

The Human Translator's (In)Visibility in Times of AI: An Self-Applied Experiment in Literary Translation

A (in)visibilidade do tradutor humano em tempos de IA: um experimento autoaplicado em tradução literária

Cynthia Beatrice Costa

Universidade Federal de Uberlândia
(UFU) | Uberlândia | MG | BR
cynthiacos@gmail.com
<https://orcid.org/0000-0002-3063-0121>

Igor Antônio Lourenço da Silva

Universidade Federal de Uberlândia
(UFU) | Uberlândia | MG | BR
ialsigor@gmail.com
<https://orcid.org/0000-0003-0738-3262>

Abstract: Based on a self-applied experiment comparing five different English–Brazilian Portuguese translations of a microfiction, we discuss the (in)visibility of the literary human translator's work in the age of AI. More specifically, we aim to tap into the interferences in meanings produced by machines, if any, and their impact on the translation process and product (particularly in terms of multiple literary text imports, adequacy, fluency, and post-editing effort). To this end, we designed a self-applied experiment in which the same source text was translated by two humans, two machine translation (MT) engines, and two generative AI chatbots. We assessed the outputs as to their similarities and differences for specific linguistic traits, and their authorial traits. Despite variations in the standard metrics BLEU and TER, both AI chatbots provided human-like outputs, except for crucial interferences that could both hinder the understanding and enjoyment of the literary text and obscure the human translators' role as both prompt creators and post-editors. Machine-related red flags included the management of deictic possessive adjectives and pronouns by both MT engines, the addition of an undesirable em dash, and the fabrication of a title by one of the AI chatbots. Our findings seem to anticipate a novel instance of translator's invisibility, challenging the human translator's degree of engagement in the task.

Keywords: literary translation and AI; translator's visibility; prompt creation; post-editing.



Resumo: Partindo de um experimento autoaplicado que compara cinco diferentes traduções inglês-português brasileiro de um microconto, discutimos a (in)visibilidade do trabalho do tradutor literário humano em tempos de IA. Mais especificamente, tivemos por objetivo investigar as interferências, se alguma, de significados produzidas pelas máquinas e o impacto destas no processo e produto tradutório (sobretudo em termos de multiplicidade de significados literários, adequação, fluência e esforço de pós-edição). O mesmo texto-fonte foi traduzido por dois humanos, dois serviços *online* de tradução automática e dois chatbots de IA gerativa. Os *outputs* foram analisados quanto às semelhanças e diferenças relativas a determinados traços linguísticos e às marcas autorais. Apesar de variações nas métricas padrão BLEU e TER, ambos os *chatbots* forneceram um resultado próximo ao do humano, exceto por intervenções cruciais que podem prejudicar o entendimento e a fruição do texto, bem como invisibilizar o papel do tradutor literário como administrador dos *prompts* e pós-editor. Os pontos problemáticos incluem a administração dos pronomes possessivos dêiticos por ambos os serviços de tradução automática, a adição de um travessão indesejável e a fabricação de um título por um dos *chatbots*. Nossos resultados sugerem uma nova possível invisibilização do tradutor, cujo grau de participação na tarefa é colocado em dúvida.

Palavras-chave: tradução literária e IA; visibilidade do tradutor; criação de *prompt*; pós-edição.

Introduction

Domestication as proposed by Venuti (1995) in his revisitation of Schleiermacher refers to a tendency toward textual fluency, which he noted in literary translations produced for the Anglo-American market in the late 20th century. Such fluency would give the reader the artificial impression that the text was originally written in English, thus promoting the erasure of its foreign origin. Another side effect of this artificial fluency would be the invisibility of the human translator, whose personal imprint disappears when the text does not show much, if any, marks of translation.

Thirty years later, the translation landscape has changed significantly. Literary human translators seem to have gained some visibility (see Barbassa, 2023; Hoyte-West, 2023). At

the same time, they now seem to find themselves threatened by the rapid and seemingly relentless rise of generative artificial intelligence (hereinafter, AI), largely available to anyone with internet access.¹ It is no longer possible to tell for sure whether a human being or a machine has translated a text, even a literary one (see Visby, 2020; Hadley *et al.*, 2019); studies have been showing that machines produce translations as acceptable as those rendered by humans (Humblé, 2019) and that the machine errors that require human post-editing (about post-editing, see Plitt; Masselot, 2010; Moorkens *et al.*, 2015; Silva *et al.*, 2015; Koglin *et al.*, 2023) are comparable to errors found in human translations (Fischer; Läubli, 2020). In a survey on machine-translated subtitles, Calvo-Ferrer (2023) found that participants were unable to accurately distinguish between ChatGPT-generated and human-generated subtitles.

State-of-the-art machines, whether in the form of AI chatbots (ChatGPT, DeepSeek, Gemini etc.), computer-aided translation (CAT) tools (MemoQ, Phrase, Trados, Wordfast, etc.), or online machine translation (MT) engines (DeepL, Google Translate etc.), can translate well enough to require a rethinking of the human translator's task. Just like in the "scandals" denounced by Venuti (1998) in the late 1990s, the human translator's (in)visibility may now take a different form: as the AI output seems so human, it can potentially erase not only the role of the human translator both in prompt creation and post-editing, but also in the act of translation itself, making the human translator's engagement in the process invisible.

Even though machines can now provide fluent text, they may also undesirably interfere with the text meanings, particularly the literary text import. To tap into such interferences, if any, and their impact on the translation process and product (particularly in terms of literary text import, adequacy, fluency, and post-editing effort), we designed a self-applied experiment whereby we compare five different translations of the same source text: one performed by humans (a literary translator, the first author, with the collaboration of an academic text translator, the second author), two provided by multilingual neural machine translation engines (namely, Google Translate and DeepL), and two provided by generative artificial intelligence chatbots (namely, ChatGPT and Gemini in their paid versions).

By adopting a qualitative approach, we analyze multiple translations of the same microfiction originally written in contemporary American English into contemporary Brazilian Portuguese, as a means of contrasting both process and product in the literary translation performed by humans and by machines. The microfiction chosen for this research was "Honeysuckle," by Ophelia Hu Kinney (2020), published on the *NYC Midnight* website, a writers' forum that hosts literary competitions. Kinney's story was one of the winners of a 2020 100-word microfiction challenge.

We tapped into common factors considered in this type of study (see Liu; Zhang; Wu, 2021) to assess approval or satisfaction with the translation renditions, namely: standard metrics BLEU and TER, adequacy and fluency of the machine output, and post-editing effort, i.e., how much cognitive and technical effort it takes to render a final, publishable version of the literary text. We also assessed the import of the literary text in the target language (Sperber; Wilson, 1995).²

¹ We could discuss the implications of using paid or free services, but we believe it does not affect our major point in this article.

² Even though we use relevance theory as our theoretical framework, this paper does not delve into the tenets of this framework, as our aim is to be as accessible as possible to the general audience in the present discussion.

As professors concerned with the education of undergraduates in translation – particularly in literary translation courses (first author) and translation technology courses (second author) –, we have witnessed the indiscriminate use of machine outputs by students without the prior acquisition of translation competence and prior understanding of the underpinnings, tenets, and limitations of translation engines. We understand that academia is (and will always be) one or two steps behind the market (Pym, 2003) and that we will never catch up to the latest developments in the language industry, especially in times of AI, predominantly developed by utilitarian computer scientists less concerned with Translation Studies, Linguistics, or Literature than with good-enough or fit-for-purpose massive data mining and/or processing. However, if we explore translation and related activities (including interpreting, pre-editing, post-editing, proofreading, and revision) as both process and product, we can be better off at training students who understand that machines deal with probabilities, but not with meanings, and that adaptive expertise (Angelone, 2022; Silva; Costa, 2020; Alves; Silva, 2021, 2024; Silva *et al.*, 2023) is of the essence in the current language industry. Since the literary text is characterized as a text with multiple imports, the human being remains center stage, playing a significant role in interpreting the source text and delivering a final translated text that is as meaningful to the target audience’s cognitive environment as the source text is to the source audience’s cognitive environment (in simple words, their prior, readily available, context-relevant knowledge). In relevance theory, the import is the overtly intended cognitive effect of an ostensive stimulus and is made up of different mental representations depending on the addressee’s prior knowledge (Sperber; Wilson, 1995).

