KOMPARATĪVISTIKAS
ALMANAHS

JOURNAL OF
COMPARATIVE STUDIES

EDITORIAL BOARD

DAUGAVPILS UNIVERSITY

INSTITUTE OF HUMANITIES
AND SOCIAL SCIENCES

JOURNAL OF
COMPARATIVE STUDIES

CONTENTS

ON PEDAGOGIC USES OF LITERARY MACHINE TRANSLATION: A CASE STUDY BASED ON THE LANGUAGE PAIR ENGLISH – ITALIAN

PAOLA BRUSASCO

Paola Brusasco, Ph.D., associate professor
Department of Languages, Literatures and Modern Cultures
"G. d'Annunzio" University of Chieti-Pescara, Italy
e-mail: paola.brusasco@unich.it
https://orcid.org/0000-0003-3128-0685

Ph.D. Paola Brusasco is an associate professor in English Language and Translation in the Department of Modern Languages, Literatures and Cultures of the University of Chieti-Pescara. Her research interests and publications are in the areas of Translation Studies, ELT, and Postcolonial Studies. Her works include the monographs "Approaching Translation. Theoretical and Practical Issues" (2013; 2016), "Writing Within, Without, About Sri Lanka: Discourses of Cartography, History and Translation in Selected Works by Michael Ondaatje and Carl Muller" (2010), and various articles, while her recent contributions in edited volumes are "Pragmatic and Cognitive Elements in Literary Machine Translation" (2022), and "Edoardo Bizzarri and 'Novellieri inglesi e americani'" (2023). She has also translated many contemporary and classic authors such as Russell Banks, Colson Whitehead, Emily Brontë, and Robert L. Stevenson, and is one of the editors of the online journal "RiTra – Rivista di Traduzione".

ABSTRACT

Translation technologies and Neural Machine Translation (NMT) have not only changed the work of translators and the skills required for the profession, but also prompted the possibility of extending MT to the traditionally human precinct of literary translation (Toral and Way 2015; Toral and Way 2018; Kuzman et al. 2019; Hadley et al. 2022, among others). While for pragmatic texts the quality of NMT output is often adequate or requiring minor editing, the greater complexity, cultural specificity, and creativity of literary texts still require a relevant degree of human intervention. However, since recourse to NMT is growing, it is crucial to retain human centrality by teaching prospective translators not only postediting strategies, but also deeper reading and interpretive skills that will allow them to detect fluent yet wrong, incoherent or stylistically poor renderings of the source text. The current paper starts from a comparison of the existing Italian translations (published by Treves 1933 and Mondadori 1965) of the first chapter from Sinclair Lewis' novel "Ann Vickers" (1933) with the same chapter translated using the online software DeepL. While the existing translations do not read too dated but inevitably contain expressions that would hardly be used today, and some mistakes, the machine-translated excerpt – despite some inadequate lexical choices and predictable shortcomings at pragmatic level – shows a remarkably fluent use of contemporary language, and a reduction of culture-specific errors that back in the 1930s and 1960s presumably derived from a limited knowledge of the Other. The aligned texts and output are analysed from a pedagogic perspective in order to 1) identify their strengths and weaknesses; 2) consider the possibility of using NMT as an aid in a retranslation/revising process that would keep most of the existing translations and only replace outdated or wrong parts; 3) promote a keener sensitivity to meaning and language use in order to avoid flattening linguistic complexity; 4) develop activities aimed at enhancing the students' ability to read and recreate a text on the basis of their physical presence in the world, time and space, i.e. activate their situated cognition when producing or post-editing translations.

Keywords: literary translation, NMT, post-editing, translation pedagogy, situated cognition

INTRODUCTION

Since 2016, Neural Machine Translation (NMT) has been increasingly adopted for commercial and institutional purposes as its performance has proved far superior to previous paradigms, i.e. rule-based and statistical MT. Moreover, deep learning, AI, and massive quantities of digital data allow translation platforms to constantly improve their outputs, providing very accurate and fluent translations in many language pairs. Although creative texts are a smaller domain, especially in comparison to the documents generated for companies and institutions, the use of NMT for language pairs including English has been investigated in that domain too, probably because of the challenges involved and the attempts to bring natural language processing (NLP) closer to human uses of language. However, after initial enthusiasm, scholars are now inclined – at least for literary and creative texts – to retain human centrality out of considerations regarding both translation quality and ethical concerns.

This article goes in that direction by envisaging MT as a support that may help the translator to refresh a literary translation that has aged; it also reflects on a didactic approach that enhances literary translation skills and at the same time sensitises students to a critical use of MT. The first paragraph synthesises previous studies on machine translation of literary texts; the second analyses the source text – American author Sinclair Lewis' novel "Ann Vickers" (1933) –, its two existing Italian translations (1933 and 1965), and DeepL's machine-translated version, with some suggestions for post-editing. The last paragraph focuses on ways to use MT for pedagogic purposes.

LITERARY TEXTS AND MACHINE TRANSLATION

Creativity and the "language writ large" (Tymoczko 2014) of literature deploy a full range of deviations from standard language, relying on ambiguity and metaphor embedded in narratives that portray everyday life as well as invented worlds. The assumption that the human mind will be able to co-construct meaning and visualise imaginary scenarios makes the literary text unforeseeable. As MT relies on algorithms determining the most frequent associations of words and the expectation of matches between

