

**Tjaša Mohar, Sara Orthaber,**

**Tomaz Onič**

University of Maribor, Slovenia

2020, Vol. 17 (1), 125-141(164)

revije.ff.uni-lj.si/elope

<https://doi.org/10.4312/elope.17.1.125-141>

UDC: 821.111(71).09-32Atwood M.:81'322.4

“I exist in two places, here  
and where you are” — Margaret Atwood

## Machine Translated Atwood: Utopia or Dystopia?

### ABSTRACT

Margaret Atwood's masterful linguistic creativity exceeds the limits of ordinary discourse. Her elliptical language contributes to interpretative gaps, while the ambiguity and openness of her texts intentionally deceive the reader. The translator of Atwood's texts therefore faces the challenge of identifying the rich interpretative potential of the original, as well as of preserving it in the target language. Witnessing the rise of artificial intelligence, a natural question arises whether a human translator could ever be replaced by a machine in translating such challenging texts. This article aims to contribute to the ongoing debate on literary machine translation by examining the translations of Atwood's "Life Stories" generated by two neural machine translation (NMT) systems and comparing them to those produced by translation students. We deliberately chose a literary text where the aesthetic value depends mostly on the author's personal style, and which we had presumed would be problematic to translate.

**Keywords:** literary machine translation; literary translation; neural machine translation (NMT); Margaret Atwood; short fiction

## Margaret Atwood v strojnem prevodu: Utopija ali distopija?

### POVZETEK

Jezikovna ustvarjalnost Margaret Atwood presega meje običajnega diskurza. Njen eliptičen jezik ustvarja interpretativne vrzeli, dvoumnost in odprtost njenih besedil pa včasih namerno zavajata bralca. Prevajalec njenih del je tako nedvomno postavljen pred izziv, kako prepoznati bogat interpretativni potencial originala in ga ohraniti v ciljnem jeziku. V času, ko smo priča vzponu umetne inteligence, se naravno postavlja vprašanje, ali obstaja upanje (oziroma strah), da bo prevajalca pri prevajanju besedil avtorjev, kot so Margaret Atwood, nekoč zamenjal prevajalnik. Članek prispeva k nadaljevanju razprave o strojnem prevajanju literature, s tem ko analizira slovenska prevoda kratke proze Margaret Atwood z naslovom »Life Stories«, dobljena z uporabo dveh nevronske strojne prevajalnikov (Google Translate in Translator.eu), in ju primerja s prevodi študentov prevajalskih študij. Za namene raziskave smo namenoma izbrali literarno besedilo, katerega estetska vrednost temelji predvsem na avtoričinem slogu in za katerega smo predvidevali, da bo predstavljal prevajalski izziv.

**Ključne besede:** strojno prevajanje književnih del; književno prevajanje; nevronske strojno prevajanje; Margaret Atwood; kratka proza

# 1 Introduction

Margaret Atwood's literary works are known for their complex and recognizable style, which, owing to her masterful linguistic creativity, frequently exceeds the limits of ordinary discourse. Her elliptical language contributes to the interpretative gaps, ambiguity and openness of the text that sometimes lead to the intentional deception of the reader. As Gadpaille (2014, 172–75) points out, the complexity of Atwood's tone, particularly her “deep irony and black humour,” can easily be misinterpreted. The translator of Atwood's texts undoubtedly faces the challenge of identifying the rich interpretive potential of the original, as well as of preserving it in the target language.

In a recent study, Borg (2016) examined how a literary translation comes into being by providing a rich description of the process, from the way the translators approach the task to the decisions and the choices they make. Among the invisible phases of the process, she explicitly lists one dedicated to polishing the style of the target language. Translators are, therefore, not simply concerned with transferring the propositional content of the source text; they also need to capture the different nuances of pragmatic meaning as well as the aesthetic value, which Schwartz & de Lange (2006, 18–19) refer to as a subjective, unique, and distinctive personal rendition of the original. A source text “is providing a certain way of knowing the world which the target text (TT), by the very act of translation [...] is invited to diversify and supplement” (Scott 2018, 18). Given the unique style of literary texts, automated translation of them is thus considered a great challenge, and for a long time the idea of literary machine translation (MT) has been rejected. Rapid development of neural machine translation (NMT) systems has increased their use by professional translators and thus affected the process of translation. The possibility of post-editing pre-translated literary texts has recently been revisited. This article aims to contribute to the ongoing debate on literary machine translation for the English-Slovene language pair by examining the translations of a selected work by Margaret Atwood produced by students of translation studies and comparing these to the output of the neural MT systems Google Translate and Translator.eu (GT and T.eu, respectively).

Accordingly, the following section will briefly examine the relevant research implemented in the field of literary neural machine translation (NMT) and then present the methodologies and approaches to human and automated evaluations of MT outputs. We then move to the description of data and explain the analytical approach adopted in this study. In the analysis, we compare the students' translations with those generated by GT and T.eu in eight linguistic and stylistic categories.

## 2 The Rise of Literary Neural Machine Translation (NMT)

That NMT can yield promising results in literary translation was demonstrated in a recent experimental study by Toral and Way (2018), who found that technology was capable of translating twelve widely known novels from English into Catalan with a 25% flawless rate (meaning that in 25% of the translated text, no post-editing intervention at the sentence level was needed). These novels were Hemingway's *The Old Man and the Sea*, Joyce's *Ulysses*,

Salinger's *The Catcher in the Rye*, Rowling's *Harry Potter and the Deathly Hallows*, Orwell's *1984* and Tolkien's *The Lord of the Rings*. In their study, Toral and Way (2018, 285) conducted a human evaluation of translations for three of the books by manually ranking the translations produced by NMT and the human translations. For two out of three of these, native speakers perceived NMT translations to be of comparable quality to those of human translations in around one-third of the cases.

One reason for this relative success is NMT's use of deep learning, a technique that according to Bonner, "teaches a computer to filter inputs through layers to learn how to predict and classify information" (2019). The main resource for training machine translation systems comprises bilingual parallel texts. Put differently, the aim of NMT is to "understand the meaning behind words and phrases by predicting word order and applying context via deep learning" (Kravariti 2018). With the rapidly growing market share of e-books, i.e., original novels and their translations, MT systems have been built that are tailored to novels.