In the following section, we provide our tests and analyses, which also reveal that expert human translators do have insecurities and must cope with several challenges in their translation process. We discuss these findings in the Discussion section while also providing the results for the standard metrics TER and BLEU.

Tests

In this section, we report on our analysis in the following order: human translation, MT outputs, and AI chatbot outputs.

Human Translation

The first step involved choosing the microfiction sample and translating it. The choice took days, because we were looking for a very short text that encompassed narrative and literary defining characteristics, such as dialogue, metaphor, and cultural marks. Once chosen through several online searches, we turned to the translation task, which took us approximately two hours and aimed at as much concision as possible (approximately 100 words to comply with the brief). In fact, the translation was mainly performed by the first author, who has been a literary translator for 15 years and has rendered nearly 50 books from English into Brazilian Portuguese, with later collaboration by the second author, an academic translator with almost 20 years of experience who has translated several articles and books in the Translation Studies, Medicine, Accounting, and Administration domains.

Human translation is not imagined outside the digital environment anymore: it is at least performed in word editing programs, such as Microsoft Word, and it often relies on online search tools, such as Google and digital dictionaries. The translation performed here was no different: following her regular practice, the first author utilized Word to write and Google search engine and Cambridge online dictionaries to clarify vocabulary. The result was as follows:

Table 1 – Human translation

Honeysuckle	Madressilva
Soon, she'll land. Her father will be shorter. Her mother's bones, lighter. "Aiyaa, it's been five years." It'll need to be said. At their apartment, she'll unpack her suitcase: gifts, vitamins, medicine for their restless legs.	Logo, ela aterrissará. Seu pai estará mais baixo. Os ossos da sua mãe, mais leves. "Aiyaa, faz cinco anos". Precisar ser dito. No apartamento deles, ela desfará a mala: presentes, vitaminas, remédios para as pernas inquietas dos pais.
Her mother will kiss her. "We have these things in China now." In two weeks, they'll repack her suitcase with the life she could've had: water spinach seeds, photographs, native honeysuckle cuttings for flowers that will make her weep.	Sua mãe a beijará. "Temos dessas coisas na China agora". Em duas semanas, arrumarão de novo a mala da filha com a vida que ela poderia ter tido: sementes de espinafre d'água, fotografias, mudas de madressilva nativa para flores que a farão chorar.
In five more years, the honeysuckle will cross her fence. Its children will grow toward the highway, and then the ocean, and then, and then...	Em mais cinco anos, a madressilva atravessará a cerca do seu jardim. Seus ramos crescerão rumo à rodovia, daí ao oceano, daí, daí...

Source: The authors.

Some choices were not obvious for this somewhat artisanal, handcrafted translation, written directly onto the computer screen with no machine assistance.³ One concerned the use of em dashes, which are more commonly applied to dialogues in Brazilian narratives, instead of quotation marks. The choice was to keep the quotation marks for stylistic reasons in microfiction, since they take up less space.

A common issue when translating a text from English into Portuguese then arose: the use of deictic possessive adjectives and pronouns. "In their apartment" could be "*Em seu apartamento*" [In their/his/her/its apartment], but this would not clarify whose apartment it is, hence the use of "*deles*" [of them]. In Brazil we often say "*na casa de*" [in the house of (someone)] to refer to a residence, even if it is an apartment or another type of dwelling. However, we decided to be as specific as the original: "*apartamento*" [apartment].

Next, another deictic possessive adjective problem emerged: "for their restless legs." Translated as "*suas pernas inquietas*" [their/his/her/its restless legs], it could be understood as the main character's legs, not her parents', hence the need for explicitation ("*pernas inquietas dos pais*" [restless legs [of the parents]]). The same problem recurs at a crucial point in the story, in "her fence" and "Its children." The first author decided once more for explicitation, making sure that

³ We could have resorted to keylogging or screen-recording applications, as in a senior thesis involving shortest short stories supervised by the second author (Domingos, 2016), but we opted to retain ecological validity as much as possible.

the reader understands that the fence is hers (i.e., the main character's) by using "*a cerca de seu jardim*" [the fence [of her garden]] – also, she thought that the explicitation of "*jardim*" [garden], even though absent (but inferable) in the original, would suggest a poetic image. In Brazilian Portuguese, there is no differentiation of the deictic possessive adjectives for inanimate beings, so "*Seus filhos*" could be more likely understood as the protagonist's children, instead of the plant's, which made the first author choose "*ramos*" [branches], making the metaphor – which we understood as the idea of generations to come – less accessible in the reader's cognitive environment but hopefully still active in its potentiality (as a weaker hypothesis).

From a cultural perspective, the choice of maintaining the spelling but adding italics to "*Aiyaa*" for the first utterance in the narrator's imagined dialogue was deliberate and conscious. Even though probably not known by many Brazilians (including us, the human translators, prior to this task), the Chinese interjection is a cultural mark in the story and, in our view, should be manifested in the target text as it is. Context allows for the reader to conclude it as an interjection, especially when used in italics. A choice could also have been made to introduce an exclamation point, but we opted not to change stylistic markers as much as possible.

Still regarding the lexicon, "water spinach" and "honeysuckle" raised insecurities. Using Google's search engine, we found that "water spinach" is *Ipomoea aquatica*, which made it easier to locate it as the literal "*espinafre d'água*," although it also goes by other names in Brazil, such as "*agrião do paraíso*" and "*agrião d'água*." Our concern was to verify if the "literal" term can allude to Chinese culture, which it does.

Honeysuckle is a generic name for several species of climbing plants, commonly found in Central and East Asia. The corresponding generic term in Brazil is *madressilva*, although we use it incorrectly for the *Lonicera japonica* species (possibly the one featured in the story), since in fact *madressilva* would more accurately refer to the European plant *Lonicera periclymenum*. We tried to confirm the information by consulting a botanical dictionary, but to no avail. Based on a notion of literary equivalence (see Britto, 2012) – replacing a generic term with an equally generic one –, we adopted *madressilva* for honeysuckle.

The use of "*daí*," common in Brazilian narratives and in orality, as a translation of "then" in the last sentence is arguably the most creative choice in our translation of the microfiction. A more obvious choice would have been "*depois*" [then], but we decided that "*daí*" would be more unexpected (perhaps more of a regionalism) and thus more authorial, something that we have discussed elsewhere (Costa; Silva, 2020) regarding the ethics of using machine translation.

At this point in the research process, we did not know how translation engines would behave with respect to these terms, nor how different (or similar) their outputs would be. In the following section, we report on the outputs provided by the online machine translation services Google Translate and DeepL.

Machine Translation Services

On the same day (July 17th, 2025) we translated the source text from scratch, we also used the online neural machine translation engines selected.

Applying machine translation to literary texts is a highly controversial topic (Costa; Silva, 2020). It involves the ethical question of authorship, in addition to the obvious impediment of the machine to finding truly creative solutions – if we both consider creativity as a

human skill (see Bohm, 1998; Claveanu *et al.*, 2013) and acknowledge that the machine neither understands nor interprets; it is based on probabilities. There is also the matter of replacing a thoroughly enjoyable human activity with a robotic one, which we have addressed elsewhere (Costa; Silva, 2020).⁴ Nevertheless, the continuous improvement of engines like Google Translate and DeepL has been undeniable and impressive. The use of such engines for translating literature has been tested for at least a decade (see Salimi, 2014; Toral; Wieling; Way, 2018). In fact, their terminology and grammatical correctness are barely an issue nowadays, but fluency and cultural accuracy can sometimes still be a challenge.⁵

Table 2 shows our first trial, using DeepL:

Table 2 – DeepL’s raw output

Soon, she’ll land. Her father will be shorter. Her mother’s bones, lighter. “Aiyaa, it’s been five years.” It’ll need to be said.	Em breve, ela vai aterrar. O pai dela vai parecer mais baixo. A mãe vai parecer mais magra. «Aiyaa, já passaram cinco anos.» É preciso dizer isso.
At their apartment, she’ll unpack her suitcase: gifts, vitamins, medicine for their restless legs.	No apartamento, ela vai desfazer a mala: presentes, vitaminas, remédios para as pernas inquietas.
Her mother will kiss her. “We have these things in China now.” In two weeks, they’ll repack her suitcase with the life she could’ve had: water spinach seeds, photographs, native honeysuckle cuttings for flowers that will make her weep.	A mãe vai beijá-la. «Agora temos estas coisas na China.» Em duas semanas, irão voltar a fazer a mala com a vida que ela poderia ter tido: sementes de espinafre-d’água, fotografias, mudas de madressilva nativa para flores que a farão chorar.
In five more years, the honeysuckle will cross her fence. Its children will grow toward the highway, and then the ocean, and then, and then...	Daqui a mais cinco anos, a madressilva irá atravessar a sua cerca. Os seus filhos irão crescer em direção à autoestrada, depois ao oceano, e depois, e depois...