items or strings in different languages, its use for creative texts seems impossible. Several studies have shown interesting results, although at present a substantial amount of human post-editing is necessary. Experiments have been carried out with both dedicated software, trained on domain-specific data, and generalist platforms like Google or DeepL, which have access to any kind of out-of-domain data. For example, Kuzman et al. (2019) evaluated excerpts of novels in English translated into Slovene using both Google and software specially trained on a small parallel literary corpus and found that the former was better – a result they interpreted as a consequence of the small size of the specialised corpus; they also hypothesised better performance if the system were trained on texts by a specific author, which however would be likely to further restrict the size of the corpus and, in the long run, would also reduce expressive possibilities. Toral at al. (2020), instead, came to the opposite conclusion in a more articulate study on texts translated from English into Catalan which involved training two MT systems on domain-specific and out-of-domain monolingual and bilingual corpora; the texts were then evaluated automatically using metrics and translators' annotations, which stated that the system trained on novels had produced more sentences equivalent in quality to human translation. Fonteyne at al. (2020) translated a novel by Agatha Christie into Dutch with Google and categorised fluency and accuracy errors; they found that 44% of the sentences did not contain errors, while the shortcomings they identified had to do with mistranslation, coherence, style and register. Brusasco (2022) analysed a 7-page long excerpt translated from English into Italian using Google, DeepL, and Microsoft, chosen because the huge database available to non-specialised platforms seemed best suited to cover the variety of topics, languages and references in contemporary novels. Assessing the output's usability from the point of view of a literary translator, she found that – despite remarkable fluency and correctness – the amount of post-editing needed and the constraints of working within and between sentences not her own slowed the task and broke the creative flow of form and content typical of the profession. A similar point was raised in other studies (Taivalkoski-Shilov 2018; Kenny and Winters 2020) that pointed out the fragmentation of the process and the loss of "voice" of the translator. Researchers also studied users' perceptions (e.g. Guerberof-Arenas and Toral 2020) by giving readers literary texts in three modalities: a raw output, a post-edited version, and a human-translated one, and investigating their narrative

engagement, enjoyment, and translation reception; MT scored the lowest and human translation the highest, with the post-edited version ranking closer to the latter, but it was also noticed that emotional engagement did not vary greatly, which suggested that, for all its limitations, at communicative level an MT output works almost as well as a text translated or post-edited by humans.

Questions are also being increasingly asked as to the desirability of perfecting MT in order to translate literature out of concern for both jobs for literary translators and the impact MT might have on the language in terms of homogenisation and simplification, which in the long run would affect how humans organise their thoughts and speech. Moreover, the risk is compounded by the fact that the data used for training MT software, and in general AI-powered models, have been proven to contain gender, age, and race bias which in turn affect MT's outputs (Montemayor 2023; Bommasani 2021; Rice et al. 2019, among others).

THE SOURCE TEXT

The source text is the first chapter of Sinclair Lewis's "Ann Vickers" (1933), a novel that follows the tomboyish adolescent protagonist from her juvenile social commitment and work with the suffragettes to motherhood and mature life. The chapter opens with a short, atemporal description of a natural environment and four children in it, almost a snapshot of the scene that is going to unfold – a kind of tableau vivant. Most of the chapter revolves around a game of "let's pretend", with the children impersonating Christopher Columbus and Queen Isabella of Spain before and during the former's expedition to find a new route to Asia. When an unknown boy joins the group, the eager and stubborn protagonist, Ann, is so charmed that she spontaneously gives up her role as Columbus. Through flash forwards, the narrator expands on Ann's independent personality while hinting at future achievements in her life.

The chapter alternates descriptions, lively dialogues, and comments by an omniscient, intrusive narrator. The dialogues reproduce children's talk effectively, with exclamations, repetitions, naïve contentions, and errors at the level of both content and pronunciation. The lexically dense descriptive passages often show an amused gaze towards the characters and the provincialism of Ann's town, as well as a somewhat

disillusioned critique of the smallness of human beings. The main semantic fields are those of playing (e.g. "play", "baseball", "raced", "snowball"), the natural environment (e.g. "river", "willows", "muddy water"), ships (e.g. "barge", "bow", "sails", "captain"), professions (e.g. "social workers", "carpenter", "teacher", "doctors", and the juxtaposition of the sexes ("boyhood", "girl", "male", "woman").

The text is highly cohesive mainly thanks to lexical cohesion (Halliday and Hasan 1976): reiteration happens via many repetitions, hyponyms and superordinates as in "revolver" and "weapon", and near-synonymy, as in "big – large – enormous", or "vessel – ship"; collocation is also present, e.g. part-part as in "shoulders – legs – eyes – hand – nose – skin", and same set, as in "boy – girl – kids – children". The chapter also shows frequent instances of pronominal reference, as in "He took *the revolver* […]. He snapped *it* open" (Lewis 1933, 6; italics mine) and, to a minor extent, ellipsis, especially in dialogues; conjunctions instead are rather sparse.

Cultural references are manifold, both as explicit terms referring to food ("oatmeal and peanut butter"), toponyms (e.g. "Watling's Island", "Main Street", "Fifth Avenue"), and religion ("Sunday school", "Episcopal", "Presbyterian", "Congregational"), and through several other references relying on the reader's background knowledge for their effectiveness (e.g. "Carl Van Doren", "Anglo-Indian proconsul", "[be] Freudian about [something]"). The ironical juxtaposition between the world experienced by the children in their game of "let's pretend" and the abundance of informal dialogues and culture-specific items engages the reader in co-constructing meaning by seeing – then filling – the gaps and recognizing what the cultural references hint at. An ideal translation would take the text towards the reader by giving them sufficient insight into the foreign culture but leave in turn a margin for their interpretive reading – an approach that would position itself midway through the strategies theorised by F. Schleiermacher (1816/2004), i.e. taking the author to the reader as against taking the reader to the author.