Considering the challenging text type, the study by Toral and Way has shown surprisingly promising performance in literary MT, possibly suggesting that literature may no longer be off-limits to technology. We have therefore chosen to test the performance of NMT on a work by Margaret Atwood ("Life Stories" from *The Tent*), where the aesthetic value of the text intrinsically relies not on content and plot, but rather on form and Atwood's individual style in the sense of Leech and Short, who define an author's style as "those aspects of linguistic choice which concern *alternative ways of rendering the same subject matter*" (2007, 31).

## 2.1 English to Slovene Combination

Studies of machine literary translation for the language combination English and Slovene, which is the language pair examined in this study, are still scarce.<sup>1</sup> According to Kuzman, Vintar and Arčan (2019), this is due to the lack of availability of resources such as translation corpora and parallel corpora of literary texts, and to the morphological complexity of the Slovene language<sup>2</sup> (see also Vintar 2013; 2018). In their study, Kuzman, Vintar and Arčan (2019) examined the productivity and output quality of literary machine translation by comparing neural literary translations for English-to-Slovene translations among different neural machine translation (NMT) models: (a) NMT Google Translate, a mixed-domain neural machine translation model trained on millions of examples, and (b) a model tailored to literature, but trained on a significantly smaller data set. Applying both automatic and human evaluation, their main objective was to explore whether such an approach could improve the performance of NMT systems. They found, however, that the mixed-domain NMT model GNMT performed better than the model tailored to literature.

<sup>1</sup> Computer-aided translation software that includes translation memories has mainly been developed for language pairs with substantial commercial interest, Slovene not being one of them. Most translations of literary works into Slovene are from English (62%), followed by French (41%) and Italian (18%). Translations into English, on the other hand, account for just 4% (Pregelj 2019). Moreover, English and Slovene belong to different language families, Germanic and Slavic, respectively.

<sup>2</sup> This observation is in line with the findings of the literary MT study for the language pair English-Russian by Matusov (2019), who found that although the quality was high enough to understand and enjoy the translated story, the NMT of German fiction into English was better than that of English into Russian.

In the aforementioned study, translation outputs were evaluated both automatically (using metrics such as BLEU or METEOR)<sup>3</sup> and by human evaluators, to compensate for their respective shortcomings.<sup>4</sup> In human evaluation, qualified translators typically check for errors in the MT output (also known as post-editing) and assess its fluency and adequacy (on detailed presentation of error classification and error typologies, see Popović 2018). One way to conduct the MT evaluation is for bilingual human evaluators to rate the MT output for meaning preservation (i.e., adequacy), grammaticality (i.e., fluency), and a combination of both (i.e., overall quality) from poor (1) to perfect (5). Evaluations also often include the temporal dimension of post-editing effort. Alternatively, for each output sentence, the evaluators choose a better version (A or B), without committing themselves to an absolute score (Popović 2018, 130).

When comparing the English-Slovene language pair, Kuzman, Vintar and Arčan found that for two out of three participants, the translation task was more time-consuming, yet they spent less time translating than post-editing, while the error analysis of literary MT identified “punctuation errors, wrong translations of prepositions and conjunctions, inappropriate shifts in verb mood, wrong noun forms and co-reference changes” (2019, 7). Translations also included semantic errors such as assigning the wrong gender to the main character, translating proper nouns as common nouns, mistranslation of idioms and ambiguous words, and other inconsistencies. Moreover, rare words were left untranslated and sometimes words were either added or omitted. Other errors included changing numbers, substituting imperial units for metric units without converting the values. As will be shown in the analytical part of this paper, our error analysis of the machine translations yielded very similar results.

### 3 Choice of Text: Margaret Atwood, “Life Stories”

Margaret Atwood’s two-page “Life Stories,” published in her 2006 collection *The Tent*, was chosen for this research for two main reasons: 1) Atwood is a world-famous author with a recognizable writing style that is a challenge to preserve in translation, mainly because of Atwood’s penchant for ambiguity; 2) her style is well represented in this piece of short fiction, which eludes detailed categorization. Atwood’s experimentation with the short fiction genre has been acknowledged by several critics (e.g., Nischik 2006; Slettedahl Macpherson 2010; Gadpaille 2018), and *The Tent* undoubtedly counts as such an experiment. Nischik points out that Atwood’s short fiction often contains other genres, such as “mini essays, ‘essay fictions’, short dialogues, dramatic monologues, and reflections” (2006, 153), and Slettedahl Macpherson extends this even to poetry by placing it “between story, essay, and prose poem” (2010, 87). “Life Stories,” indeed, reads like a mini essay; it could also be defined as a dramatic monologue or as reflections by its first-person narrator. Owing to the reflective mood of the piece, the narrator’s voice could easily be mistaken for the author’s; however, Slettedahl Macpherson argues that the pieces published in *The Tent* contain only “snippets

---

<sup>3</sup> BLEU stands for BiLingual Evaluation Understudy, and METEOR for Metric for Evaluation of Translation with Explicit ORdering.

<sup>4</sup> Human evaluation is perceived as slow, expensive and inconsistent, whereas automated evaluations are still used at the sentence level, are often subject to linguistic shortcomings of a specific MT system, and assume that there is a single correct translation (Way 2018).

of Atwoodian panache, masquerading as knowledge about the author,” yet at the same time “slipping back from genuine revelation” (2010, 87). Gadpaille finds that in *The Tent* “Atwood was freed from genre restrictions, and thus able to play with concepts from psychology, philosophy and environmental science, while adapting familiar genre motifs and structures, with amoeba-like flexibility” (2018, 19). In “Life Stories” the narrator is “taking apart” her life story, and in doing so inevitably touches upon psychological and philosophical concepts. In combination with Atwood’s writing style, these add to the considerable complexity of the literary work; therefore, the NMT systems and the students (trainee translators) are expected to encounter translation problems.

#### 4 Methodology and Analytical Approach

Eight master’s students in translation studies (English-Slovene) translated “Life Stories”<sup>5</sup> as a take-home assignment within their regular course work. Before translating it, they were given some basic information about Atwood and her work; prior to this task, they had translated three Atwood prose poems from her collection *The Door* (2007), and their translations were discussed in class. The students had no other previous experience with translating Atwood’s works.