Source: The authors.

This raw output—which contains some lexical choices that hint at European Portuguese (e.g., “aterrar,” “autoestrada”) instead of Brazilian Portuguese – led to several considerations. The first one concerns the frequent use of the verb “*ir*” in the present tense as an auxiliary form that functions as the future of the main verb (as in “*vai aterrar*” [[she] goes[present] to land/ground]), unlike the human choice for the morphologically realized (“-á”) future tense in the main verbs themselves (e.g., “*aterrissará*” [land [future]]). This difference causes a change in style. The passage “*irão voltar a fazer a mala*” [[they] go[future] to return to packing the suitcase] turned out particularly truncated, with three verbs in a row.

⁴ Much has been said about the dangers of substituting human activity with tasks performed mainly by robots. “I believe that we underestimate the powers of AI at our own peril,” as Visby (2020, p. 28) puts it.

⁵ Collins (2018) says the following about the state-of-the-art technology: “What of the world of artificial intelligence as it is currently presented to the public? This is a world of *artificial intelligence*, available through the newspapers, books and films. [...] Brilliant programming can make computers do a lot in the way of simulating language understanding and responding, as the modern devices show – and they are going to get much better. It is still not quite the real thing, however. But showing that it is not quite the real thing, and it is really important to do so, will become harder and harder and the danger of the Surrender will come ever closer.”

We have “*A mãe vai parecer mais magra*” (The mother goes[present] to look thinner) as a translation of “Her mother’s bones, lighter.” This translation may at first seem to be an interpretation/inference (lighter bones are a sign of thinness), restricting the import and entailing a slight change in punctuation (which was intentionally used in the human translation as a means of resembling the original syntax with the resources available in the target language). However, we should stress that the output is by no means indicative of an actual understanding of the machine, but rather a probability-oriented choice based on its training dataset. Besides, in literary terms, this machine rendering is deprived of iconicity and narrows down the possible interpretations and imageries (i.e., literary text import) by the target reader.

In the first sentence of the dialogue, DeepL used the adverb “já” [already], which intensifies the completion of an event in time, and realized the narrator’s following verb in the present tense (“*É preciso dizer isso*” [[It] is needed to say that]). Here we have two arguably poor “strategies” (in fact, only lexicogrammatical realizations, as the machine does not use a strategy in the stricto sensu): (1) the use of the adverb can be understood as an unwelcome intrusion, since the interjection “*Aiyaa*,” also rendered by DeepL, already expresses the idea (albeit implicitly) that it has been a long while; (2) the rendition of the verb group “*É preciso dizer*” in the present tense in the following sentence (“*É preciso dizer isso.*”) sounds “translated” to informed ears due to the unnecessary use of the demonstrative pronoun “*isso*” [that], which not only is usually implicit in such context, but also renders the text unnecessarily ambiguous as to its deixis (i.e., what is “that?”).

DeepL did not specify for whom the medicine for restless legs was (“*remédios para as pernas inquietas*”), whereas – as described in the previous section – this was a concern to the human translator, who resorted to explicitation. The deictic possessive adjective and pronoun usage was almost entirely different from that used in the human translation; DeepL rendered “*a sua cerca*” [their/his/her/its fence], which does not necessarily indicate that the fence belongs to the main character, as it could also belong to her parents. The rendition “*Os seus filhos*” [their/his/her/its children] is unnecessarily confusing/ambiguous: whose children are they? In Brazilian Portuguese they can be understood, for instance, as the main character’s children.

Overall, the first impression is that the translation produced by DeepL is readable and mostly error free from a syntactic perspective (apart from some lexical choices that are reminiscent of European Portuguese and punctuation choices that are uncommon in Brazilian Portuguese, namely the Latin Aspa and the placement of the full stop within the quotation marks). However, such a translation is not satisfactory from a literary perspective, as it tells a different story, one that is unnecessarily ambiguous (including for literary standards) and arguably unnatural in the target language. As a result, the DeepL translation would not serve as a substitute for human translation. Post-editing work would involve too many interventions to consider that the machine output would be useful in this case.

Now, we turn our attention to the Google Translate output in Table 3.

Table 3 – Google Translate’s raw output

<p>Soon, she’ll land. Her father will be shorter. Her mother’s bones, lighter. “Aiyaa, it’s been five years.” It’ll need to be said. At their apartment, she’ll unpack her suitcase: gifts, vitamins, medicine for their restless legs.</p> <p>Her mother will kiss her. “We have these things in China now.” In two weeks, they’ll repack her suitcase with the life she could’ve had: water spinach seeds, photographs, native honeysuckle cuttings for flowers that will make her weep.</p> <p>In five more years, the honeysuckle will cross her fence. Its children will grow toward the highway, and then the ocean, and then, and then...</p>	<p>Em breve, ela pousará. Seu pai será mais baixo. Os ossos de sua mãe, mais leves. “Aiyaa, já faz cinco anos.” Precisa ser dito. No apartamento deles, ela desfará a mala: presentes, vitaminas, remédios para as pernas inquietas. Sua mãe a beijará. “Agora temos essas coisas na China.” Em duas semanas, eles reembalarão a mala dela com a vida que ela poderia ter tido: sementes de espinafre-d’água, fotografias, mudas de madressilva nativa para flores que a farão chorar. Em mais cinco anos, a madressilva cruzará a cerca. Seus filhotes crescerão em direção à rodovia, e depois ao oceano, e depois, e depois...</p>
---	--

Source: The authors.

A mistake in the machine translation is noticeable in the very first paragraph: the rendition “*será*” [be[future, permanent]] instead of “*estará*” [be[future, transitory/punctual]] in the second sentence. Such a problem with the Brazilian Portuguese verbs “*ser*” (related to permanent situations/characteristics/events) and “*estar*” (transitory ones) is expected, since both meanings are realized by the verb “to be” in English.

In the second paragraph, Google Translate renders the same adverb “*já*” [already] as DeepL did. It also uses the verb group “*Precisa ser dito*” ([It] needs to be said) in the present tense, while the original predicts a future event/obligation.

Two phrases stand out in this translation: the verb “*reembalarão*” in future tense [will encase], rarely used in the context of packing luggage, and “*Seus filhotes*” ([Their/His/Her/Its] cubs/pups/kittens) for “Its children,” which sounds unusual since neither plants nor humans have such offsprings. Tentatively, we can only but assume that the engine rendered the latter phrase due to the possessive “Its,” assuming that the noun “children” probabilistically refers to the offspring of an animal in this context (the “Its” in question).

We also highlight how Google Translate used deictic possessive adjectives and pronouns, which generally led to the same specificities deliberately chosen by human translators (except for whose legs are referred to in the unspecific, general “remédios para as pernas inquietas” [medicine for restless legs]). Overall, Google Translate’s output came closer to the human translation, but it would not replace it in the sense of being publishing-ready. It would require some post-editing corrections, such as the use of the verb “*estar*” (to be, transitory/punctual) in the second sentence and the use of a less artificial-sounding verb for packing (instead of “*reembalarão*” [will encase]). Besides, it did not render a text that could be considered as creative as the human translation, mostly attaining to the source text syntax and most probable lexical equivalents in Brazilian Portuguese.

In the next section, we analyze the performance of the AI chatbots.

Chatbots

The main difference between using a machine translation engine and using a chatbot is the possibility in the latter of asking for what one wants or needs in detail, i.e., by providing an unabridged, very specific prompt in natural language. When writing the prompt, however, we tried not to anticipate potential issues, focusing first on a common translation command in Brazilian Portuguese: “Boa tarde! Por favor, traduza o microconto a seguir para o português brasileiro contemporâneo, respeitando o estilo e o conteúdo da narrativa” [Good afternoon! Please translate the following microfiction into contemporary Brazilian Portuguese, respecting both the style and content of the narrative]. We obtained the following answer from ChatGPT:⁶ “Claro! Por favor, envie o microconto que você deseja que eu traduza. Estou pronta para começar assim que você compartilhar o texto” [Of course! Please send the micro-story you want me to translate. I’m ready to start as soon as you share the text].⁷ The input text was then sent and returned in seconds.

