns' (German: '[…] *da fand ich mich / Grad' auf dem Richtplatz zwischen beiden Säulen*', Kruse 1876, 28). The post-edited version read: 'Up the steps I found myself / On the place of execution between the two columns'. Another example is '*Ab*. Behind him *Steno*' (German: '*Ab. Hinter ihm Steno*', Kruse 1876, 21), which was post-edited as 'He leaves. Steno follows him out'. Untranslated words in Italian found in the source text were not considered as comprehensibility issues, since the researcher understands the language.
- Mistranslations: Where the translation in the raw output was pragmatically incorrect. An example is 'We didn't need to punish Steno at all; / The slight cancels itself out' for '*Wir brauchten Steno gar nicht zu bestrafen; / Die Kränkung hebt sich auf*' (Kruse 1876, 23). These verses were post-edited as 'We did not need to punish Steno at all; / The offence will be forgotten'. Another one is the German verse '*Kein finst'rer Ernst hält seiner Laune Stand*' (Kruse 1876, 44), translated as 'And he's not a man to be taken lightly'. The post-edited version was 'No seriousness can withstand his humour'.

Every comprehensibility issue encountered during the post-editing process was colour coded. The identification and categorisation of each instance was inevitably subjective, since what may be an issue for one

reader may not be so for another, and some of the issues were not clear-cut cases in terms of which category they fitted into. A case in point is the above-mentioned example of *Flatterhans*, where the issue could well have been classified in the 'Non-translations' category. Consequently, the number of issues per category should not be considered in absolute terms. What is indicative is the proportion between them, which sheds light on the overall nature of the challenge posed by using MT to understand a nineteenth-century German historical drama.

Findings and analysis

The table below lists the total number of comprehensibility issues encountered in the raw output provided by DeepL, together with the number of occurrences for each category described above.

As can be seen from Table 2.1, a total of 380 comprehensibility issues were encountered. Of these, the most frequent were semantic issues and mistranslations. Lexical issues were less frequent, but their percentage is still significant. The prevalence of these categories is to be expected, since they are the ones that most depend on context and require the pragmatic competence of a human translator. An unexpected indication from these results is the relatively high percentage of additions, which, despite not impeding overall comprehension, had to be checked every time to make sure that they were not found in the source text and that they were not a stylistic expedient used by the author.

The most significant result of this study from a CL perspective is that the translation of Heinrich Kruse's historical tragedy *Marino Faliero*

Category	Frequency	Percentage
Semantic	100	26.32%
Mistranslations	86	22.63%
Lexical	64	16.84%
Additions	52	13.68%
Grammatical	28	7.38%
Non-translations	20	5.26%
Word order	16	4.21%
Cultural	14	3.68%
Grand Total	**380**	**100%**

Table 2.1 Occurrences of comprehensibility issues in raw output, in descending order

using DeepL achieved the goal of allowing the researcher to obtain direct access to the text despite not knowing the source language. The raw output required post-editing to clarify comprehensibility issues, as described above. However, considering that the text is in verse form and was written in nineteenth-century German, and considering its length, the number of issues was deemed reasonable. The post-editing process required time and external assistance from German specialists and a reviser. The latter's intervention was especially important for the identification of some expressions that had not been adequately rendered by DeepL and could not be identified as such by the researcher. A case in point was the rendering of '*Dem Dogen mag die Galle überlaufen*' (Kruse 1876, 7), which in the raw output was rendered as 'The Doge's gall may run' and was post-edited as 'The Doge's gall may be too much'. However, the reviser noted that the source text refers to the expression '*jemand läuft die Galle über*', which resulted in the final translation reading 'The Doge may well be very angry'. Such instances were counted as semantic incomprehensibility issues, since they altered the meaning of the specific phrase. Apart from possible better translation alternatives, the reviser identified only 9 mistranslations, such as the example above, and 16 partially correct renderings, mostly due to missing modifiers such as adjectives and adverbs. At the end of the process, the translation was deemed adequate for research purposes, allowing for a reliable analysis of the narrative elements of the text, which was the ultimate scope of the exercise.

The applicability of the model above is subject to many conditions. Firstly, its feasibility depends on the time at the researcher's disposal. The time element is crucial when deciding whether translating a work this way would be advisable or not, and if the process would benefit the CL research project at hand. In most cases, the time-consuming transcription phase would be much shorter, since most texts are printed in fonts for which OCR is widely available and only revision would thus be required. A more important factor is the availability of human assistance in the source language. In this study, the researcher could rely on the help of German scholars and a literary translator to solve issues that could not be settled by an online search. If the source text had been written in a lesser-used language for which no human assistance was available, the final target text would not have been validated and thus would have been unreliable. Other limitations may be the length of the source text and its stylistic characteristics, especially if it relies heavily on metaphors, allegories, allusions, neologisms, unconventional use of language, and the sound and rhythm of the words.

