

University of Groningen

To be or not to be

Guerberof-Arenas, Ana; Toral, Antonio

Published in:
Target

DOI:
[10.1075/target.22134.gue](https://doi.org/10.1075/target.22134.gue)

IMPORTANT NOTE: You are advised to consult the publisher's version (publisher's PDF) if you wish to cite from it. Please check the document version below.

Document Version
Publisher's PDF, also known as Version of record

Publication date:
2024

[Link to publication in University of Groningen/UMCG research database](#)

Citation for published version (APA):
Guerberof-Arenas, A., & Toral, A. (2024). To be or not to be: A translation reception study of a literary text translated into Dutch and Catalan using machine translation. *Target*, 36(2), 215-244.
<https://doi.org/10.1075/target.22134.gue>

Copyright

Other than for strictly personal use, it is not permitted to download or to forward/distribute the text or part of it without the consent of the author(s) and/or copyright holder(s), unless the work is under an open content license (like Creative Commons).

The publication may also be distributed here under the terms of Article 25fa of the Dutch Copyright Act, indicated by the "Taverne" license. More information can be found on the University of Groningen website: <https://www.rug.nl/library/open-access/self-archiving-pure/taverne-amendment>.

Take-down policy

If you believe that this document breaches copyright please contact us providing details, and we will remove access to the work immediately and investigate your claim.

Downloaded from the University of Groningen/UMCG research database (Pure): <http://www.rug.nl/research/portal>. For technical reasons the number of authors shown on this cover page is limited to 10 maximum.

To be or not to be

A translation reception study of a literary text translated into Dutch and Catalan using machine translation

Ana Guerberof-Arenas and Antonio Toral

University of Groningen

This article presents the results of a study focusing on the reception of a fictional story by Kurt Vonnegut translated from English into Catalan and Dutch in three conditions: machine translated, post-edited, and human translated. Participants ($n = 223$) rated the three conditions using three scales: narrative engagement, enjoyment, and translation reception. The results show that human translation had higher engagement, enjoyment, and translation reception in Catalan, compared to the post-edited and machine-translated translations. However, Dutch readers scored the post-edited translation higher than the human and machine translation, and the highest engagement and enjoyment scores were reported for the original English version. We hypothesize that when reading a fictional story in translation, not only are the condition and the quality of the translation key to understanding its reception, but also the participants' reading patterns, reading language, and, potentially, the status of the source language in their own societies.

Keywords: literary translation, translation reception, machine translation, post-editing, narrative engagement, enjoyment, comprehension, creativity

1. Introduction

As part of a larger research project¹ that explores creativity and the use of machine translation (MT) in the production of translated literary texts, this article details

1. Creativity and Narrative Engagement of Literary Texts Translated by Translators and Neural Machine Translation (CREAMT); see <https://cordis.europa.eu/project/id/890697>.

the results of a study exploring translation reception when mediated by technology, in this case neural MT.

Translation reception is an under-explored area in literary translation and further research is needed (Walker 2021) to understand translation from a wider perspective (including its impact on the user), but also to understand how new tools in the translation process might affect the reading experience of the final users of translation. Although there is some published work that considers the translation of literary texts versus original texts (Kruger 2013; Walker 2020), the impact MT has on the literary translation process and, subsequently, on readers has not been thoroughly explored yet.

In our previous work (Guerberof-Arenas and Toral 2020), we selected a short story, “Murder in the Mall” (Nuland 1995), to compare different translation modalities using a narrative engagement scale (Busselle and Bilandzic 2009), enjoyment questions (Hakemulder 2004), and a translation reception questionnaire that we designed. The methodology used was partially borrowed from Mangen and Kuiken (2014), who compared engagement between online and on-paper reading. Our results showed that professional translation without MT, defined here as human translation (HT), scored higher for creativity than post-editing (PE) and MT, and that HT scored higher in narrative engagement and translation reception and slightly lower than PE in enjoyment, while the results for MT were significantly lower than for HT and PE. We hypothesized that this difference between HT and PE is due to a higher creativity score in HT and, therefore, that a different source text with more creative challenges for translators would increase the gap between the conditions if reception were to be considered.

With this hypothesis in mind, we expand the cohort of participants in this study to 223, include another language combination, English to Dutch, and, more importantly, choose a text in which, unlike our pilot experiment, style prevails over action.

2. Related work

As mentioned in Section 1, there is existing work that considers the reception of translated literary texts using eye-tracking (in combination with other methods). Kruger (2013) analyzes the reading and processing of domesticating and foreignizing translation strategies by readers of translated children’s books, and looks at dwell time, fixation count, first-fixation duration, and glance count for areas of interest that reflect these strategies. Walker (2020) considers the salient features

of literature (used for foregrounding)² to determine whether readers' cognitive effort is equivalent between source text and target text, using the measures of gaze duration and total fixation duration. He finds that these salient features result in "greater diversity in the levels of visual attention (cognitive effort) exhibited by a range of different readers" (ibid., 312). Although these studies do not deal directly with MT or PE, they signal the relevance of considering salient features of literary texts in combination with users' responses.