Table 4 shows the ChatGPT’s raw output.

Table 4 – ChatGPT’s raw output

Soon, she’ll land. Her father will be shorter. Her mother’s bones, lighter. “Aiyaa, it’s been five years.” It’ll need to be said.	Logo, ela aterrissará. O pai estará mais baixo. Os ossos da mãe, mais leves. — Aiyaa, já se passaram cinco anos. — Será preciso dizer.
At their apartment, she’ll unpack her suitcase: gifts, vitamins, medicine for their restless legs.	No apartamento deles, ela abrirá a mala: presentes, vitaminas, remédio para as pernas inquietas.
Her mother will kiss her. “We have these things in China now.” In two weeks, they’ll repack her suitcase with the life she could’ve had: water spinach seeds, photographs, native honeysuckle cuttings for flowers that will make her weep.	A mãe lhe dará um beijo. — Já temos essas coisas na China agora. Em duas semanas, eles farão de novo as malas dela, com a vida que ela poderia ter tido: sementes de espinafre-d’água, fotografias, mudas de madressilva nativa — flores que a farão chorar.
In five more years, the honeysuckle will cross her fence. Its children will grow toward the highway, and then the ocean, and then, and then...	Em mais cinco anos, a madressilva cruzará a cerca. Seus filhos crescerão em direção à rodovia, depois ao oceano, e depois, e depois...

Source: The authors.

The ChatGPT’s translation stands out as straightforward and usable, requiring no major post-editing intervention, except for debatable issues of taste and style that could easily be adjusted with a new prompt. Among few problematic issues, we noticed: (1) the use of em dashes, which had not been requested; (2) the lack of specification of who has restless legs, an occurrence also noted in DeepL; and (3) the translation of “*Seus filhos*” as “Its children,” which is not an error at all, just a different choice when compared to the human’s output.

⁶ ChatGPT 4.1, paid version.

⁷ It has been noted that chatbots respond better when asked politely (see Wright, 2024), despite controversial allegations that “Saying ‘please’ and ‘thank you’ to ChatGPT costs OpenAI millions” (Carroll, 2025; see also Butts, 2026).

Noticeably, the first paragraph is almost identical to the human translation, with the exception that, like the machine translation engines, ChatGPT also rendered the adverb “já” [already]. The remainder of the translation is also similar to the human translation, but we highlight the lack of creativity in the translation of “then” as “depois.” Overall, however, no significant errors could be found that would require some tweak or rewriting, except for the insertion of the em dash after “*madressilva nativa*,” which is not only absent in the original but is neither necessary nor welcome in current literary translation practices in Brazil. We discuss the use of en and em dashes in ChatGPT outputs in the next section.

Subsequently, we applied the same prompt in Brazilian Portuguese to Gemini. It responded in English, except for the word “*microconto*” (microfiction): “Sure, I can help you with that! Please provide the *microconto* you would like translated. I’m ready to translate it into contemporary Brazilian Portuguese, maintaining its style and content.” Table 5 provides the raw output obtained.

Table 5 – Gemini’s raw output

	O Retorno
Soon, she’ll land. Her father will be shorter. Her mother’s bones, lighter. “Aiyaa, it’s been five years.” It’ll need to be said.	Logo ela vai pousar. O pai estará mais baixo. Os ossos da mãe, mais leves. “Aiyaa, já se passaram cinco anos.” Isso precisará ser dito.
At their apartment, she’ll unpack her suitcase: gifts, vitamins, medicine for their restless legs.	No apartamento deles, ela vai desempacotar a mala: presentes, vitaminas, remédios para as pernas inquietas deles.
Her mother will kiss her. “We have these things in China now.” In two weeks, they’ll repack her suitcase with the life she could’ve had: water spinach seeds, photographs, native honeysuckle cuttings for flowers that will make her weep.	A mãe vai beijá-la. “Temos essas coisas na China agora.” Em duas semanas, eles vão refazer a mala dela com a vida que ela poderia ter tido: sementes de agrião-d’água, fotografias, mudas de madressilva nativa para flores que a farão chorar.
In five more years, the honeysuckle will cross her fence. Its children will grow toward the highway, and then the ocean, and then, and then...	Em mais cinco anos, a madressilva vai atravessar a cerca dela. Seus filhos crescerão em direção à rodovia, e depois ao oceano, e depois...

Source: The authors.

Here we witness an interesting issue that had not been raised previously: the hallucination of the title, which was precisely what made us realize that we had not included the title of the story as an input. Gemini “created” a title that was not included, a trend of “creative” additions (namely, hallucinations according to some, but creativity according to others, probably naïve users) by AI that is already widespread and feared. We address this in the discussion section.

The output lacks standardization of the verb form: sometimes future is realized through the auxiliary verb “ir” (“vai” [go]) in present tense + main verb in infinitive, sometimes future is realized through a morpheme attached to the very end of the verb. Like all other tools, Gemini rendered the adverb “já” [already] – which intensifies the idea of long time or enhances the idea that the event is fully complete – and the unnecessary, albeit grammatically correct, deictic demonstrative pronoun “Isso” [This/That]. However, unlike the machine

translation engines, Gemini provided unproblematic renditions for the deictic possessive pronouns, except for the deictic ambiguity in “*Seus filhos*” [Their/His/Her/Its children].

An interesting deviation in Gemini’s translation as compared to the others is the use of “*agrião d’água*” as a rendition for “water spinach,” which is not incorrect from a botanical perspective. Overall, however, the translation produced by Gemini is readable and requires relatively less post-editing – even though verbs would have to be adjusted – than the outputs from Google Translate and DeepL, but significantly more than ChatGPT. Besides, no creativity was found as in the more oral-like human rendition “*daí*” for the neutral temporal deixis “then.”

In the next section, we discuss these findings.

Discussion

When analyzing the machine outputs, two main questions stood out: (1) how readable and close to human translation the generated texts were, especially that of ChatGPT, and (2) how vocabulary issues considered and reconsidered during human translation were resolved similarly by the machine, but in a matter of seconds. To address these two questions, we first delved into two gold standard metrics applied to test the quality of machine translation outputs, namely Bilingual Evaluation Understudy (BLEU) and Translation Error Rate (TER). The rationale underlying both metrics is that the closer the machine translation output is to that of a professional human translator, the better the engine is. BLEU provides a number between 0 and 1 (which we standardized to a 0–100 scale), with figures tending to 1 (or 100) representing higher similarities to a human translation provided as reference, while TER measures (in a 0 to 100 scale) indicate the number of edits needed to change an engine’s output so that it exactly matches a human-translated reference, with figures tending to 0 representing higher similarities. Table 6, elaborated by Heider Carlos Ribeiro da Cruz, an MA student oriented by the second author at Federal University of Uberlândia (UFU)’s Graduate Program in Linguistics (PPGEL/UFU), provides the test results for all engines’ outputs compared to our human translation as reference.

Table 6 – BLEU and TER results for all machine outputs

Metric \ Machine	Google	DeepL	Gemini	ChatGPT
BLEU	58.90	39.50	49.43	45.48
TER	33.66	60.40	38.61	47.52

Source: Elaborated by MA student Heider Carlos Ribeiro da Cruz.

According to the metrics, Google Translate seemed to perform better than all the others, while DeepL seemed to have the worst performance. Meanwhile, ChatGPT and Gemini performed similarly according to BLEU, although Gemini had a clearly better TER performance. From the post-editor point of view, this means that Google Translate would require the fewest number of edits and DeepL would require the highest number of edits, with ChatGPT and Gemini falling somewhere in between.

Nevertheless, caution is needed when using such metrics, especially for literary texts. As we aimed at authorship, multiple imports and creativity, we could not expect that the machines would perform well in the tests. Besides, neither of them account for intelligibility nor grammatical correctness. That is the strength of our exploratory, qualitative analysis reported in the previous section and further elaborated on in this section.

Several aspects are to be considered for quality in terms of both intelligibility and grammatical correctness of the machines' outputs. In the specific case reported here, the leap in quality is palpable between online machine translation engines (namely, DeepL and Google Translate) and chatbots (specifically, ChatGPT). Chatbots seem to translate significantly better than other tools – a conclusion also drawn by other studies, such as Bistafa (2024) and You *et al.* (2025), who compared AI chatbot outputs to machine translation engines in the rendering of difficult texts into English. You *et al.* (2025) also reported the poor performance of DeepL.