Conclusion

Using MT to access literary works in a language which a researcher cannot understand comes with many caveats. Before making such an attempt, the feasibility of obtaining a useful target text must be carefully evaluated. Not all literary genres and texts can be machine-translated usefully for comparative purposes. Despite the significant improvements brought about by neural MT, especially for the most widely used languages, the raw output must be relied upon with great caution, because not all mistranslations are easily identifiable. This renders the interpretation of the target text much more difficult; mistakes can easily be made. If a source text is strongly characterised by figurative language, wordplay, puns, allusions and other stylistic elements, it would not be a suitable candidate for MT. The time, effort and expertise required to post-edit the target text would make the exercise unviable, and anything short of professional human literary translation would not be enough to attain reliable access to the source text. However, the MT post-editing of Kruse's historical drama *Marino Faliero* was successful enough to provide a reliable target text. The length and format of the text –around three thousand unrhymed verses mainly consisting of dialogue – made it suitable for the case study. The fact that the source text was written in the nineteenth century did not have a significant bearing on the results, since most comprehensibility issues were not determined by archaic or obsolete words. In most cases, the obstacles to comprehension were due to an inadequate link between the phrase or sentence being translated and its wider context. This gave rise to semantic, syntactic and grammatical issues, as well as outright mistranslations that hindered reading and comprehension of some passages of the target text.

Notwithstanding its shortcomings, which could be expected given that it is a work in progress, MT should not be ignored as a tool for research from a CL perspective. As long as expectations are not too high and objectives are set to realistic levels, it can give researchers access to texts in languages that were hitherto inaccessible to them. In the case study described above, the researcher had two options: to keep aloof from Kruse's drama or to try MT and see if he could work with its output. The result was useful and could help one gain insights on the way that Kruse dealt with his subject matter, which could then be compared with strategies used by other writers. Translation technology is not a tool that is often looked upon favourably by literary scholars, but in certain cases it may be useful to overcome the language barriers that have limited the reach of researchers in CL for a long time.

Notes

1 No English translations have been found in the catalogues of the British Library and the Library of Congress.
2 Back translations and post-editing consider the context of the passage where the relevant words, phrases or sentences are found.

References

Apter, Emily. 2006. *The Translation Zone: A new comparative literature*. Princeton, NJ: Princeton University Press.

Bassnett, Susan. 1993. *Comparative Literature: A critical introduction*. Oxford: Blackwell.

Bassnett, Susan, and André Lefevere, eds. 1990. *Translation, History and Culture*. London: Cassell.

Bermann, Sandra. 2009. 'Working in the And Zone: Comparative literature and translation'. *Comparative Literature* 61: 432–46.

Bernheimer, Charles, ed. 1995. *Comparative Literature in the Age of Multiculturalism*. Baltimore, MD: Johns Hopkins University Press.

Besacier, Laurent. 2014. 'Traduction automatisée d'une oeuvre littéraire: une étude pilote'. In *21ème Traitement Automatique des Langues Naturelles*, Association for Computer Linguistics, 389–94.

Castilho, Sheila, Joss Moorkens, Federico Gaspari, Iacer Calixto, John Tinsley, and Andy Way. 2017. 'Is neural machine translation the new state of the art?' *Prague Bulletin of Mathematical Linguistics* 108: 109–20.

Castilho, Sheila, and Natália Resende. 2022. 'Post-editese in literary translations', *Information* 13, no. 2: #66. https:// doi.org/10.3390/info13020066.

Daems, Joke, Sonia Vandepitte, Robert J. Hartsuiker, and Lieve Macken. 2017. 'Identifying the machine translation error types with the greatest impact on post-editing effort', *Front Psychol* 8: #1282. https://doi.org/10.3389/fpsyg.2017.01282.

D'hulst, Lieven. 2007. 'Comparative literature versus translation studies: Close encounters of the third kind?' *European Review* 15, no. 1: 95–104. https://doi.org/10.1017/S1062798707000099.

Even-Zohar, Itamar. 1990. 'The position of translated literature within the literary polysystem', *Poetics Today* 11, no. 1: 45–51. https://doi.org/10.2307/1772668.

Genzel, Dmitriy, Jakob Uszkoreit, and Franz Och. 2010. '"Poetic" statistical machine translation: Rhyme and meter'. In *Proceedings of the 2010 Conference on Empirical Methods in Natural Language Processing*, edited by Hang Li and Lluís Màrquez, 158–66. Cambridge, MA: Association for Computational Linguistics.