THE TRANSLATIONS INTO ITALIAN

The first Italian translation, by Lila Jahn, was published in 1933, the same year as the source text, by Treves, a publishing house founded in Milan in 1861 and active until 1939. The language is

very fluent, but to a contemporary reader several aspects are likely to sound dated and/or overly elevated, such as:

- the translation of the names of the main characters and of the song "Jingle Bells", a common trait of the past which became the rule under the fascist regime (1922–1943), keen on eliminating foreign words from the Italian language. This translational convention was abandoned decades ago;
- the use of subject pronouns *egli, ella, essi* [he, she, and they, respectively], now only found in very formal writing, such as some legal and bureaucratic documents, academic essays, or in literary texts meant to sound old, while usually they are either omitted, because the inflected form of verbs allows the reader to identify the subject, or replaced by their object forms, respectively *lui, lei, loro* [him, her, them], especially in dialogues or in passages characterized by an informal style;
- contracted forms that are no longer in use, except possibly in some dialects, like *diss'egli* [said he], usually rendered as *disse lui*, or even just *disse* with the elision of the pronoun if it is clear from the context, or *ch'essa* [that she], which today would probably be *che lei*.
- unusual lexical choices like *allorché* [when], *sinanche* [even]; *non trovò mai a pentirsi* [she never regretted] instead of *non ebbe mai a pentirsi* or, less formally, *non se ne pentì mai;*
- the elision of the final vowel in verbs, e.g. *raccoglier*, instead of *raccogliere* [to pick up]; *cader* instead of *cadere* [to fall]; *d'aver*, instead of *di avere* [of having], and adjectives such as *simil* for *simile* [similar];
- marked syntax and, consequently, unusual word order, as in *la sua vita svolgendosi* [her life unfolding].

In dialogues, Jahn places the definite article before names, a regional trait that today is mainly limited to either locally connoted texts or comic effects; however, dialogues are overall lively and natural in their reproduction of situated oral exchanges.

The other translation, by Isabella Leonetti, appeared in 1965 for Mondadori, currently the biggest publishing group in Italy. The thirty years separating it from the first translation are perceivable in the language, which is much more similar to current use, although the text shows the same domesticating approach to proper names and culture-specific elements like food and "Jingle Bells". Overall, Leonetti seems to stress gender by choosing more marked lexical

items, e.g. "The three boys" is rendered as *I tre maschi* [the three males], whereas Jahn (1933) uses the more generic *ragazzi* [boys, or boys and girls]. The former is probably motivated by the fact that *ragazzi* is a plural form that can also include girls (much as "children") and the translator is trying to compensate for the gender-specificity of "boyhood", appearing in the same paragraph, which is lost in Italian as no precise equivalent exists, and its possible rendering as *fanciullezza* conveys the meaning of "young age" but lacks any reference to gender. Similarly, Leonetti translates one of the boys' objections to Ann wanting to play the role of Christopher Columbus "And you're only a girl!" (Lewis 1933, 4) as *"E poi sei una donna"* [After all you're a woman], thus emphasizing gender, while a more literal rendering of "girl" as *ragazza* or *bambina* might have suggested that the problem was the protagonist's young age. These lexical choices point towards the translator's loyalty (Nord 2007) to Lewis's focus on both gender roles in early twentieth-century America and the protagonist's eagerness to follow her inclinations and find her place in society.

DeepL's output is on the whole very readable, although there are lexical errors, inconsistencies with verb tenses, an invented word (*catamonti* calqued on "catamounts"), and – predictably – problems with dialogues, affecting both naturalness and the correct rendering of the cohesive device of ellipsis. On the other hand, since it relies on billions of data, the text sounds more modern, names are not translated, as is done today, and food is rendered correctly, an aspect that readers are likely to recognise given the much wider circulation of products and the familiarity with other countries' eating habits. The translation is often fairly literal and at times results in non-fluent or even wrong Italian, but it must be said that there are also cases in which sentences are reorganised effectively. An example of the first type of translation is "Till her day and moment, […]" (Lewis 1933, 8), rendered literally as *Fino al suo giorno e al suo momento*, which might work in a very specific context where the day and moment are clearly identified and refer to a very precise time, but here the sense is "before", "in the past", and it could be rendered as *precedentemente*, as both Leonetti and Jahn did. By contrast, there are instances of effective reorganisation, e.g.:

> There are but frayed cords binding such ambitious, out-stepping American girls as Ann, not only to their native villages, but also to their families, unless they are of recent Jewish or German or Italian origin. (Lewis 1933, 9)

Le ragazze americane ambiziose e intraprendenti come Ann non hanno che corde sfilacciate che le legano non solo ai loro villaggi d'origine, ma persino alle loro famiglie, a meno che non siano di recente origine ebraica, tedesca o italiana.
[Ambitious and dynamic American girls like Ann have but frayed cords binding them not only to their native villages [...]]]

As expected, dialogues are not fluent, a problem that can be traced to several reasons. First, context is crucial to interpret utterances, but machine translation is based on segmentation and – despite attempts to extend its attention to multisentence sequences (Popel 2020) – the contextual span it takes into account is usually one sentence (Rothwell et al. 2023, 104). Second, a dialogue is a situated communicative process in which participants have a goal that is co-constructed in a certain environment on the basis of (at least some) shared knowledge: the amount of information known, the relationship between interlocutors, their use of intonation and body language, and their experience of the world determine language and the degree of explicitness required, but none of these elements can be perceived, and hence recreated, by the software, which, moreover has no access to the physical and emotional experience of humans. Finally, from the point of view of cohesion, dialogues are likely to contain frequent cases of reference, substitution and ellipsis, which rely on the interlocutor's (and the reader's) ability to interpret the relations between elements, but interpretation is not part of the process applied by MT. For example, after the protagonist has accepted to play Isabella of Spain, she announces she will also play Columbus:

> "Now, I'm going to be Columbus!"
> "You are *not*," protested Winthrop. "I'm Columbus! You can't be Iserbella *and* Columbus! And you're only a girl. You gimme that revolver!"
> "I am, too, Columbus! I'm the best Columbus. So now! Why, you can't even tell me the names of Columbus's ships!"
> "I can too!"
> "Well, what were they?"
> "Well, I can't just --- Neither can you, smarty!"
> "Oh, I can't, can't I!" crowed Ann. "They were the Pinto and the Santa Lucheea and --- and the Armada!"
> "Gee, that's right. I guess she better be Columbus" [...].
> (Lewis 1933, 4)

DeepL's output is extremely literal, apart from (1), so no other backtranslation is provided. Unnatural and/or incorrect parts are shown in bold and commented upon below following the numbers attributed to each case.