One relevant aspect is the chatbots' improved handling of Brazilian Portuguese deictic possessive adjectives and pronouns when compared to the machine translation engines. Both ChatGPT and Gemini seemed to “understand” (i.e., probabilistically capture, unveil) that specifying or clarifying deictic possessive pronouns or adjectives is important for the reader's comprehension and enjoyment of the text in Brazilian Portuguese.

Assuming that fluency is a desirable feature of the target text, the machines' outputs could be assessed as to whether they could potentially lead to self-effacement – that is, if an expert human speaker of the target language would fail to determine whether the raw outputs are “human-like.” Based on this criterion, we would argue that the four machine-generated translations are close enough to achieving a “human-like” status if we assume that readers both have no access to the source text and proactively engage in finding coherence in the target text. Still, easiness in post-editing all raw outputs could be placed in a continuum in the following order, from easiest to most difficult: ChatGPT, Gemini, Google Translate, and DeepL. This order is consistent with both our qualitative analysis in the previous section and our quantitative analysis of Table 1 in this section. It is also consistent with You *et al.* (2025), who reported DeepL's poor performance and the closeness of some AI raw outputs to human translators' renditions.

However, none of the raw outputs of our self-applied experiment would be considered a polished, finished text, that is, one that is undeniably ready for immediate publication. Therefore, none of them would appear to us to have been written by a professional literary human translator, but they could be easily mistaken as renditions of translation students in their early courses. The only exception seems to be DeepL's raw output, because of its trace of European Portuguese and use of Latin Aspas. In fact, for a short text like the one used here, post-editing DeepL's raw output would require more effort and generate more doubt and insecurity (e.g., is this vocabulary correct, or should we verify each choice?) than human translation from scratch, thus requiring more post-editing effort both cognitively and technically. At the other end of the spectrum, ChatGPT (literally, Chat Generative Pre-Trained Transformer) seems to be a potentially useful tool, confirming what has been demonstrated in other studies (Bistafa, 2024; Alkodimi; Alqahtani; Al-Wasy, 2024; Spoturno, 2024).

Interestingly, one could argue that the long, time-consuming, thoughtful and carefully deliberate strategies of human translators are easily defeated by seconds of machine data processing that arrive at the exact same solutions. Words like “*madressilva*” for “honeysuckle” and “*espinafre d'água*” for “water spinach” were researched and generated uncertainty in the human translators but were quickly resolved by the machines, with the obvious dis-

claimer that we had the solutions confirmed before using the machines, i.e., our reasoning would be quite different, had we used the machines first. It is possible that this would not have been the case with a longer text, but the human translators' frustration was evident: our immediate reaction was to laugh nervously and question our role as experts (something that, again, is debatable if we consider that we assessed the machine outputs only after confirming vocabulary and terminology thoroughly). Evidently, we do reaffirm the importance of investigating terminology, of making conscious, deliberate decisions and of knowing how to explain them if necessary (see Costa; Silva, 2020; Silva, 2021; Silva; Alves, 2020) as an ethical attitude in translation. In the case discussed here, however, the fact is that all the research made no difference to the result if we take a solely utilitarian, pragmatic, productivity perspective that naïvely embraces the machine and accepts its output as valid without any need for factual and terminological validation; in fact, the difficulties of the process cannot be perceived through a mere comparison of the products. This by no means should be interpreted as our Surrender to the machine (Collins, 2018):

As it is, the big danger facing us is [...] failing to notice computers' deficiencies when it comes to appreciating social context and treating all consequent mistakes as our fault. Thus, much worse, and much more pressing, than the danger of being enslaved by enormously intelligent computers, is our allowing ourselves to become the slaves of stupid computers – computers that we take to have resolved the difficult problems but that, in reality, haven't resolved them at all.

[...]

The Surrender will mean destroying the most distinctive things about what it is to be human – natural language use with all the flexibility and context-sensitivity that makes the use of language so rich.

As we have discussed elsewhere (Costa; Silva, 2020), we still believe that literary translation is a human activity *par excellence*. It is not about merely resolving “the difficult problems,” but rather testing and defying the target natural language to its limit in an endeavor of guided production of meaning (Halliday, 1992), so that an instantiation of such a language system shows unforeseen, unpredictable richness in import-oriented lexicogrammatical choices, which may or may not allow influences of the source language system to shine through in the target language.

Another issue that frustrated us as human translators was the use of the Chinese interjection “*Aiyaa*” in all raw outputs. No machine provided a Brazilian Portuguese interjection for the word nor hallucinated or provided any other type of solution (at least in formatting, for instance). When we chose the microfiction, we had anticipated that the presence of a lexical item from a culture that is foreign to both cultures in question (Anglophone and Lusophone, more specifically the United States and Brazil) would potentially generate different outputs, leading us to a discussion about the machine's potential inability to handle cultural markers. Handling cultural markers is a very common notion when considering the human translator as a mediator: “[Human t]ranslators are increasingly taking on roles as language consultants, cultural mediators, and even data curators. They provide value by addressing nuanced linguistic and cultural aspects that AI systems struggle with, such as idioms, humour, tone, and context-specific references” (Sun; Liu; Moratto, 2025, p. 5). However, running counter to our expectation, our findings seem to show that the machine can handle cultural markers in

some contexts (at least when the input contains cultural markers that are unusual to both source and target languages, which remains to be further explored in future research).

A key issue that emerged in this brief test was the undesirable “interference” of AI in the target text’s lexicogrammatical realizations (and consequently imports) as compared to those in the source text. Chatbots have shown a “creative” tendency, one that is generally a concern surrounding the fight against fake news today (see Metz; Weise, 2025) and that can also cause problems in translations, especially long ones, as the machine adds information that was not there (e.g., as in the case of hallucinations) and tends to generate more text than necessary. While the realization of the adverb “já” [already] may be seen as a preference for formulaic sequences (see Bestgen, 2022) – one that, nonetheless, has meaningful implications, as it intensifies the completion of an action in time, restricting potential implicatures –, ChatGPT and its inclusion of em dashes, and particularly Gemini and its creative title for the story, can be seen as examples of unwanted machine interference. AI’s so-called hallucinations seem to be worsening (Metz; Weise, 2025), and psychology experts have been warning against the dangers of their fabrications: “So, use ChatGPT at your own peril. Just as I would not recommend collaborating with a colleague with pseudologia fantastica, I do not recommend ChatGPT as an aid to scientific writing” (Emsley, 2023, p. 2).

What is most relevant for us, in the domain of Translation Studies, is that interference, perhaps counterintuitively, points not to AI’s invisibility, but rather to a heavy-handed performance that may reveal more of the engagement of the machine than that of the human in the translation process. It seems that the current language industry and its clients are less doubtful about the level of AI use than about the level of human engagement. Ethical, human-oriented, quality-driven concerns have been placed on hold for the sake of costs reduction and increased productivity in a context of tryouts or trial and error until the launch of a more disruptive technology or the emergence of scandals derived from poor AI translation use and output management. Human translators are assumedly expected to use the machine, but the public just does not know how, how much, when, and where it will be used. There may be a few exceptions, such as expert human translators known for their distaste for machines, but even in these cases there is no way to predict or control the human translators’ processes in the isolation of their offices. Nor can we gauge how ethical translators (especially translation agencies) have been in their disclaimers to customers about AI shortcomings and their actual use and management of AI in their activities. Like Giddens (1990) used to say, we can only trust in the expert system and hope for the best while the juggernaut (a metaphor that we can apply to AI) is going down the hill.

Visible “ticks” are already noticeable in AI outputs – additions, in particular. One of these, recurrent in translations from English into Brazilian Portuguese, is the use of em dashes in sentences originally punctuated differently, as shown in our self-applied experiment with ChatGPT. This phenomenon has already been widely noticed by users and frequently generates online debates, such as “Is an em dash (—) proof of AI manipulation?” on Reddit, currently with 66 comments.⁸ One respondent even noted the chatbot’s “disobedience” (incompliance) regarding this use: “It’s strange that even if you tell chatGPT not to use em dash, it keeps using

⁸ IS AN EM dash (—) proof of AI manipulation? [S. I.], Oct. 5, 2024. Reddit: BaggyBoy. Available at: https://www.reddit.com/r/ChatGPT/comments/1fx12q1/is_an_em_dash_proof_of_ai_manipulation/. Accessed on: Aug. 12, 2025.

it.” In a thread among Brazilians on the same website, a user vented: “Every day I tell GPT to never use em dashes, but GPT doesn’t care about what I want.”⁹ In the current 172 comments, fellow users tried to help the post’s owner improve their prompt to avoid this infamous “disobedience.” ChatGPT’s “stubbornness” has also been reported in academic settings (Emsley, 2023).