The same text was then translated by two NMT systems, Google Translate and Translator.eu, both freely available, multilingual, automated translation tools. GT was first launched in 2006; a decade later it switched to a neural machine translation engine GNMT. This means its architecture “consists of two recurrent neural networks, one to consume the input text sequence and one to generate translated output text” (Wu et al. 2016, 1). Currently, the service is available for 103 languages, including Slovene, and translates 5000 characters per request, which is impractical for longer texts, e.g., novels. Similarly, Translator.eu is a publicly available online translation tool that provides translation into 40+ languages. It currently translates 1000 characters per request.

At the first stage of the research analysis, the students’ translations of “Life Stories” were checked and evaluated by the course leader, who is an experienced literary translator; a teacher of English and American literature with over fifteen years of experience in teaching literature as well as literary translation classes; and an English language instructor and practicing translator with ample experience in teaching translation courses. The evaluators first rated the student translations for meaning preservation, grammatical correctness and overall quality, focusing also on potential translation shifts and their nature. The student translations were then compared to the control version that consisted of translation solutions established by the evaluators as the most accurate in terms of semantics, grammar and style (in the example section referred to as Ctrl). In the following, the translations generated by the GT and T.eu translation tools were examined by the three evaluators according to the same criteria that were used for the student translations and compared to the control version. Based on the results, eight evaluation categories were established in order to identify the grammatical, semantic and/or stylistic shortcomings of the translation tools. Finally, the students’ translations were compared to those offered by the translation tools and the control.

---

<sup>5</sup> The text has 498 words (2808 characters with spaces).

## 5 Contrastive Textual Analysis

In this section we present the results of our small-scale experiment. The errors and deficiencies identified in the translations of “Life Stories,” provided by GT and T.eu, were manually classified into eight linguistic and stylistic categories.

### 5.1 Short, Non-Idiomatic Sentences, Simple Syntax

As examples [1–4] illustrate, from among the linguistic and stylistic categories dealt with in this study, both MT systems performed well in translating short, non-idiomatic sentences with simple syntax; these were translated semantically and grammatically correctly.

- [1] Orig:<sup>6</sup> *It's mostly a question of editing.*  
GT and Ctrl: *Večinoma gre za vprašanje urejanja.*  
T.eu: *To je predvsem vprašanje urejanja.*

A translation similar to that by GT appears in six out of eight of the student translations and in the control translation. The T.eu version is also semantically acceptable.

- [2] Orig: *It helps if there are photos.*  
GT, T.eu and Ctrl: *Pomaga, če obstajajo fotografije.*

GT and T.eu translated the verb “to be” as “obstajati” (“to exist”), which is common in Slovene when translating the English phrase “there is/there are”; three out of eight students also opted for this solution. Two students kept the verb “biti” (“to be”), two used the verb “imeti” (BT:<sup>7</sup> “if we have photos”), and one simplified the sentence into “photos help.” These solutions are not favoured; the first one lacks stylistic naturalness in Slovene, while the latter two unnecessarily change the original sentence and thus affect Atwood’s writing style.

- [3] Orig: *I spent my childhood.*  
GT and T.eu: *Otroštvo sem preživel.*  
Ctrl: *Otroštvo sem preživela.*

Here, both NMT systems took “I” as masculine, which is visible from the Slovene past participle, while all students as well as the control assume a female subject. The narrator’s gender is, in fact, not revealed in the original text; however, owing to the autobiographical impression an average reader gets when reading “Life Stories,” the female form seems more appropriate. Some students had problems translating this sentence: one changed the word order (“Preživela sem otroštvo”; BT: “I survived childhood”), which changes the original meaning, while four others mistranslated the verb “spend” as “zapraviti” (BT: “I wasted my childhood”), possibly owing to the complexity of the context in which the sentence appears.

- [4] GT: *Ko začnete, je zabavno.*  
T.eu: *Ko začneš, je zabavno.*  
Ctrl: *Ko enkrat začneš, je zabavno.*

---

<sup>6</sup> Atwood’s original.

<sup>7</sup> Back-translation.

In GT's translation of [4], "you" is a plural pronoun, while in T.eu's version, it is singular. Since no formal difference exists between the two in English, the grammatical number must be deduced from the context. In [4], however, "you" is generic, so either form can be used in Slovene. All students as well as the control opted for the singular form. Additionally, five students added the adverb "enkrat" ("Ko enkrat začneš"), which roughly corresponds to the English adverb "once" in this context. This addition also appears in the control.

## 5.2 Unconventional or Elliptical Syntax

Atwood's unconventional use of language is well manifested in "Life Stories," which contains numerous sentences with unusual syntax or ellipsis, and these, even short ones, proved to be problematic for NMT; they also presented a challenge to the students.

- [5]     Orig: *I was born, I would have begun, once.*  
           GT: *Rodil sem se, nekoč bi začel.*  
           T.eu: *Rodil sem se, jaz bi začel, enkrat.*  
           Ctrl: *Rodila sem se, bi nekoč začela.*

It seems that both NMT systems translated the two clauses ("I was born," "I would have begun") separately, thus producing incongruent translations in Slovene. Since the first clause is direct speech and the second one a reporting sentence – the omitted quotation marks being Atwood's stylistic choice – the Slovene word order of the latter must be adapted. The adverb "once" needs to be included in the second clause in order to avoid ambiguity (see, also, Ctrl). Slovene uses the same form of conditional for the present and the past; without the adverb "nekoč" it is not clear that "bi začela" refers to the past. In this case, style needs to be sacrificed for the sake of clarity. GT did place the adverb "once" in the standard position but did not adapt the word order, while T.eu translated it as a separate item. Most students also failed to do justice to the original sentence; only three preserved the meaning. One student mistranslated the adverb "once" as "včasih" (BT: "sometimes"), while the other four translations contain additions or change the meaning. The difficulty probably stems from their (mis)reading of the original and consequent failure to make use of the main advantage they have over NMT, i.e., the ability to take into consideration Atwood's particular style.