> "Ora **diventerò** (1) Colombo!   [I will become Columbus]
>
> "**Non lo sei**," (2) protestò Winthrop. "Io sono Colombo! Non puoi essere Iserbella e Colombo! E sei solo una ragazza. Dammi quel revolver!"
>
> "**Lo sono anch'io**, (3) Colombo! Sono il miglior Colombo. **E così adesso**! (4) **Perché,** (5) non sai nemmeno dirmi i nomi delle navi di Colombo!"
>
> "**Posso farlo anch'io**!" (6)
>
> "E quali erano?"
>
> "**Beh io non posso proprio** (7)… nemmeno tu, furbacchion**e**!"
>
> "**Oh, non posso non posso!**" (8)**,** si affrettò a dire Ann. "Erano il Pinto e il Santa Lucheea e… e l'Armada!"
>
> "Accidenti, è vero. Credo **sia meglio che sia** (9) Colombo," […].

As it is, the exchange is hardly understandable since short answers and elliptical forms are mismatched. *Diventerò* [I will become] (1) is not wrong in itself, but in spoken informal Italian the future tense is rarely used, and in a context like this, a child announcing their character impersonation would probably use the present and the verb *fare* [do] instead of *diventare* [become], so the translation could be *Ora Colombo lo faccio io*. The boy's short answer "You are *not*" shows emphasis through the use of italics, and cohesion is realised through ellipsis; the literal version in Italian (2), however, does not relate to (1) in tense, because the answer is in the present (while the previous sentence was translated using the future), nor in meaning, as it contains the verb "be", not "become". (3) is ambiguous: by substituting *lo* for Columbus, strictly speaking the particle would be referring back to the last sentence containing the verb "be", which does not make sense as it refers to the condition of being a girl. Viable alternatives in current Italian could be *Faccio anche Colombo* [I'll be Columbus too] or *Anch'io faccio Colombo* [I too will be Columbus – i.e. not just you]. Interjections (4) and (5), literally translated, sound very unnatural and could be replaced by exclamations like *Insomma!* [Well/For heaven's sake!] and *Figurati* [Just think]. (6), (7), and (8) contain the inadequate choice of one of the two translatants of "can", *posso* and *so*, the former referring to

permission, the latter to ability or knowledge, which is what is meant here: challenged by Ann to name the three caravels, the boy claims he knows them but, when pressed, he tries to get away by suggesting she does not know them either. Only sentence (5), translated with the verb *sai* [you know], conveys the correct meaning, while in (6), (7), and (8) – although the cohesion resulting from the five repetitions of "can/can't" is visually reproduced – the meaning is lost. As a matter of fact, in both previous translations Jahn and Leonetti detach themselves from literalism and create lines that are natural and refer correctly to the notion of "knowing". Finally, the use of two subjunctives in (9) is not only unlikely by a child but it also introduces a high level of formality absent in the slangish "she better be".

MT works by selecting the words that, on the basis of the data on which the software is trained, are most likely to appear near each other – no interpretive process takes place. In dialogues – both in real life and in fiction – much of the meaning is made through the embodied experience of the world participants have had and are having, but software has no access to that situated cognition and therefore cannot recognise nor adequately translate certain situations typical of human life (Brusasco 2022).

METHODOLOGY

In order to assess the potential use of the MT version as either a pedagogical tool in the training of literary translators or an editing aid in view of the republication of an existing translation, the ST and each of the three Italian translations were aligned using an online alignment tool; then, the segments were checked for uniformity, and a file was created with four columns containing all versions side by side in order to compare them. Table 1 (below) shows a section of the file with the aligned texts. The character styles used for certain words or strings of words identify different types of translation problems: underlined signals errors in meaning; italics points at problems with naturalness, register, and/or language that has aged; bold is used for a broad category that includes hardly detectable errors, i.e. words or segments that are plausible but do not take into account intratextual connections or contain changes in word order that alter either meaning or register; underlined bold, instead, indicates particularly adequate solutions.