In our self-applied experiment’s microfiction there was no need for an en or em dash, nor would a professional literary human translator (likely) have retained it in post-editing. Had it been retained, however, it could have been an indicator or a reminder to editors and other informed readers of the (potential) use of AI.

To test ChatGPT’s “obedience” or compliance, we repeated the microfiction translation prompt on another day (July 31st, 2025), including at the end: “sem adição de traços e travessões” (without adding en or em dashes). The raw output was as follows:

Table 7 – ChatGPT’s new raw output

Soon, she’ll land. Her father will be shorter. Her mother’s bones, lighter. “Aiyaa, it’s been five years.” It’ll need to be said.	Em breve, ela vai pousar. Seu pai estará mais baixo. Os ossos de sua mãe, mais leves. “Aiyaa, já se passaram cinco anos.” Isso precisará ser dito.
At their apartment, she’ll unpack her suitcase: gifts, vitamins, medicine for their restless legs.	No apartamento deles, ela vai desempacotar sua mala: presentes, vitaminas, remédio para as pernas inquietas.
Her mother will kiss her. “We have these things in China now.” In two weeks, they’ll repack her suitcase with the life she could’ve had: water spinach seeds, photographs, native honeysuckle cuttings for flowers that will make her weep. In five more years, the honeysuckle will cross her fence. Its children will grow toward the highway, and then the ocean, and then, and then...	Sua mãe vai beijá-la. “Agora temos essas coisas na China.” Em duas semanas, vão repor a mala com a vida que ela poderia ter tido: sementes de espinafre d’água, fotografias, estacas de madressilva nativa para flores que a farão chorar. Em mais cinco anos, a madressilva vai atravessar sua cerca. Seus filhos vão crescer em direção à estrada, depois para o oceano, e depois, e depois...

Source: The authors.

Several points caught our attention when analyzing this new raw output. To begin with, it is quite different from the translation ChatGPT itself had presented a couple of days before. As instructed, no en or em dashes were used (not even in the narrators’ imagined dialogue, as used by ChatGPT in the first translation and which is customary in Brazilian Portuguese narratives); the chatbot did not practice its infamous “disobedience,” which may only have happened by chance, at random. However, the output’s quality dropped significantly (even though the TER and BLEU metrics did not differ much from those found for the previous ChatGPT translation: 48.41 vs. 47.52 and 45.64 vs. 45.48, respectively). Particularly, verb tenses would need to be adjusted in post-editing, as much as in the case of all other raw outputs reported in the previous section; the use of the word “*estaca*” [cuttings] and, later, “*estrada*” [road] may be considered less acceptable in the context than “*mudas*” [seedlings] and “*rodovia*” [highway], respectively, in the previous translation. Although we do not explore the issue in profundity in this study, this test indicates that the quality of ChatGPT’s raw outputs may sometimes look

⁹ TODO dia eu digo para GPT para nunca usar travessões, mas GPT não se importa com o que eu quero. [s. l.], Mar. 11, 2025. Reddit: reliablepayerthead. Available at: https://www.reddit.com/r/ChatGPT/comments/1j8vq4z/every_single_day_i_tell_gpt_to_never_use_em/?t=pt-br. Accessed on: Aug. 12, 2025.

impressive (as was the case in the first round), but it does not guarantee standardization and predictability, since different iterations produced very different raw outputs. Besides, it also indicates that automatic metrics like BLEU and TER may not be sensitive enough to capture relevant nuances in the raw outputs. In fact, such measures entirely disregard the multiplicity of translation choices one has at one's disposal depending on context, personal choices, contract, etc. (something widely discussed within Translation Studies). They have only one or a few human translations used as a standard to compare with (an automatic method that is not only widely accepted as efficient among Natural Language Processing scholars and practitioners, but is also appealing and understandable to society in general).

Considering Venuti's (1995) approach to domestication as a tendency toward text uniformity, human translators who use AI risk becoming domesticators even when that is not their deliberate wish, since post-editing purportedly requires standardization of verb tenses, elimination of unnecessary additions, and a general "cleanup" of the text to make it more uniform while trying not to surrender to an allegedly unnecessary or undesirable eagerness to rephrase sentences completely. In this operation, the roles are reversed: the AI would be creative, disobedient, "rebel," while the human translator would act as a mere diligent proofreader.

Updating Venuti's (1995, p. 17) words to our current conundrum, if what the literary human translator performs in post-editing is "rewriting them [texts] in the transparent discourse that prevails" in a given language, their control over the translation project itself is compromised. Even if the prompt is managed specifically, it is clear that AI does not always comply and is sometimes quite insistent in its lexicogrammatical choices, as they do need to conform to probabilistic combinations.

In literary translation, therefore, the challenge is to find ways to use AI without being used by it; to maintain the human translator's creativity – and a domesticating or foreignizing project, if they so desire – without falling into the persistent traps of the pre-programmed machine. Spoturno (2024) has shown that teaching AI management in literary translation is becoming increasingly unavoidable in the educational environment. As we have contended elsewhere (Costa; Silva, 2020), using AI in literary translation is no longer a matter of feasibility, but rather a matter of ethical choice. AI hallucinations and unusual lexicogrammatical choices can only be considered creative if they first pass a translator's judgement and validation process; otherwise, they are nothing but random fancy errors that happened to shine through as creative choices for uninformed monolingual readers. That is why – as impressed as we were with the quickly obtainable raw outputs in this self-applied experiment, especially against the meticulous, time-consuming work we performed during human translation – we would still hesitate to use the machine for texts like this. We can justify our creativity, but not the machine's. In fact, even though we used the phrase "lexicogrammatical choices" for the raw machine outputs, what we have is only a probabilistically based rendition. The human translators are still the only ones holding the power to make linguistic choices, as they are still the ones who really can interpret the source text and account for the import produced in the target text against all other meaning potentials available in the target language that they opted not to instantiate. In other words, accountability may well not be an issue for the machine, but it fortunately is for humans.

Final Remarks

As outlined in our Introduction, we designed a self-applicable experiment to tap into undesirable interferences produced by machine translation engines and AI chatbots in their raw outputs as compared to the meanings of the source literary text and the lexicogrammatical choices we had made in our own human translation of such text. We qualitatively analyzed process and product aspects (particularly, literary text import, adequacy, fluency, and post-editing effort) of all raw outputs and compared them to those involved in our own human translation of the same input. Running counter to our expectations, the number of undesirable interferences was not as striking as we had anticipated.

Following this simple step-by-step process provoked in us some not-so-easy reflections on our craft. The first shock was realizing how similar the translations were to each other and to our own (even though that was not captured through the automatic metrics BLEU and TER). However, looking at them closer, we concluded that the majority would still need much human intervention to become satisfactorily publishable, except for ChatGPT's first raw output, which would still invite adjustments, but fewer and arguably less important for the literary text import. Still, ChatGPT can fabricate words and punctuation that mismatch the original wording, which adds an extra need for attention to the human translator's post-editing process.

Literary human translators are not easily replaceable (nor could a simple, ungeneralizable self-applied experiment such as the one performed here ever be indicative of such a drastic outcome). Experiments like the one performed here simply point to the reasonable fact that AI may be a gain in the translation field, one in which the role of the human translator as an expert is more relevant than ever. As recently noted by Sun, Liu, and Moratto (2025), new specializations have emerged or are yet to emerge on the interface between human and AI translation processes: "Another avenue for adaptation is the expansion of the human translator's role to include that of a cultural mediator and communication specialist" (Sun; Liu; Moratto, 2025, p. 12). In fact, looking at the contemporary AI-related language phenomena requires a manifold approach, including, among others, considerations regarding adaptive expertise (Angelone, 2022) and distributed cognition (Alves; Jakobsen, 2021) to account for the translator's agency in the process, as well as linguistic, stylistic and literary analyses of raw outputs and post-edited outputs. Even though we did not explicitly state that we included all such considerations, they are hinted at throughout this article. In fact, we endeavored to be as humble and sincere as possible to acknowledge our limitations and our attempts at adapting to a new era. Translation trainers should follow suit to have a better understanding of both the impact of AI on our craft and our agency in using AI in our work, as well as challenge society (especially associations and government agencies) to demand and provide a clear notion of accountability for the translations circulating in this new era. As a matter of fact, who is to be held accountable, especially in the event of disasters, biases, unintelligible texts, fake pieces of information etc.?