The Digital Opinions on Translated Literature project, DIOPTRA-L (Kotze et al. 2021), focuses on the salience of the fact of translation in readers' opinions: What are the main concepts, and emotional and evaluative parameters used by readers who post their opinions on translated books on Goodreads? Kotze et al. (2021) find that the fact of translation is not particularly salient for these readers and, when translation is mentioned, it tends to accompany a poor review. The authors speculate that translation might act as a 'scapegoat' for readers' preferences. Although it varies by language combination, translation is more salient for those who read a translation in English than in other languages.

In Guerberof-Arenas and Toral (2020), narrative engagement, enjoyment, and translation reception are measured in a cohort of eighty-eight Catalan readers, and differences are found between the translation conditions. HT scores higher in narrative engagement and translation reception, and slightly lower than PE in enjoyment. However, there are no statistically significant differences between HT and PE for any of these variables. MT, unsurprisingly, has the lowest engagement, enjoyment, and translation reception scores. However, certain categories in the scale, such as attentional focus, emotional engagement, and narrative presence, do not show statistically significant differences across the conditions, which could be related to the nature of the story readers were presented with, where action was more relevant than style. It is also clear from the results that readers enjoyed PE marginally more or as much as HT, even if they appreciate that HT was a better translation.

Colman et al. (2022) report the preliminary results of an eye-tracking study in which participants read a full novel in Dutch, alternating between MT and the published translation. The authors compare the reading process of participants reading both versions and find that the MT version has an increased number and duration of fixations, as well as increased gaze durations and shorter saccade amplitudes. The study focuses on overall reading effort and error typologies; however, it does not analyze any specific creative traits or the impact of post-edited texts on the reader.

2. In literary studies, foregrounding refers to the effect of certain language features that serve to change the attention of the reader.

In a recent publication, Whyatt, Lehka-Paul, and Tomczak (2023) carry out a translation reception study for English–Polish translation with twenty participants who read two translations of a product description text, one being high quality and the other low quality. They use eye-tracking measures to correlate translation effort and translation quality with reading effort. They find that high-quality translations are easier to read but they do not find any straightforward relationship between translation effort and reading effort. This exploratory study is a part of larger research project called “Reading and Reception of Mediated (Translated) Text” (Read Me), which investigates the relationship between the translation process and the reading experience.

Stasiomiti and Sosoni (2022) examine the reception of marketing and literary texts in the English to Greek language combination with 100 readers. They find that the PE creative texts do not seem to differ from the HT texts in the number of errors and in the feedback from the readers. However, the translations were created by translation students, some of whom were active translators, but no professional literary translators were engaged in the experiment.

Research into the reception of audiovisual translation has been gaining prominence since the work of D’Ydewalle and colleagues in the 1980s (D’Ydewalle 1984; D’Ydewalle, Van Rensbergen, and Pollet 1987; D’Ydewalle and Van Rensbergen 1989). Kruger (2018) and Orrego-Carmona (2018) offer detailed overviews of how viewers behave when reading subtitles (e.g., analyzing standard, intralingual, and reversed subtitling). These overviews also focus on topics like the presentation of subtitles, experience with humour, accessibility studies, audio-description, and dubbing versus subtitling. Relevant to the current article, Ortiz Boix (2016) examines two conditions (HT and PE) in the translation of wildlife documentaries. The results of a panel of experts and fifty-six end-users establishes no significant differences between the two conditions. Hu, O’Brien, and Kenny (2021) investigate the reception of subtitles in massive open online courses from English to Chinese. They find that viewers show higher reception scores in the PE condition, though the difference with MT is not statistically significant. Viewers appreciate the HT condition, but it does not generate the highest scores. The eye-tracking data does not clearly show whether one condition requires more cognitive effort than the other.

3. Overall reading experience: Methodology

The CREAMT project has two main axes. Results of the first, which focuses on creativity in translated texts, are reported in Guerberof-Arenas and Toral (2022).

In this article, we present the results of the second axis, which is driven by the following research questions:

- RQ₁: Do users reading translated material produced under different translation conditions have different reading experiences?
- RQ₂: Does the readers' experience vary between languages (in this case, Dutch and Catalan)?

To answer these questions, the following experimental setting was used.

3.1 Source text

The chosen literary text had to meet the following conditions. It had to:

1. have a higher creative potential than "Murder in the Mall" (used in Guerberof-Arenas and Toral [2020]; see Section 1),
2. not be present in the data used to train the NMT engines used in the study,
3. be engaging enough to measure reading experience in a wider audience,
4. be short enough to be read in approximately thirty minutes,
5. not be too outdated so different age groups could engage with it, and
6. not infringe copyright laws.

We selected a Kurt Vonnegut story from the Project Gutenberg Literary Archive Foundation.³ "2BR02B" was published in January 1962 in the digest magazine *Worlds of IF Science Fiction* and was later included in the collection *Bagombo Snuff Box* (Vonnegut 1999). This story is set in a futuristic world where death has been eradicated and people live forever; the only caveat is that for one person born, another one must die. Vonnegut centres the story on several characters: a father to be, a doctor, a government official, a hospital orderly, and a painter. These characters exemplify different attitudes towards the world they are immersed in. In creating a new world, Vonnegut invents terms, roles, institutions, and expressions to describe it.