- [6]     Orig: *Rip, crumple, up in flames, out the window.*  
           GT: *Odtrgajte, zdrobite, v plamenih, skozi okno.*  
           T.eu: *Rip, crumple, v plamenih, skozi okno.*  
           Ctrl: *Strgaj, zmečkaj, zažgi, vrzi skozi okno.*

The elliptical sentence in [6], also very Atwoodian, is mainly problematic for translation because of the ambiguity concerning verb moods as well as its combination with the ellipsis. Owing to their form, the verbs "rip" and "crumple" can be read as imperatives or infinitives in the original, while conscious choices must be made in the translation process. Moreover, the ambiguity of the subsequent elliptic phrases must also be carefully worded in Slovene for the interpretive potential to be preserved. T.eu left the words "rip" and "crumple" untranslated, while the GT's solutions produce a semantic shift ("odtrgajte" means to "rip (something)

off,” “zdrobite” means “to crumble”). The elliptical and idiomatic clause “up in flames” is translated by both systems as “in the flames,” while in the control version, the verb “zažgi” (BT: “burn”) has been used, which sounds better than the more direct version “v ogenj;” it also preserves parallelism in the list of imperatives. For coherence and clarity, a verb has also been added to the ellipsis “out the window” (BT: “throw out of the window”). Six out of eight students translated “rip” and “crumple” as verbs: one in first-person singular and four in second-person singular (two as the imperative mood). One student translated the verbs as infinitives, which sounds awkward in Slovene; and two students changed them into adjectives (“ripped,” “crumpled”), which creates a semantic shift.

In examples [7–9], non-standard syntax or ellipsis appears in long complex sentences, which proved to be even more problematic for NMT. Deficiencies were also identified in students’ translations.

[7] *Orig.: They should have spotted the photographer in the bushes, they shouldn't have chewed with their mouths open, they shouldn't have worn the strapless top, they shouldn't have yawned, they shouldn't have laughed: so unattractive, the candid denture.*

*GT: Fotografja bi morali opaziti v grmovju, ne bi smeli žvečiti z odprtimi usti, ne bi smeli nositi vrha brez naramnic, ne bi smeli zehati, ne bi se smejali: tako neprivilačno, odkrito proteza.*

*T.eu: Morali bi opazil fotograf v grmovju, ki jih ne bi smeli žvečiti z odprtimi usti, ki jih ne bi smeli nositi brez naramnic vrhu, ne smejo imeti yawned, ne bi smeli smejati: tako neprivilačna, iskren Denture.*

*Ctrl: Morali bi opaziti fotografja v grmovju, ne bi smeli žvečiti z odprtimi usti, ne bi smeli nositi majice brez naramnic, ne bi smeli zehati, ne bi se smeli smejati: kako neprivilačno je, če se vidi zobna proteza.*

This sentence contains comma splices – that is, independent clauses joined by commas, the last one being an elliptical clause. The solution provided by Translator.eu does not make sense in Slovene, which proves that the complex sentence has not been translated as a unit. Google Translate’s solution is better, yet still deficient. In the elliptical clause “so unattractive, the candid denture,” the word order is unusual because of the post-modification, and the MT systems were unable to identify the relation between the words (the correct form of adjective would be “neprivlačna,” since “denture” is feminine in Slovene). In addition, the elliptical clause is translated literally, and although the meaning of individual words is conveyed correctly, such syntax is not acceptable in Slovene. In the control version we therefore opted for an if-clause and did not preserve the ellipsis: “kako neprivilačno je, če se vidi zobna proteza” (BT: “it’s so unattractive if the denture is showing”).

Other elements of the sentence also proved to be problematic for NMT, for instance, the modal verb “should” in the past tense: GT translated the clauses containing “should” correctly except for the fifth one, where “shouldn’t have laughed” is changed to “wouldn’t laugh” (instead of modality in the past present conditional is used). T.eu translated only one should-clause correctly, while two were translated as relative clauses, in one the wrong tense was used, and one is ungrammatical. Also, in both MT versions the noun “top” is

translated as “vrh,” signifying the top of a tree or a mountain rather than a piece of clothing, as in the original. Additionally, T.eu left words untranslated, used the wrong case of the noun (“photographer”), and made pronoun resolution errors (in the ellipsis the first adjective is feminine and the second masculine, although both refer to the same noun).

While the students did not encounter translation problems in other parts of the sentence, the elliptic clause “so unattractive, the candid denture” proved to be problematic for most of them. Two students translated it as a full clause, of which one opted for an if-clause, as in the control version. The other students preserved the ellipsis. The adjective “candid” was translated by most of the students as: “razkrita” or “odkrita” (BT: “revealed,” “disclosed”), which is one of the possible dictionary meanings, but does not collocate with the noun “proteza” (“denture”).

[8] Orig: *But snip, snip, away go mother and father, white ribbons of paper blown by the wind, with grandparents tossed out for good measure.*

GT: *Toda šcipajte, ostržite, mati in oče, bele trakove papirja, ki jih je pihal veter, in jih babice in dedki za dobro mero vrgli.*

T.eu: *Ampak snip, snip, proč Pojdi mati in oče, beli trakovi papirja prepihani z vetrom, s starimi starši vrgla ven za dober ukrep.*

Ctrl: *Ampak škrt, škrt, že odideta mama in oče, veter odnese bele trakove papirja, za povrh se znebiš še starih staršev.*

Both the above machine translations are grammatically and semantically problematic: in fact, they fail to make sense. Although the quality of the students’ translations is much higher, some errors and deficiencies appear there as well. While both NMTs failed to acknowledge that “snip” is an onomatopoeic noun indicating the sound of snipping (cutting with scissors), rather than a verb (GT translated it as a verb, T.eu left it untranslated), all students correctly translated it as an onomatopoeic noun; three found a good solution in Slovene (“škrt, ck, resk”), the other solutions were inappropriate or stylistically marked (“rez, šip, štric”).

In the case of the elliptical clause “white ribbons of paper blown by the wind,” the Slovene language prefers a different word order in which “wind” would come first, as well as the active instead of the passive voice (Ctrl: “veter odnese bele trakove papirja”; BT: “the wind carries away white ribbons of paper”). While all students opted for the active form, five left the word order unchanged, leaving “white ribbons” at the beginning of the sentence, which shifts the emphasis to the wind.

In the phrase “with grandparents tossed out for good measure,” Slovene also prefers the active voice (Ctrl: “za povrhu se znebiš še starih staršev”; BT: “for good measure you get rid of the grandparents”). Five students changed the passive form into active, one preserved it, while two opted for an adjective (“zavrženi stari starši,” BT: “discarded grandparents”), which tweaks the meaning in Slovene. Additionally, two students omitted the phrase “for good measure.”