Besides prompt creation, professional human translators will be required not only to evaluate the machine's output but to teach it – to society in general and want-to-be translators in particular – in an increasingly more critical and skilled manner (both linguistically and literarily, in the case of complex fiction). Consequently, a shift from product to process may be taking place, since human translators will show their expertise in what happens before – choice of engine or chatbot, prompt creation – and adjustments during the translation process, more

than in the product itself. However, as it has always been the case in the language industry, both professional translators and translation trainers have never engaged proactively in the design of translation technologies or anticipated the impact of Natural Language Processing evolutions on their works. As we believe that this will remain in the near to the far future, we should rely on the fact that adaptive expertise – alongside language/cultural proficiency and metalanguage (Costa; Silva, 2023) – has always been our best card to play in moments of inflection. Adapting to this ever-changing scenario does not entail Surrender (Collins, 2018) but rather reassures that human beings are in control of the machines and translators are the best gatekeepers available when it comes to validating or defying the machine outputs. All in all, natural languages are human in nature, but in an industry resorting to inexpensive, high-quality AI raw outputs, only expert translators will be the last men standing as relevant. This, in turn, calls for improved training and deliberate practice (Ericsson, 2001) aimed at language/cultural proficiency, metalanguage, and adaptive expertise.

The risk, as we found in our discussion, is turning the task of the human translator – prompting, post-editing, revising – invisible once again. This leads us, finally, to the unsurmountable issues regarding the relationship between AI and translation: authorship, agency, and accountability, which deserve further consideration in future studies. In human translation, as demonstrated by the annotations of the human translators reported in our self-applied experiment, choices can be reasoned, deliberated and, if necessary, duly justified. Even if a machine arrives at the same result, the “reasoning” behind it will remain unclear, or rather, the mere unfolding of probabilistic automated decisions. It is up to us as translators and translation scholars to make society aware of this, so that AI and machine translation consumers start challenging the machine juggernaut instead of blindly believing in artificial intelligence (to use Collins’s words) and in the for-profit language industry players’ artificial stories. It is also our responsibility to train and educate want-to-be translators to deal with AI and the ever-changing translation market, as wisely advised by Esqueda and Silva (in press). We – Translation Studies scholars and professional translators – need to reclaim our status as gatekeepers in the language industry instead of surrendering to the invisible forces of AI and the heavy hand of computer scientists only concerned with vectorial distances that provide acceptable, grammatical, fluent combinations based on massive language models, without any accountability for text content, text style, authorship, copyrights, cultural references, and even validity and unbiasedness of arguments.

Acknowledgements

The second author is thankful for the productivity scholarship provided by the National Council for Scientific and Technological Development (CNPq, Process No. 316193/2023-2).

References

- ALKODIMI, Khaled Abkar; ALQAHTANI, Osama Abdulrhman; AL-WASY, Baleigh Qassim. Human-AI Collaboration in Translation and Back Translation of Literary Texts. *Journal of Social Studies*, v. 30, n. 2, p. 173-192, July 2024.
- ALVES, Fabio; JAKOBSEN, Arnt Lykke. Grounding Cognitive Translation Studies: Goals, Commitments and Challenges. In: ALVES, Fabio; JAKOBSEN, Arnt Lykke (org.). *The Routledge Handbook of Translation and Cognition*. New York: Routledge, 2021. p. 545-554.
- ALVES, Fabio; SILVA, Igor Antônio Lourenço da. Bridging Paradigms to Approach Expertise in Cognitive Translation Studies. In: MUÑOZ, Ricardo; SUN, Sanjun; LI, Defeng (org.). *Advances in Cognitive Translation Studies*. Singapore: Springer, 2021. v. 1, p. 1-20.
- ALVES, Fabio; SILVA, Igor Antônio Lourenço da. Expertise in Interpreting as an Interlingual Reformulation Skill. In: MELLINGER, Christopher (org.). *The Routledge Handbook of Interpreting and Cognition*. New York: Routledge, 2024. v. 1, p. 32-51.
- ANGELONE, Erik. Weaving Adaptive Expertise into Translator Training. In: MASSEY, Gary; HUERTAS-BARROS, Elsa; KATAN, David (org.). *The Human Translator in the 2020s*. New York: Routledge, 2022. E-book.
- BARBASSA, Juliana. 'Building Something Together': Translators Discuss Their Art. *The New York Times*, July 2, 2023. Available at: <https://www.nytimes.com/2023/06/28/books/review/literary-translator-roundtable-discussion.html>. Accessed on: Sept. 19, 2025.
- BESTGEN, Yves. Comparing Formulaic Language in Human and Machine Translation: Insight from a Parliamentary Corpus. In: PARLACLARIN, 3.; LANGUAGE RESOURCES AND EVALUATION CONFERENCE (LREC), 13., 2022, Marseille. *Proceedings [...]*. Marseille: European Language Resources Association (ELRA), 2022. p. 101-106. Available at: <https://aclanthology.org/2022.parlaclarin-1.14.pdf>. Accessed on: July 31, 2025.
- BISTAFA, Sylvio R. Translating Scientific Latin Texts with Artificial Intelligence: The Works of Euler and Contemporaries. *Euleriana*, v. 4, n. 2, p. 149-171, 2024.
- BOHM, David. *On Creativity*. London: Routledge, 1998.
- BRITTO, Paulo Henriques. *A tradução literária*. Rio de Janeiro: Civilização Brasileira, 2012.
- BUTTS, Dylan. Sam Altman Defends AI Resource Usage: Water Concerns 'Fake,' and 'Humans Use Energy Too'. *CNBC*, Feb. 23, 2026. Available at: <https://www.cnbc.com/2026/02/23/openai-altman-defends-ai-resource-usage-water-concerns-fake-humans-use-energy-summit.html>. Accessed on: Mar. 5, 2026.
- CALVO-FERRER, José Ramón. Can You Tell the Difference? A Study of Human vs Machine-Translated Subtitles. *Perspectives*, v. 32, n. 6, p. 1115-1132, 2023.
- CARROLL, Shannon. Saying 'Please' and 'Thank You' to ChatGPT Costs OpenAI Millions, Sam Altman Says. *Quartz*, Nov. 3, 2025. Available at: <https://qz.com/open-ai-sam-altman-chatgpt-gpt4-please-thank-you-1851777047>. Accessed on: Mar. 5, 2026.
- COLLINS, Harry. *Artificial Intelligence: Against Humanity's Surrender to Computers*. Cambridge: Polity, 2018.