ST	DeepL	L. Jahn (1933)	I. Lionetti (1965)
While they debated, there came into that willow grove, that little leaf-littered place holy to boyhood, a singing girl.	Mentre discutevano, arrivò in quel saliceto, quel *piccolo luogo illuminato dalle foglie* e sacro alla **fanciullezza**, una ragazza che cantava.	[...] e ferveva la discussione quando nella *selvetta* dei salici, o meglio quel poco di terreno tutto ingombro di foglie ch'era propriamente **il bosco sacro dei ragazzi,** capitò invece una ragazza che cantava.	Mentre discutevano, nel boschetto di salice soffice di foglie, angolino sacro ai **fanciulli**, arrivò canterellando una ragazzina.
"Jiminy, there's Ann Vickers. She'll be Iserbella," said Winthrop.	"Grillo, ecco Ann Vickers. **Sarà** *Iserbella*", disse Winthrop.	- *Eccotela*! C'è l'*Anna* Vickers. **Può farla lei** l'Isabella – disse Winthrop.	"*Urca*, c'è Anna Vickers!" disse Winthrop. **"Isabella allora la fa lei".**
"Ah, no, gee, she'll hog the whole thing," said Ben. "But I guess she can play Iserbella better than anybody."	"Ah no, *caspita*, si prenderà tutto", disse Ben. "*Ma credo che possa interpretare Iserbella meglio di chiunque altro*".	- Ma va, **quella ci rovina tutto** – disse Ben. Però credo anch'io che per l'Isabella andrebbe **meglio lei di noialtri.**	"Ma va no. Ci rovina tutto!» disse Ben. "Però **mi sa** che Isabella la saprebbe fare meglio lei di **un altro**."
"Ah, she can not! She's no good at baseball.	"**Ah, non può**! Non è brava a giocare a baseball".	- **Macché**, non va bene. Anche a baseball *non è buona affatto.*	"Oh, no, *impossibile*! Non vale niente a baseball."
"No, she ain't much good at baseball, but she threw a snowball at Reverend Tengbom."	"No, **non è molto brava a baseball, ma ha tirato una palla di neve al reverendo Tengbom".**	- No, non è *gran che buona a baseball*, **ma è stata lei a** *gettare* **quella palla di neve al reverendo Tengbom.**	"No, *non a baseball*, però ha tirato una palla di neve al reverendo Tengbom."
"Yes, that's so, she threw that snowball."	"Sì, è così, ha lanciato quella palla di neve".	Già, questo è vero. **La palla di neve l'ha** *gettata lei*.	"*Già*, è vero, **la palla l'ha tirata.**"
The girl stopped before them, arms akimbo--a chunk of a girl, with sturdy	La ragazza si fermò davanti a loro, <u>con le braccia alzate</u>: un pezzo di	La ragazza, le mani sui fianchi, si fermò dinanzi a loro. Era un bel pezzo di	La ragazzina si fermò davanti a loro, mani sui fianchi, un pezzo di

shoulders and thin legs. Her one beauty, aside from the fresh clarity of her skin, was her eyes, dark, surprisingly large, and eager.	ragazza, con spalle robuste e gambe sottili. La sua unica bellezza, a parte la freschezza della pelle, erano gli occhi, scuri, sorprendentemente grandi e <u>desiderosi</u>.	ragazzina, con le spalle forti e le gambe magre. A parte la chiara freschezza della pelle, l'unica sua bellezza erano gli occhi, due occhi scuri, **vivaci** e straordinariamente grandi.	ragazzina con le spalle forti e le gambe sottili. Di bello, oltre al chiarore fresco della pelle, aveva gli occhi, scuri, incredibilmente grandi e <u>ardenti</u>.
"Come on and play <u>Iserbella</u> 'n' Columbus," demanded Winthrop.	"Vieni a giocare a *Iserbella* e Colombo", <u>chiese</u> Winthrop.	- Vieni a giocare <u>all'Isabella</u> e a Colombo – <u>chiese</u> Winthrop.	"Dài, vieni a giocare a Colombo e Isabella" <u>la invitò</u> Winthrop.
"I can't," said Ann Vickers. "I'm playing Pedippus."	"Non posso", disse Ann Vickers. "Sto giocando a Pedippus".	- Non posso – disse Anna Vickers. – Sto giocando a Pedippo.	**<u>"Non posso, sto giocando a Pedippo."</u>**
"What the dickens is Pedippus?"	"**<u>Che diavolo è Pedippus?</u>**".	**<u>- E chi diavolo è questo Pedippo?</u>**	E chi diavolo è **Pedippo**?
"He was an ole hermit. Maybe it was Pelippus. […] and he gave up all **the joys of the flesh** and he went and lived in the desert on-- oh, **on oatmeal and peanut butter** and so on and so forth, in the desert, and prayed all the time."	"Era un vecchio eremita. Forse era Pelippus. […] **<u>rinunciò a tutte le gioie della carne e andò a vivere nel deserto</u>** *con*… oh, farina d'avena e **<u>burro di arachidi e così via</u>**, nel deserto, e **<u>pregava tutto il tempo</u>**".	- Era un santo eremita**. Pedippo o forse** <u>anche</u> **Pelippo**, non so. […] e così aveva abbandonato le gioie della carne, e stava *continuamente* nel deserto, vivendo di semplice avena, sicuro… d'avena <u>e d'olio di noci</u>, e *non smettendo* mai di pregare.	"Era un santo eremita. *Pelippo, forse* […] lasciò *tutte* le gioie della carne e andò nel deserto e visse di <u>farinate e marmellata di noccioline</u> eccetera eccetera, *là* nel deserto, e **<u>pregava tutto il tempo.</u>**

Table 1. A section of the file with the aligned versions of Chapter 1

<u>Underlined: meaning</u>
italics: naturalness; register
<u>underlined bold: fluent rendering</u>
bold: ineffective word order; intratextual incoherence; hardly detectable error

As can be seen even from this short excerpt, Jahn's version tends to be wordier, a trend that is confirmed by the total length of the chapter: 2,514 words for a source text of 2,195, while Leonetti's is 2,175 and DeepL 2,192. It has to be said, though, that

in the passage quoted above Jahn succeeds in recreating the colloquiality and intentions of the characters. For example, in box 3, by using *meglio lei di noialtri* [better she than any of us] the translator makes the line sound very natural and restricts its scope to the boys in the scene, while the other two versions are correct but suggest that Ann can play Isabella better than anyone else in absolute terms. When one of the boys counters the remark about Ann's hopelessness at baseball by pointing out that she threw a ball at the Reverend, Jahn's wording emphasises that it was her who did it, thereby explicitating the boy's admiration. DeepL produces a less marked sentence, which is however fluent.