[9] Orig: *Farewell crumbling turrets of historic interest, farewell icebergs and war monuments, all those young stone men with eyes upturned, and risky voyages teeming with germs, and dubious hotels, and doorways opening both in and out.*

GT: *Oprostitvene propadajoče kupole zgodovinskega zanimanja, poslovilne ledene gore in vojni spomeniki /.../*

T.eu: *Zbogom razpada turrets zgodovinskega pomena, zbogom ledenih gorah in vojnih spomenikov /.../*

Ctrl: *Zbogom razpadajoči stolpiči zgodovinskega interesa, zbogom ledeniki in vojni spomeniki /.../*

As in example [8], both MT systems failed to provide a sensible translation of this sentence, which corroborates recent findings by Tezcan, Daems and Macken (2019) that NMT performance decreases with sentence length and complexity. Among other things, GT translated “farewell” in both cases as an adjective (but as two synonyms) defining the subsequent noun, while T.eu translated it correctly but failed to match the words with the appropriate case suffixes. Also, it left the noun “turrets” untranslated. The students generally had no problems translating this sentence, which might suggest that for human translators, long sentences might be easier to translate than short ones, where we need to fill in the blanks.

### 5.3 Pronominal Anaphora Beyond the Sentence

The following cases demonstrate that both NMTs have problems with translation of pronouns that replace a referent in the previous sentence. This proves that they cannot consider the context beyond the sentence, which is true for MT in general. This communicative example shows a human translator’s advantage over MT: while the latter focuses on one sentence at a time, the former, in principle, approaches each sentence with the knowledge and comprehension of the whole text.

[10] Orig: *I don’t mean I’m putting it together; no, I’m taking it apart.*

GT: *Ne mislim, da ga sestavljam; ne, razločim ga.*

T.eu: *Ne mislim, da sem ga dal skupaj; ne, to bom vzel narazen.*

Ctrl: *S tem ne mislim, da jo sestavljam; ne, razstavljam jo.*

The pronoun “it” refers to “my own life story” from the previous sentence. The noun “story” (“zgodba”) is feminine in Slovene, yet both NMTs missed this and used the masculine pronoun. Additionally, T.eu replaced the second person pronoun “it” with the demonstrative pronoun “this,” which puts emphasis on the subject that does not exist in the original sentence. The translation of multiword verbs appearing in this sentence is analyzed in section 5.5.

[11] Orig: *Into the ground with you, my tender fur-brained cats and dogs, and horses and mice as well: I adored you, dozens of you, but what were your names?*

GT: *V zemljo s tabo, moje nežne mačke in psi s kožuhi, pa tudi konji in miši. Oboževal sem te, na desetine vas, ampak kako so bila vaša imena?*

Ctrl: *Pod zemljo z vami, moje ljube kosmatoglave mačke in psi, pa tudi konji in miši: oboževala sem vas, bilo vas je na ducate, ampak kako vam je že bilo ime?*

In this sentence, the pronoun “you” is part of an apostrophe in cataphoric use. Even within the sentence GT failed to offer the right pronoun: the first and the second “you” are incorrectly

translated as singular pronouns, while the last one is plural, probably owing to “dozens of,” which the software recognized as a plural marker. The T.eu solution is even weaker and is therefore omitted here. The students have, of course, translated the pronouns correctly in both [10] and [11], but encountered problems with the phrase “dozens of you.” All translated it literally (“na ducate vas”), which sounds awkward in Slovene and is thus stylistically marked. In the control, we opted for a verbal clause (BT: “there were dozens of you”).

## 5.4 Present Participle

In the following examples the translation of the participle was identified as problematic for NMTs, owing to their inability to correctly identify the relation among parts of speech within or beyond the sentence.

[12] Orig: *What was it like to breathe so heavily, as if drugged, while rubbing up against strange leather coats in alleyways?*

GT: *Kako je bilo dihati tako močno, kot da bi se drogiral, medtem ko se v uličicah drgnil ob čudne usnjene plašče?*

Ctrl: *Kako je že bilo takrat, ko sem globoko dihala, kot bi bila omamljena, medtem ko sem se drgnila ob tuje usnjene plašče v stranskih ulicah?*

GT produced a meaningful translation of the first part of the sentence, while failing to do this in the part with the participle (“rubbing”) – the system could not identify the referent to which it relates. Considering the ability of the English language to be ambiguous, combined with Atwood’s careful wording of the question in sentence [12], this is not a surprise; it is, however, a serious translation challenge, since Slovene is an inflectional language, and the inflections frequently prevent ambiguity in the sense the latter can be implemented in English. While “rubbing” in GT is changed into a finite verb form, which is a common way to translate participles into Slovene (since they are less frequent than in English), the verb should be reflexive in this case (“drgniti se” instead of “drgniti”). T.eu’s translation fails to meet the standards of coherence and meaningfulness and is therefore omitted here.

The problems the students encountered here were mostly stylistic in nature: six students translated the first part as “Kako je že bilo globoko/težko dihati ...,” which is stylistically awkward, since Slovene prefers subordinate clauses to infinitives in such cases. The other two students used a similar temporal clause as the control. The part “as if drugged” is also translated as a subordinate clause in the control version. Four students opted for the same solution, while the other four chose “kot na drogah” or “kot omamljena,” which is closer to the original in structure but weaker in style.

In example [13], the participle connects the extended reporting clause to reported speech, and this proved a challenge to both software tools.

[13] Orig: *So that’s what she looked like, we say, connecting the snapshot to the year of the torrid affair.*

GT: *Tako je izgledala, pravimo, da je posnetek povezoval z letom burne afere.*

T.eu: *Torej, to je tisto, kar je izgledala, smo rekli, povezuje posnetek v leto vroče afere.*

Ctrl: *Takšna je torej bila videti, rečemo in povežemo fotografijo z letom žgoče afere.*

Both machine translations are grammatically inaccurate, T.eu's product is even beneath the threshold of sensible syntax. GT's reported clause is acceptable, while T.eu wrongly put it in the past tense. Again, however, the software failed to figure out the relation between the clause containing reported speech and the reporting clause with the participle "connecting"; GT translated the latter as a subordinate clause (BT: "that he connected the snapshot with the ..."), while T.eu translated it as a separate full clause with a missing subject (BT: "connects the snapshot with ..."). Both tools obviously translated the sentence clause per clause and were unable to consider it as a unit. On the other hand, the students' solutions for this sentence are generally fine. As in the control, they all opted for turning the non-finite structure into a finite one (either a coordinated or a subordinated clause; both solutions are possible) and preserving the link to the previous clause. Two students, however, wrongly used the imperfect verb form "connect" ("povezovati") instead of the perfect one ("povezati").