- COSTA, Cynthia Beatrice; SILVA, Igor Antônio Lourenço da. Annotating Translation as a Meaningful Metacognition-Oriented Practice in the Translation Classroom. *Revista de Letras*, v. 1, p. 1-16, 2023.
- COSTA, Cynthia Beatrice; SILVA, Igor Antônio Lourenço da. On the Translation of Literature as a Human Activity par Excellence: Ethical Implications for Literary Machine Translation. *Aletria: Revista de Estudos de Literatura*, v. 30, p. 225-248, 2020.
- DOMINGOS, Livia Cremones. *O desafio da tradução de minicontos da coletânea The World's Shortest Stories: uma tradução comentada com foco na metarreflexão e no processo tradutório*. 2016. 73 f. Senior Thesis (Bachelor's Degree in Translation) – Universidade Federal de Uberlândia, Uberlândia, 2016.
- EMSLEY, Robin. ChatGPT: These Are Not Hallucinations – They're Fabrications and Falsifications. *Schizophrenia*, v. 9, n. 52, p. 1-2, 2023.
- ERICSSON, K. Anders. Expertise in Interpreting: An Expert-Performance Perspective. *Interpreting*, v. 5, n. 2, p. 187-220, 2001. DOI: <https://doi.org/10.1075/intp.5.2.08eri>.
- ESQUEDA, Marileide Dias; SILVA, Igor Antônio Lourenço da. An Autoethnonarrative of Charting Virtual Worlds: A Pedagogical Case Study of Training in Game Translation and Localization at a Brazilian University. *Mutatis Mutandi*, [2026]. In press.
- FISCHER, Lukas; LÄUBLI, Samuel. What's the Difference Between Professional Human and Machine Translation? A Blind Multi-Language Study on Domain-Specific MT. In: ANNUAL CONFERENCE OF THE EUROPEAN ASSOCIATION FOR MACHINE TRANSLATION, 22., 2020, Lisboa. *Proceedings* [...]. Lisboa: European Association for Machine Translation, 2020. p. 215-224.
- GIDDENS, Anthony. *The Consequences of Modernity*. Stanford: Stanford University Press, 1990.
- GLAVEANU, Vlad *et al.* Creativity as Action: Findings from Five Creative Domains. *Frontiers in Psychology*, Lausanne, v. 4, p. 1-14, 2013.
- HADLEY, James *et al.* (ed.). *The Qualities of Literary Machine Translation*. Dublin: Machine Translation Summit, 2019.
- HALLIDAY, Michael Alexander Kirkwood. Language Theory and Translation Practice. *Revista Internazionale di Tecnica della Traduzione*, n. 0, p. 15-25, 1992.
- HOYTE-WEST, Antony. On the Visibility of Literary Translators Working from Smaller Languages into English: Some Observations from the United Kingdom. *Analele Universității „Ovidius” Constanța*, v. XXXIV, n. 2, p. 235-250, 2023.
- HUMBLÉ, Philippe. Machine Translation and Poetry: The Case of English and Portuguese. *Ilha do Desterro*, Florianópolis, v. 72, n. 2, p. 41-56, 2019.
- KINNEY, Ophelia Hu. Honeysuckle. *NYC Midnight – 100-word Microfiction Challenge*, 2020. Available at: <https://www.nycmidnight.com/stories/honeysuckle>. Accessed on: July 31, 2025.
- KOGLIN, Arlene; MOURA, Willian Henrique Cândido; MATOS, Morgana Aparecida de; SILVEIRA, João Gabriel Pereira da. Quality Assessment of Machine-Translated Post-Edited Subtitles: An Analysis of Brazilian Translators' Perceptions. *Linguistica Antverpiensia, New Series – Themes in Translation Studies*, Antwerp, v. 22, p. 41-60, 2023.
- LIU, Ming; ZHANG, He; WU, Guanhao. Fine-Grained Human Evaluation for English-to-Chinese Machine Translation: a Case Study on Scientific Text. *ArXiv*, Sept. 13, 2021. Available at: <https://arxiv.org/pdf/2110.14766>. Accessed on: Aug. 7, 2025.

METZ, Cade; WEISE, Karen. A. I. Is Getting More Powerful, but Its Hallucinations Are Getting Worse. *The New York Times*, May 5, 2025. Available at: <https://www.nytimes.com/2025/05/05/technology/ai-hallucinations-chatgpt>. Accessed on: Aug. 10, 2025.

MOORKENS, Joss; O'BRIEN, Sharon; SILVA, Igor Antônio Lourenço da; FONSECA, Norma Barbosa de Lima; ALVES, Fabio. Correlations of Perceived Post-Editing Effort with Measurements of Actual Effort. *Machine Translation*, v. 29, p. 1-18, 2015.

PLITT, Mirko; MASSELOT, François. A Productivity Test of Statistical Machine Translation Post-Editing in a Typical Localization Context. *The Prague Bulletin of Mathematical Linguistics*, Neuchâtel, n. 93, p. 7-16, 2010. Available at: <https://ufal.mff.cuni.cz/pbml/93/art-plitt-masselot.pdf>. Accessed on: Aug. 10, 2025.

PYM, Anthony. Redefining Translation Competence in an Electronic Age. In *Defence of a Minimalist Approach*. *Meta: Translators' Journal*, Montréal, v. 48, n. 4, p. 481-497, 2003.

SALIMI, Jonni. *Machine Translation of Fictional and Non-Fictional Texts: an Examination of Google Translate's Accuracy on Translation of Fictional Versus Non-Fictional Texts*. 2014. 22 f. Thesis (Bachelor's Degree in English Linguistics) – Department of English, Stockholms Universitet, Stockholm, 2014.

SILVA, Igor Antônio Lourenço da. Translation, Expert Performance and Cognition. In: ALVES, Fabio; JAKOBSEN, Arnt Lykke (org.). *The Routledge Handbook of Translation and Cognition*. New York: Routledge, 2021. p. 1-17.

SILVA, Igor Antônio Lourenço da; ALVES, Fabio. Desenvolvendo a subcompetência estratégica: convergência entre os elementos da competência tradutória. In: ESQUEDA, Marileide Dias (org.). *Ensino de tradução: proposições didáticas à luz da competência tradutória*. Uberlândia: EDUFU, 2020. p. 126-152.

SILVA, Igor Antônio Lourenço da; BORGES, Isabel; MAGALHÃES, Janayna Rocha; MORAIS, Cecília Franco. Formação emancipatória de tradutores: uma autoetnografia. In: MALTA, Gleiton (org.). *Ensino de tradução e de línguas: reflexões e propostas didáticas*. Campinas: Pontes Editores, 2023. v. 1, p. 103-130.

SILVA, Igor Antônio Lourenço da; COSTA, Cynthia Beatrice. A tradução automática de textos literários e o comportamento do tradutor em formação: reflexões para o ensino da tradução. In: PEREIRA, Germana Henriques; COSTA, Patrícia Rodrigues; SILVA, Rodrigo D'Ávila B. (org.). *Formação de tradutores: desafios da sala de aula*. Campinas: Pontes Editores, 2020. v. 9, p. 143-166.

SILVA, Igor Antônio Lourenço da; SCHMALTZ, Marcia; ALVES, Fabio; PAGANO, Adriana; WONG, Derek; CHAO, Lidia; LEAL, Ana Luísa V.; QUARESMA, Paulo; GARCIA, Caio. Translating and Post-Editing in the Chinese – Portuguese Language Pair. *Translation Spaces*, v. 4, p. 145-169, 2015.

SPERBER, Dan; WILSON, Deirdre. *Relevance: Communication and Cognition*. 2nd ed. Oxford: Wiley-Blackwell, 1995.

SPOTURNO, María Laura. Traducción literaria e inteligencia artificial: consideraciones para la formación universitaria. *Cadernos de Tradução*, Florianópolis, v. 44, n. 1, p. 1-26, 2024.

SUN, Sanjun; LIU, Kanglong; MORATTO, Riccardo. Introduction: Navigating the Paradigm Shift-Translation Studies in the Age of AI. In: SUN, Sanjun; LIU, Kanglong; MORATTO, Riccardo (ed.). *Translation Studies in the Age of Artificial Intelligence*. London: Routledge, 2025. p. 1-17.

TORAL, Antonio; WIELING, Martijn; WAY, Andy. Post-Editing Effort of a Novel with Statistical and Neural Machine Translation. *Frontiers in Digital Humanities*, Lausanne, v. 5, p. 1-9, 2018.

VENUTI, Lawrence. *The Scandals of Translation: Towards an Ethics of Difference*. London: Routledge, 1998.

VENUTI, Lawrence. *The Translator's Invisibility: a History of Translation*. London: Routledge, 1995.

VISBY, Morten. The Future Relationship of Literary Translation and AI: Reflections from CEATL President. *Counterpoint*, n. 8, p. 33-36, 2020. Available at: https://ceatl.eu/wp-content/uploads/2023/04/Counterpoint_2020_04_article_08.pdf. Accessed on: July 31, 2025.

WRIGHT, Webb. Please Be Polite to ChatGPT. *Scientific American*, July 25, 2024. Available at: <https://www.scientificamerican.com/article/should-you-be-nice-to-ai-chatbots-such-as-chatgpt/>. Accessed on: July 31, 2025.

YOU, Mu *et al.* How Well Can State-of-the-Art Machine Translation Systems Render a 16th-Century Chinese Novel? *Cadernos de Tradução*, v. 45, n. 3, Special Issue, p. 1-24, 2025. Available at: <https://www.scielo.br/j/ct/a/yqp6PPtcD5pKXGYgMWhjRWk/?lang=en>. Accessed on: Aug. 20, 2025.