On the whole, sheer errors in DeepL's output are not many: *illuminato dalle foglie* [lit by leaves] poetically conjures the image of golden leaves on the grove floor and might therefore go unnoticed, but the source text actually describes a layer of leaves without referring to its colour; *Grillo* [cricket], to render "Jiminy", an exclamation of surprise of dismay, could be dismissed as a hallucination – the term used to refer to MT translations that are completely off the target – but since the software selects words according to patterns of co-occurrence, the reason probably lies in Walt Disney's character Jiminy Cricket in the fairy tale "Pinocchio". The paragraph describing Ann (seventh cell in Table 1 above) contains two errors that without a comparison with the source text might go unnoticed: the girl is shown with "arms akimbo" – a self-confident, defiant posture – and "eager eyes", but in the translation she has *braccia alzate* [raised arms] and *occhi desiderosi* [eyes full of desire], which suggests surrender or even the preaching attitude of the old hermit she mentions, and adds a sexual connotation absent from the source text. These apparently minor changes actually result in a different characterization of the protagonist and partly contradict her attitude and words in the following dialogues with her friends. While it is true that *desiderosi* can translate "eager", the context suggests opting for something that expresses intensity, curiosity, vivacity, especially because the same adjective appears in the opening paragraph of the novel to describe the children in the scene: "Four children, sharp-voiced, and innocent and eager […]" (Lewis 1933, 1). DeepL maintains *desiderosi* in both occurrences, as ideally should happen in the name of consistency, but the lexical item is inadequate in both contexts. Lionetti translates the first occurrence as *avidi* [avid/greedy] and the second as *ardenti* [ardent/glowing], thereby disrupting the internal connection and adding both a negative connotation and an aspect of love or passion respectively. Jahn

shifts the meaning of the first occurrence to *allegri* [cheerful] but captures the second through the choice of *vivaci* [lively].

Disseminated through Lewis's first chapter are a number of deliberate mistakes in the children's lines meant to benevolently mock their mastery of certain areas of knowledge they are not fully familiar with – "Iserbella" instead of Isabella; "concert" instead of "consort" when Ann, impersonating Queen Isabella tries to speak in an elevated style; "pagodas" as dwellings of the Indians, etc. In the two human translations, conceptual errors are maintained, but those based on spelling and pronunciation have been corrected, probably because Italian is a phonetic language, with the result that some humour is lost. DeepL retains "Iserbella" and translates "concert" literally as *concerto*, which is correct but meaningless in the sentence; a solution might be *consorziato* [consortium member], which would be a plausible distortion of a formal, unusual word by a child.

The textual aspects discussed so far are among the translation problems that the PACTE research group (Process in the Acquisition of Translation Competence and Evaluation) identifies as "rich points":

- Linguistic problems: lexical (non-specialised) and morphosyntactic
- Textual problems: coherence, cohesion, text type and genre, and style
- Extralinguistic problems: cultural, encyclopaedic and subject-domain knowledge
- Problems of intentionality: difficulty in understanding information in the source text (speech acts, presuppositions, implicature, intertextual references)
- Problems relating to the translation brief and/or the target-text reader (affecting reformulation) that, from a functionalist point of view, would affect all Rich Points. (PACTE 2014, 90)

Interestingly, the rich points highlighted by the PACTE group as linguistic items to which trainee translators should pay attention largely coincide with the errors in the MT output, an aspect that can provide guidance in devising pedagogic activities.

SOME PEDAGOGIC CONSIDERATIONS AND PROPOSALS

The cases illustrated above are in line with previous studies (Fonteyne et al. 2020; Guerberof Arenas and Toral 2020; Brusasco 2022, among others) showing that – despite neural machine translation's improvements and better outputs – the complexity, multilayeredness, and pragmatic dimensions of literary texts still require the interpretive and stylistic skills of a human translator. However, since MT is currently the most widely used translation technology (Rothwell et al. 2023) and is likely to continue improving thanks to artificial intelligence and the expansion of training data, it seems wise to include at least some basic MT literacy in the training of literary translators, and find ways of using it that can assist the human translator while keeping his/her centrality.

A hypothesis is to use MT as an aid when an old but overall good translation needs to be refreshed in view of a new publication, as could be the case with "Ann Vickers". Since revision is a long job, with far lower fees than translation itself, the alignment of the existing translation(s) and the MT output might provide the translator/reviser fast, adequate solutions for sections previously identified as dated or containing mistakes. The procedure, which could be considered a literal application of Emmerich's notion of translation as "a form of translingual editing by which a translator both negotiates existing versions and creates a new one of her own" (2017, 2), is explored in a co-authored article (Brusasco and Taivalkoski-Shilov, forthcoming).

The suggestion put forward here, instead, is to consider MT's shortcomings and use them in (literary) translator training in order to both alert students to potential errors should they postedit an MT output, and enhance their focus on textual dimensions that contribute to making (their) translated texts pragmatically sound and consistent in interpretation. One of the crucial aspects is the contrast between the speed of MT and the slow pace of a human translator interpreting a text and honing the target language after all the elements contributing to the shaping have been taken into consideration. Such craftsmanship is endangered by changed reading and writing habits, and by the growing exposure to AI-generated texts. The spread of NMT itself contributes to a levelling of styles, genre conventions, lexical choices and overall complexity: TAUS 2016 "Translation Technology Landscape Report" predicted that "the world will get accustomed to what we

call Fully Automatic Useful Translation (FAUT) and will more and more accept this as the norm for standard translation" (Rothwell et al. 2023, 214). Luckily, texts are still ranked differently depending on their value, purpose, and risk potential, not only related to the quality of the translation, but to the consequences of their use in – for example – medical, financial or legal settings; therefore, language service providers adopt different protocols varying from fully automated translation to postediting, to professionals working with the aid of technology, which means that "useful translation" has not become the norm yet. Except for urgent information, like life-saving instructions in emergencies, and possibly for totally functional texts like instruction manuals, language and translation should retain their full range of complexity and expressiveness, a point that is particularly true in the case of creative texts. As Tymoczko (2014) noted, literary texts are ideal for training because of the variety and pliability of language, as well as the representation of numberless communicative situations and human experiences. Training activities, therefore, are geared towards the acquisition of procedural knowledge, but a strong emphasis is placed on cognition and metacognitive competence, i.e. "the ability to self-regulate cognitive processes that contribute to goal achievement and the professional success of translators" (Pietrzak 2022, 3). Moreover, the activities suggested are deliberately slow in order to both re-ground reading and writing and make the most of the time dedicated to the acquisition of knowledge and skills. Unless differently specified, activities are not intended in any specific sequence since the choice and order will be determined by the text to be translated or revised.