## 5.5 Multiword Verbs

Multiword verbs are another category that proved to be more difficult for neural machine translation than for the students.

[14] Orig: *I'm working on my own life story.*

GT and T.eu: *Delam na svoji življenjski zgodbi.*

Ctrl: *Ukvarjam se z lastno življenjsko zgodbo.*

Both software tools produced semantically and grammatically correct translations of the multiword verb, but the formulation is not equivalent in terms of formality level. Surprisingly, all but one student provided the same translation of "to work on" as the MTs. While the meaning is not radically changed, the register is lowered in this solution. In the control, we therefore replaced "delati" with "ukvarjati se" ("deal with"), a solution that was also provided by one of the students. Additionally, we replaced the pronoun "svojo" ("my") with "lastno" ("own") for emphasis, which is closer to the original ("my own").

The pair of antithetic multiword verbs from the example [10] ("put together, take apart") was translated literally by T.eu: "put together" as "dati skupaj" and "take apart" as "vzeti narazen," which is roughly correct in both cases but stylistically poor. GT found a better solution for "put together" ("sestavljati"), but the translation of "take apart" is semantically incorrect ("razločiti"). The students had no problems translating these two verbs.

## 5.6 Other Phraseological Units

Like multiword verbs above, idioms are "not predictable from the literal meaning of their parts" (Hladnik 2017, 25) and thus represent a similar obstacle for machine translation, as will be shown in the examples below.

- [15] Orig: *Maybe we just want to be in charge, of the life, no matter who lived it.*  
 GT: *Mogoče bi radi samo vodili življenje, ne glede na to, kdo ga je živel.*  
 T.eu: *Mogoče bi samo rada bila glavna, življenja, ne glede na to, kdo je živel.*  
 Ctrl: *Morda preprosto želimo imeti nadzor nad življenjem, ne glede na to, kdo ga je živel.*

The GT version is grammatically accurate but contains a semantic shift. The idiom “to be in charge” can indeed be translated as “voditi” (“lead”), but not in this case, where it means “have control over,” or in Slovene “imeti nadzor.”<sup>8</sup> The T.eu’s version (“biti glaven”) is one of the possible translations of the idiom, yet not acceptable in this case. However, the T.eu’s translation of the sentence as a whole is ungrammatical and meaningless. As for the students’ translations, only one student chose the same solution as appears in the control version; the others provided different versions of “be in charge of,” all of which are semantically correct but stylistically weak.

- [16] Orig: *What was all the fuss about?*  
 GT: *Zakaj je bila vsa prepir?*  
 T.eu: *Kaj je bilo ves ta hrup?*  
 Ctrl: *Čemú je bil ves ta cirkus?*

The situation is not very different for machine translation if the expression is colloquial, as is the case here. Not only did both MT engines experience problems finding the right meaning of “fuss” (GT translated it as “fight” or “argument,” T.eu as “noise”), but the sentences are also ungrammatical. The students had no difficulties understanding this sentence but provided different translations of “fuss” (such as “pomp,” “direndaj,” “cirkus,” “razburjanje”), all of which are semantically correct.

- [17] Orig: *I’m coming unstuck from scrapbooks, from albums, from diaries and journals, from space, from time.*  
 GT: *Odhajam iz beležk, albumov, dnevnikov, dnevnikov, veselja, časa.*  
 T.eu: *Prihajam odlijepio iz scrapbooks, od albumov, iz dnevnikov in revij, iz veselja, od časa.*  
 Ctrl: *Osvobajam se spominskih knjig, albumov, dnevnikov in zapiskov, prostora, časa.*

Here both machine translations are semantically inaccurate (the one by T.eu is also ungrammatical), while the students understood correctly the phraseological unit “to come unstuck,” although some of their solutions failed to meet the standards of stylistic appropriacy. Only two students provided an acceptable solution (“osvobajati se”; BT: “break free from”), which appears also in the control.

<sup>8</sup> The seemingly trivial idiom is often not easy to translate into Slovene, which proved the case when the same students translated Alice Munro’s short story “Tricks” as a take-home assignment. The suggestions for the translation of the idiom in the original sentence – “He would not leave that brother in charge for very long.” – varied from *odgovornega* (responsible), to *glavnega* (in charge) and “ne bi mu prepustil odgovornosti” (BT: he would not be left with the responsibility), which are semantically correct but stylistically marked in this context. Four students, however, chose the solution that appears in the published translation and seems the most appropriate, considering the context: “Brata ne bi pustil dolgo samega.” (BT: He would not leave the brother alone for long).

## 5.7 Part of Speech Recognition

Recognizing the right part of speech proved to be problematic for MTs when this was not obvious from the suffix, i.e., when the form itself could fit more than one word class (and when syntax and/or context are the decisive factors).

- [18] Orig: *No more choices for the people in them – pick this one, dump that one.*  
GT: *Ni več izbire za ljudi, ki so v njih – izberi to, odloži tisto.*  
T.eu: *Nič več izbire za ljudi v njih-pick ta, smetišče, da je eden.*  
Ctrl: *Nič več izbire za ljudi na njih – vzemi to, zavrzi tisto.*

GT recognized that “dump” in [18] is a verb in imperative mode and translated it as “odloži,” while T.eu identified it as a noun and translated it as “smetišče,” which makes the translation nonsensical. Additionally, in the T.eu version the word “pick” is left untranslated, while “that one” is translated as “that it is one.”

- [19] Orig: *goodbye well-thumbed tears and scabby knees, and sadness worn at the edges*  
GT: *[...] ter žalost, ki se je nosila na robovih*  
T.eu: *[...] in žalosti nosili na robovih*  
Ctrl: *[...] in na robovih oguljena žalost.*

In example [19], “worn” is an adjective, yet both MT engines translated it as the past tense of the verb “wear.” Again, the GT translation is grammatically correct, while the T.eu one fails to make sense. The students had no problems with part of speech recognition.