Greene, Erica, Tugba Bodrumlu, and Kevin Knight. 2010. 'Automatic analysis of rhythmic poetry with applications to generation and translation'. In *Proceedings of the 2010 Conference on Empirical Methods in Natural Language Processing*, edited by Hang Li and Lluís Màrquez, 524–33. Cambridge, MA: Association for Computational Linguistics.

Hermans, Theo, ed. 1985. *The Manipulation of Literature: Studies in literary translation*. London: Croom Helm.

Jones, Ruth, and Ann Irvine. 2013. 'The (un)faithful machine translator'. In *Proceedings of the 7th Workshop on Language Technology for Cultural Heritage, Social Sciences, and Humanities*, 96–101. Sofia: Association for Computational Linguistics.

Kenny, Dorothy, and Marion Winters. 2020. 'Machine translation, ethics and the literary translator's voice', *Translation Spaces* 9, no. 1: 123–49. https://doi.org/10.1075/ts.00024.ken.

Kruse, Heinrich. 1876. *Marino Faliero: Trauerspiel in fünf Aufzügen*. Leipzig: S. Hirzel.

Kruse-Jarres, Jürgen D. 2008. *Heinrich Kruse: Journalist und Schriftsteller Ein Kämpfer für den Liberalismus im 19. Jahrhundert*. Düsseldorf: Grupello Verlag.

Lefevere, André. 1982. 'Mother courage's cucumbers: Text, system and refraction in a theory of literature', *Modern Language Studies* 12, no. 4: 3–20. https://doi.org/10.2307/3194526.

Lefevere, André. 1992. *Translation, Rewriting and the Manipulation of Literary Fame*. London: Routledge.

Lefevere, André. 1995. 'Introduction: Comparative literature and translation', *Comparative Literature* 47, no. 1: 1–10.

Moorkens, Joss, Antonio Toral, Sheila Castilho, and Andy Way. 2018. 'Translators' perceptions of literary post-editing using statistical and neural machine translation', *Translation Spaces* 7, no. 2: 240–62. https://doi.org/10.1075/ts.18014.moo.

Ning, Wang, and César Domínguez. 2016. 'Comparative literature and translation: A cross-cultural and interdisciplinary perspective'. In *Border Crossings: Translation studies and other disciplines*, edited by Yves Gambier and Luc van Doorslaer, 287–308. Amsterdam: John Benjamins.

Popović, Maja. 2018. 'Error classification and analysis for machine translation quality assessment'. In *Translation Quality Assessment: From principles to practice*, edited by Joss Moorkens, Sheila Castilho, Federico Gaspari, and Stephen Doherty, 129–58. Cham: Springer.

Savenkov, Konstantin. 2019. 'State of the machine translation', Intento blog, 19 June 2019. Accessed 3 April 2024. https://blog.inten.to/state-of-the-machine-translation-june-2019-e3ffb457b76c.

Taivalkoski-Shilov, Kristiina. 2019. 'Ethical issues regarding machine(-assisted) translation of literary texts', *Perspectives* 27, no. 5: 689–703. https://doi.org/10.1080/0907676X.2018.1520907.

Tee, Carlos G. 2012. 'Towards a new and more constructive partnership: The changing role of translation in comparative literature', *Fu Jen Studies: Literature and Linguistics* 45: 1–21.

Toral, Antonio, and Andy Way. 2015. 'Translating literary text between related languages using SMT'. In *Proceedings of NAACL-HLT Fourth Workshop on Computational Linguistics for Literature*, edited by Anna Feldman, Anna Kazantseva, Stan Szpakowicz, and Corina Koolen, 123–32. Denver, CO: Association for Computational Linguistics.

Toral, Antonio, and Andy Way. 2018. 'What level of quality can neural machine translation attain on literary text?' In *Translation Quality Assessment: From principles to practice*, edited by Joss Moorkens, Sheila Castilho, Federico Gaspari, and Stephen Doherty, 263–87. Cham: Springer.

Toral, Antonio, Martijn Wieling, and Andy Way. 2018. 'Post-editing effort of a novel with statistical and neural machine translation', *Frontiers in Digital Humanities* 5: #9. https://doi.org/10.3389/fdigh.2018.00009.

Venuti, Lawrence. 2008. *The Translator's Invisibility: A history of translation*. 2nd ed. London: Routledge.

Zanettin, Federico. 2017. 'Issues in computer-assisted literary translation studies', *inTRAlinea*, Special Issue: Corpora and Literary Translation, edited by Titika Dimitroulia and Dionysis Goutsos. Accessed 3 April 2024. https://www.intralinea.org/specials/article/2254.