A top-down translation-oriented text analysis (Nord 2005) combined with close reading strategies still plays an important role when quality is at stake, be it for translation or for postediting. A clear view of the text in its entirety, with a detailed mapping of lexical relations and cohesive devices, an evaluation of staticity versus dynamism as conveyed by prevalent stative as against dynamic verbs and adjectives, the prevalence of certain word classes and lexical items, levels of meaning, information flow, intertextuality, reverberations throughout the texts – all this allows students to experience the text(ure), while careful consideration of culture-specific elements and implicatures will determine the degree of mediation necessary in order for the prospective reader to fully appreciate the text. Such analysis should be preceded and followed by experiential reading, approaching the text as a whole and allowing it to connect with personal experiences and previous

26

knowledge of other texts. This step is meant to allow students to become aware of more levels of signification and to activate language heard or used in the stories and situations evoked.

As noted above, reasonable predictions can be made about the "tricky" spots in a text – PACTE's "rich points" (2014): developing students' awareness of the translation problems a source text poses during the preliminary analytical phase will guide them in their decision-making. The same awareness, however, could lead to more effective revision or post-editing. A study conducted by Volkart et al. (2022), showed that students corrected only about 50% of the errors present in the MT output, which may partly depend on the fact that, unless forced to, they tended to ignore the source text and carry out a predominantly monolingual post-editing. Considering that NMT has increased the number of hardly detectable errors, it would be important for students to know what to look for when post-editing so as not to be mislead by fluent readability. Within a comparative approach, preliminary source-text analysis followed by moving back and forth between ST and MT output, as well as – when available – previous versions by human translators, is likely to sharpen the students' sensitivity to rich points, widen their expressive range and increase correctness in the case of post-editing.

Translation skills and sharper reading in view of post-editing may be promoted via completion activities in both languages to promote hypothesis formation and expectations. This entails working at microlevel with incomplete sentences, and at macrolevel – with missing paragraphs to be supplied by students on the basis of the context. An active approach to reading co-constructs meaning along the way, and it is crucial that translation students develop an ability to stop at any moment trying to imagine what may come next, be it while reading a narrative text or an instruction manual. This would be particularly useful in cases where keeping close to the source text results in a target text that is fluent and apparently correct, but whose meaning ends up deviating from what is logical, as in the examples mentioned above – "eager eyes" rendered as *ardenti* by Leonetti and "arms akimbo" as *braccia alzate* by DeepL. Incidentally, creating logical and linguistic expectations may be considered the human equivalent to the basic principle of NMT – selecting the word(s) that are most likely to appear near each other, but endowed with experience and creativity.

Closely connected to this would be asking students to develop a sort of "mental script" portraying the environment, the situation

and the participants, in order to produce dialogue which sounds natural and cohesive (see the problems highlighted in the examples above). They could ask questions such as "In what situation might this sentence be used?", and "What do people (or what would I) say in that situation?".

More general activities include focusing on unusual collocations, both in the ST in view of translation, and in an MT output as an exercise towards more accurate PE. Corpora in both languages will clarify if a lexical pair is a collocation or a creative deviation from norm that can be recreated or maintained without jeopardising meaning, or, in the case of MT, a glitch that has to be solved. Reading and writing comparable texts in the target language, carrying out extra-linguistic research on the text topics, dramatization in the case of literary texts, followed by moments of reflection on how translation strategies and procedures were affected, are other activities conducive to a wide and varied approach to translation and post-editing.

CONCLUSION

Technology has changed the world of translation, focusing on quantity, i.e. speed of delivery. NMT has outperformed previous models and its output is approaching human quality. Literary texts, however, still challenge NMT because they require interpretation, creativity and experience in and of the world, hence human translators. Yet, the economic advantages of NMT and the growing acceptance of "Fully Automated Useful Translation" (Rothwell et al. 2023, 214) go against the slow, high-quality-searching activity of the literary translator. Against this backdrop, the present article is a reflection on some uses of NMT in specific areas of literary translation, namely for revision in view of republication (Brusasco and Taivalkoski-Shilov, forthcoming) and for pedagogic purposes. Examples drawn from the existing Italian versions of Lewis's "Ann Vickers" (1933) compared with DeepL's output have been discussed to point out both the lexical areas where the software improves the existing text and the items or passages which instead require the interpretive action of a human translator. The findings have prompted some pedagogic considerations: given the rapidly evolving situation, some basic MT literacy can be useful to literary translators too, who might use it as a tool to hone their skills and enhance their linguistic sensitivity; at the same time, however, familiarising themselves with the post-editing of creative texts may

also be a way of retaining some human control on a pervading technology.

REFERENCES

1. Bommasani, R. et al. (2021). "On the Opportunities and Risks of Foundation Models." *ArXiv*. https://doi.org/10.48550/arXiv.2108.07258 Available at: <https://crfm.stanford.edu/assets/report.pdf> (accessed August 2023).

2. Brusasco, P. (2022). "Pragmatic and Cognitive Elements in Literary Machine Translation: An Assessment of an Excerpt from J. Polzin's 'Brood' Translated with Google, Deepl and Microsoft Bing." In: Hadley J. L., Taivalkoski-Shilov, K., Teixeira, C. S. C. and Toral, A. (eds.) *Using Technologies for Creative-text Translation*. London – New York: Routledge, 139–160.