## 5.8 Atwood’s Neologisms and Personal Style

Atwood’s linguistic creativity is, unsurprisingly, a considerable obstacle for human translators, let alone MTs (see, for example, Somacarrera 2013 on the challenges of translating Atwood into Spanish). Both tools used in this research failed in translating words coined by Atwood or used in her text in an unusual way, for instance, the adjective “well-thumbed tears,” which could at least be classified as an epithet. While T.eu leaves “well-thumbed” untranslated, GT translates it as “dobro nabrekli solze” (BT: “well-swollen tears”), which changes the meaning of the adjective. Four students provided acceptable solutions, similar to that of the control (“obrabljene solze”), while four students provided semantically inaccurate translations. Another neologism can be found in example [20]:

- [20] Orig: *The livers of the lives in question had their chances, most of which they blew.*  
GT: *Jetrnice zadevnih življenj so imele svoje možnosti, med katerimi so večino raznesle.*  
T.eu: *Jetra iz življenja so imeli svoje možnosti, večina jih je razneslo.*  
Ctrl: *Tisti, ki so omenjena življenja živeli /.../*

In this context, “livers” is used in an original way: it is not the plural of “liver,” an organ of the body, but rather means “people who live certain lives.” T.eu opts for the organ (“jetra”), while GT translates it as “jetrnice” (plural of “jetrnica,” a sausage stuffed with rice and liver),

which is – despite the polyptoton – a radical change of meaning. While the first part of GT’s translation is at least grammatically correct (“jetrnice so imele svoje možnosti”), the one produced by T.eu is not: “Jetra ... so imeli,” instead of: “... so imela.” Both software tools also failed to grasp the figurative meaning of the verb “to blow” in this sentence, which is “to fail”; T.eu translated it as “raznesti,” meaning “to erupt” or “to explode,” while GT used the verb “raznesti” in a rarer use, as in “raznesti možnosti” (BT: “to spread the possibilities”). The students provided the correct meaning of “livers” in Slovene; however, three of them used the literal translation “živeči,” which is rarely used in Slovene in this sense. Slovene prefers subordinate clauses in such cases; that is why we opted for a relative clause in the control (BT: “those who lived the lives in question”). Five students also provided such a solution.

The ending of Atwood’s “Life Stories” proved to be a major translation challenge, particularly because of its stylistic features. The sentence “I was born” is reduced by one word in each of the two subsequent steps, and what remains is a non-verbal clause, the personal pronoun “I,” which is the shortest possible word in English, yet, paradoxically, the central and most crucial person in any life story – the life owner:

[21] Orig:	<i>I was born.</i>	GT:	<i>Rodil sem se.</i>	T.eu:	<i>Rodil sem se.</i>	Ctrl:	<i>Jaz sem se rodila.</i>
	<i>I was.</i>		<i>Jaz sem bil.</i>		<i>bil sem.</i>		<i>Jaz sem bila.</i>
	<i>I.</i>		<i>JAZ.</i>		<i>I.</i>		<i>Jaz.</i>

Slovene allows several translation options for “I was born”: “Jaz sem se rodila.” (exposed pronoun, marked use, usually for emphasis; used by two students and the control version); “Rodila sem se.” (stylistically the strongest option; used by one student and both NMTs, although the NMTs used the masculine form of the past participle); “Rojena sem bila.” (passive form, stylistically marked; used by two students); “Jaz sem bila rojena” (exposed pronoun, passive form; used by two students); “Bila sem rojena” (passive form, stylistically the weakest solution; chosen by one student). Since Slovene as an inflectional language does not favour the exposed pronoun (the verb or past participle suffix reveals the grammatical person, number and gender), the inclination in neutral context would be preferable: “Rodila sem se. / Bila sem. / Jaz.” However, this translation loses the aesthetically pleasing initial anaphora as well as the decreasing, possibly anticlimactic parallelism that on the textual level depicts the conclusion of one’s life. Opting for the neutral translations without the exposed pronoun pushes the translator into an inevitable loss of one or more prominent stylistic features. As shown, the two translation tools were unable to approach these features as a unit; GT used a combination of the neutral version in the first sentence and the exposed pronoun version in the second, while T.eu used the neutral translation in both sentences but left the pronoun “I” untranslated. The two students who began the first sentence with “*Jaz sem se rodila*” were able to keep the anaphora in the second step and finish with the pronoun I: “*Jaz sem bila. / Jaz.*” This is also the solution in the control version.

## 6 Concluding Remarks

The analysis of MTs’ translations of Atwood’s “Life Stories” has shown that their performance was satisfactory only in the case of short, non-idiomatic sentences with simple structure,

while their performance decreased with the length and complexity of the sentences. Atwood's personal style comprising the unconventional use of language (e.g., non-standard or elliptical syntax) and her neologisms proved to be a particularly insurmountable obstacle for machine translation; these were also a major challenge to the student translators. Overall, the NMTs manifested shortcomings previously identified in studies of literary MT such as accuracy and robustness (e.g., dealing with rare words like “snip” and “turret” by failing to translate them; switching between masculine and feminine gender; co-reference changes; mistranslation of phrasal verbs and idioms, disambiguation, etc.), while the problems identified in the student translations were mostly not grammatical or semantic but stylistic in nature. GT, however, performed better than T.eu in producing meaningful, coherent sentences.

NMT systems will undoubtedly continue to improve, owing to their ability to learn. However, the leading principle in NMT development, in the sense of “teaching” it to choose the more appropriate (i.e., the more frequently or widely used) lexical item when there are several, is intrinsically in conflict with the stylistic aspects of literature, where the more frequent or widely used lexical choices are not favoured, or where the most precious aspects of a text are not at the lexical level.

The extent to which NMT systems can be useful in translating literature is highly dependent on the complexity of the literary work. Atwood's unique personal style, which needs to be reproduced in translation, is particularly salient in her shorter works (mini fiction, prose poetry). It depends on avoiding the conventional use of language and sometimes on experimentation, which is different from books to which the readers are attracted either by the plot or by a writing favouring simple structures (e.g., Hemingway, J. K. Rowling). We can thus expect that despite further improvements in output quality, translating authors such as Atwood will continue to present a major challenge for machine translation; the same is true for inexperienced human translators, such as students.

The existing studies of NMT for the English-to-Slovene combination are scarce and merely quantitative. This study allows the reader to gain insight into individual cases and therefore contributes substantially to understanding the nature of machine translated output, rather than merely demonstrating the success rate of the machine translation. By providing a detailed translation analysis of the selected work by Atwood, we bring attention to the difficulty of literary translation in cases of stylistically challenging prose, and the bridging of language gaps between Atwood's text(s) and the Slovene reader(s).