3. Brusasco, P. and Taivalkoski-Shilov, K. (forthcoming). "New Uses for Translation Technology? Revising 'The Job' in Italian and 'Silent Spring' in Finnish with the Help of Translation Memory and Machine Translation." *Palimpsestes*.

4. Emmerich, K. (2017). *Literary Translation and the Making of Originals*. New York, London: Bloomsbury.

5. Fonteyne, M., Tezcan, A. and Macken, L. (2020). "Literary Machine Translation under the Magnifying Glass: Assessing the Quality of an NMT-Translated Detective Novel on Document Level." In: Calzolari, N. et al. (eds.) *Proceedings of the 12th Language Resources and Evaluation Conference*. Marseille: European Language Resources Association, 3790–3798. Available at: <https://aclanthology.org/2020.lrec-1.468> (accessed August 2023).

6. Guerberof Arenas, A. and Toral, A. (2020). "The Impact of Post-editing and Machine Translation on Creativity and Reading Experience." *Translation Spaces* 9 (2), 255–282.

7. Hadley, J. L., Taivalkoski-Shilov, K., Teixeira, C. S. C. and Toral, A., eds. (2022). *Using Technologies for Creative-text Translation*. London, New York: Routledge.

8. Halliday, M. A. K. and Hasan, R. (1976). *Cohesion in English*. London: Longman.

9. Kenny, D. and Winters, M. (2020). "Machine Translation, Ethics and the Literary Translator's Voice." *Translation Spaces* 9 (1), 123-149.

10. Kuzman, T., Vintar, Š. and Arčan, M. (2019). "Neural Machine Translation of Literary Texts from English to Slovene." In: Hadley, J., Popovic, M., Afli, H. and Way, A. (eds.) *Proceedings of the Qualities of Literary Machine Translation*. Dublin: European Association for Machine Translation, 1–9. Available at: <https://aclanthology.org/W19-73.pdf> (accessed April 2024).

11. Lewis, S. (1933). *Ann Vickers*. New York: Doubleday, Doran & Co.

12. Lewis, S. (1933). *Anna Vickers. Romanzo di una donna americana*. Trans. L. Jahn. Milano: Treves.

13. Lewis, S. (1965). *Anna Vickers*. Trans. I. Leonetti. Milano: Mondadori.

14. Montemayor, C. (2023). *The Prospect of a Humanitarian Artificial Intelligence: Agency and Value Alignment.* London: Bloomsbury Academic.

15. Nord, C. (2005). *Text Analysis in Translation. Theory, Methodology, and Didactic Application of a Model for Translation-Oriented Text Analysis*. 2nd edition. Amsterdam, New York: Rodopi.

16. Nord, C. (2007). "Function Plus Loyalty: Ethics in Professional Translation." *Génesis. Revista Científica do ISAG* 6, 7–17.

17. PACTE Research Group. (2014). "First Results of PACTE's Group Experimental Research on Translation Competence Acquisition: The Acquisition of Declarative Knowledge of Translation." *MonTI Special Issue – Minding Translation*, 85–115. Available at: <http://www.e-revistes.uji.es/index.php/monti/article/view/1702> (accessed July 2023).

18. Pietrzak, P. (2022). *Metacognitive Translator Training: Focus on Personal Resources*. Cham: Springer Nature Switzerland – Palgrave Macmillan.

19. Popel, M. (2020). "CUNI English-Czech and English-Polish Systems in WMT20: Robust Document-Level Training." In: Barrault, L. et al. (eds.) *Proceedings of the Fifth Conference on Machine Translation*. Association for Computational Linguistics, 269–273. Available at: <https://aclanthology.org/2020.wmt-1.28> (accessed September 2023).

20. Rice, D., Rhodes, J. H. and Nteta, T. (2019). "Racial Bias in Legal Language." *Research & Politics* 6, 2. https://doi.org/10.1177/2053168019848930 (accessed August 2023).

21. Rothwell, A., Moorkens, J., Fernández-Parra, M., Drugan, J. and Austermuehl, F., eds. (2023). *Translation Tools and Technologies*. London, New York: Routledge.

22. Schleiermacher, F. (2004). "On the Different Methods of Translating." In: Venuti, L. (ed.) *The Translation Studies Reader*. Oxon – New York: Routledge, 43–63.

23. Taivalkoski-Shilov, K. (2018). "Ethical Issues Regarding Machine(-assisted) Translation of Literary Texts." *Perspectives. Studies in Translatology* 27 (5), 689–703. Available at: <www.tandfonline.com/doi/full/10.1080/0907676X.2018.1520907> (accessed June 2023).

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

25. Toral, A. and Way, A. (2015). "Machine-assisted Translation of Literary Text: A Case Study." *Translation Spaces* 4 (2), 240–267.

26. Toral, A. and Way, A. (2018). "What Level of Quality Can Neural Machine Translation Attain on Literary Text?" In: Moorkens, J., Castilho, S., Gaspari, F. and Doherty, S. (eds.) *Translation Quality Assessment. Machine Translation: Technologies and Applications*. Cham: Springer, 263–287.

27. Tymoczko, M. (2014). "Why Literary Translation is a Good Model for Translation Theory and Practice." In: Boase-Beier, J., Fawcett, A. and Wilson, P. (eds.) *Literary Translation: Redrawing the Boundaries*. Basingstoke: Palgrave Macmillan, 11–31.

28. Volkart, L., Girletti, S., Gerlach, J., Mutal, J. D. and Bouillon P. (2022). "Source or Target First? Comparison of Two Post-editing Strategies with Translation Students." *Journal of Data Mining and Digital Humanities*, *Towards robotic translation?* https://doi.org/10.46298/jdmdh.9067 (accessed July 2023).