## References

- Atwood, Margaret. (2006) 2007. “Life Stories.” In *The Tent*, 3–5. New York: Anchor.
- Bonner, Anne. 2019. “What is Deep Learning and How Does it Work?” *Towards Data Science*, September 7, 2019. <https://towardsdatascience.com/what-is-deep-learning-and-how-does-it-work-f7d02aa9d477>.
- Borg, Claudine. 2016. “A literary translation in the making: an in-depth investigation into the process of a literary translation from French into Maltese.” PhD diss., Aston University. [https://publications.aston.ac.uk/id/eprint/30855/1/Borg\\_C.\\_2017.pdf](https://publications.aston.ac.uk/id/eprint/30855/1/Borg_C._2017.pdf).
- Gadpaille, Michelle. 2018. “Sci-Fi, Cli-Fi or Speculative Fiction: Genre and Discourse in Margaret Atwood's ‘Three Novels I Won't Write Soon.’” *ELOPE* 15 (1): 17–28. <https://doi.org/10.4312/elope.15.1.17-28>.

- . 2014. "Thematics and its Aftermath: A Meditation on Atwood's Survival." *Primerjalna književnost* 37 (3): 165–77.
- Hladnik, Marko. 2017. "Variability in the Syntax of Idioms." *ELOPE* 14 (2): 25–37. <https://doi.org/10.4312/elope.14.2.25-37>.
- Kravariti, Alexandra. 2018. "Machine translation: NMT translates literature with 25% flawless rate." *Translate Plus*, February 9, 2018. <https://www.translateplus.com/blog/machine-translation-nmt-translates-literature-25-flawless-rate/>.
- Kuzman, Taja, Špela Vintar, and Mihael Arčan. 2019. "Neural Machine Translation of Literary Texts from English to Slovene." In *The Qualities of Literary Machine Translation*, edited by James Hadley et al., 1–9, Dublin: Machine Translation Summit XVII. <https://www.aclweb.org/anthology/volumes/W19-73/>.
- Leech, Geoffrey, and Mick Short. 2007. *Style in Fiction*. Harlow: Pearson.
- Matusov, Evgeny. 2019. "The Challenges of Using Neural Machine Translation for Literature." In *The Qualities of Literary Machine Translation*, edited by James Hadley et al., 10–19, Dublin: Machine Translation Summit XVII. <https://www.aclweb.org/anthology/volumes/W19-73/>.
- Nischik, Reingard. 2006. "Margaret Atwood's Short Stories and Shorter Fictions." In *The Cambridge Companion to Margaret Atwood*, edited by Coral Ann Howells, 145–60. Cambridge: Cambridge University Press.
- Popović, Maja. 2018. "Error Classification and Analysis for Machine Translation Quality Assessment." In *Translation Quality Assessment, Machine Translation: Technologies and Applications 1*, edited by Joss Moorkens et al., 129–58. Cham: Springer. [https://doi.org/10.1007/978-3-319-91241-7\\_7](https://doi.org/10.1007/978-3-319-91241-7_7).
- Pregelj, Barbara. 2019. "Nekaj podatkov za oris profila slovenskih prevajalk in prevajalcev." In *Bela knjiga o prevajanju* 2018, 65–77. Ljubljana. [http://www.belaknjigaoprevajanju.si/files/2019/03/Bela\\_knjiga\\_cela.pdf](http://www.belaknjigaoprevajanju.si/files/2019/03/Bela_knjiga_cela.pdf).
- Schwartz, Ros, and Nicolas de Lange. 2006. "A dialogue: on a translator's interventions." In *The Translator as Writer*, edited by Susan Bassnett and Peter Bush, 9–19. London, New York: Continuum.
- Scott, Clive. 2018. *The Work of Literary Translation*. Cambridge: Cambridge University Press.
- Slettedahl Macpherson, Heidi. 2010. *The Cambridge Introduction to Margaret Atwood*. Cambridge: Cambridge University Press.
- Somacarrera, Pilar. 2013. "Contextual and Institutional Coordinates of the Transference of Anglo-Canadian Literature into Spain." In *Made in Canada, Read in Spain: Essays on the Translation and Circulation of English-Canadian Literature*, edited by Pilar Somacarrera, 21–53. London: Versita.
- Tezcan, Arda, Joke Daems, and Lieve Macken. 2019. "When a 'sport' is a person and other issues for NMT of novels." In *The Qualities of Literary Machine Translation*, edited by James Hadley et al., 40–49, Dublin: Machine Translation Summit XVII. <https://www.aclweb.org/anthology/volumes/W19-73/>.
- Toral, Antonio, and Andy Way. 2014. "Is Machine Translation Ready for Literature?" In *Proceedings of translating and the computer 36*, edited by João Esteves-Ferreira et al., 174–76. Geneva: Editions Tradulex. <http://www.tradulex.com/varia/TC36-london2014.pdf>.
- . 2018. "What Level of Quality Can Neural Machine Translation Attain on Literary Text?" In *Translation Quality Assessment. Machine Translation: Technologies and Applications*, vol. 1, edited by Joss Moorkens et al., 263–87. Cham: Springer.
- Vintar, Špela. 2018. "Terminology Translation Accuracy in Statistical versus Neural MT: An Evaluation for the English-Slovene Language Pair." In *Proceedings of the Eleventh International Conference on Language Resources and Evaluation (LREC 2018)*, 34–37. European Language Resources Association. Miyazaki, Japan.
- . 2013. "Uvodnik: o rojstvu korpusa SPOOK in njegovih prvih sadovih." In *Slovenski prevodi skozi korpusno prizmo*, edited by Špela Vintar, 6–13. Ljubljana: Ljubljana University Press.
- Way, Andy. 2018. "Quality Expectations of Machine Translation." In *Translation Quality Assessment. Machine Translation: Technologies and Applications 1*, edited by Joss Moorkens et al., 159–78. Cham: Springer. [https://doi.org/10.1007/978-3-319-91241-7\\_7](https://doi.org/10.1007/978-3-319-91241-7_7).
- Wu, Yonghui, Mike Schuster et al. 2016. "Google's neural machine translation system: Bridging the gap between human and machine translation." Ithaca: Cornell University. <https://arxiv.org/abs>.