The text contains 123 paragraphs, 234 segments, and 2548 words. The text has a Flesch Reading Ease index of 79 and a Flesch-Kincaid Grade Level of 4.6, which indicate that the text is not an overly difficult text to read for an English speaker. To our knowledge, this story has not been translated and published in Catalan or Dutch.

3. See <https://www.gutenberg.org/files/21279/21279-h/21279-h.htm>.

3.2 Target texts

The short story was translated from English into Dutch and Catalan using three modalities: a human translation (HT), post-edited MT output (PE), and raw MT output. These languages were selected for two reasons: convenience of sample and availability of two MT engines already trained and tested at University of Groningen. The HT and PE versions were provided by four professional translators (two into Dutch and two into Catalan) who specialize in literary translation.⁴

To reduce the effect of individual translator style in the experiment (i.e., the risk that a reader would engage more with one translator's work because they preferred that translator's style) each professional translated and post-edited fifty percent of the text, and then the text was aggregated. Therefore, each of the two translation conditions, HT and PE, contained the aggregated translations of these two translators per language.

The translations were reviewed by five professional translators, who were asked to assess the translations' novelty (number of creative shifts) and acceptability (number of errors) in order to obtain an overall assessment of creativity in the translations. The reviewers were unanimous in ranking HT as an extremely good translation, MT as an extremely poor translation, and PE as neither a good nor a poor translation (Guerberof-Arenas and Toral 2022). The MT contained more errors than the other two conditions combined, and the PE also contained a strikingly higher number of errors than the HT (double the errors in Catalan and three times in Dutch), even if the same two translators worked on the HT and PE. Furthermore, the HT had the highest number of creative shifts, followed by the PE, and then the MT. The Dutch versions contained more errors than the Catalan versions, but a similar number of creative shifts.

The target texts presented to the participants in the current study were the translated texts that did not include the changes from reviewers.

3.3 Participants

The criteria for the inclusion of participants were that they were frequent readers of fiction, native Catalan and Dutch speakers, and eighteen years or older. The participants were offered ten euros compensation for their time. To recruit participants that fit the profile, forums for readers were targeted to avoid any bias

4. To recruit the translators, two databases were consulted: *Expertisecentrum Literair Vertalen* and *Associació d'criptors en llengua catalana*. Some translators recommended others who were contacted based on availability. The translators used in our previous experiment (Guerberof-Arenas and Toral 2020) were recruited as reviewers for Catalan.

towards translation scholars, practitioners, or students, but without necessarily excluding these readers.

The experiment was advertised on the following sites: Goodreads (Lectura en català Group,⁵ Fanatieke Nederlandse Lezers,⁶ Netherlands Flanders Group),⁷ Hebban,⁸ Relats en català (a web forum for readers and writers in Catalan),⁹ Facebook (Catalans in Ireland,¹⁰ Catalans in Holland,¹¹ Samenlezenisleuker,¹² Leesclub In de boekenkast),¹³ reading clubs organized by libraries in Catalonia¹⁴ and in the Netherlands such as Senia¹⁵ and Literaire Studentenvereniging Flanor.¹⁶ It was also distributed to writing schools such as Escola d'Esriptura de l'Ateneu Barcelonès, and to Dutch-speaking members of the European Association of Creative Writing Programmes. To obtain more participants, a group of students studying Catalan and Dutch philology were targeted through different professors at Universitat Oberta de Catalunya, Universitat Rovira I Virgili, and University of Groningen. Some participants received the link to our study from other participants through private messaging. The advertisements were posted at different intervals during the time the questionnaire was active, from 7 September 2021 to 21 January 2022.

3.4 Reading conditions

All readers were randomly assigned a translation condition. READINGA corresponded to the PE version, READINGB to the MT version, and READINGC to the HT version. Since English fiction is also often read in the original in the Netherlands (see, e.g., NOS 2022), we decided to add an additional condition READINGD, English condition, to the Dutch group. Therefore, in the end, each Catalan reader was assigned one of three conditions and each Dutch reader one

5. See <https://www.goodreads.com/group/show/61003-lectura-en-catal>.

6. See <https://www.goodreads.com/group/show/79675-fanatieke-nederlandse-lezers>.

7. See <https://www.goodreads.com/group/show/223-netherlands-flanders-group>.

8. See <https://www.hebban.nl/community>.

9. See <http://relatsencatala.cat/>.

10. See <https://www.facebook.com/groups/CatalansIrlanda>.

11. See <https://www.facebook.com/groups/203190096419372>.

12. See <https://www.facebook.com/groups/451488498379185>.

13. See <https://www.facebook.com/groups/1574870836130714>.

14. See <https://biblioteques.gencat.cat/>.

15. See <https://www.senia.nl/pages/Senia/Home>.

16. See <https://www.flanor.nl/en/home>.

of four. In this article, the name of the translation condition is used for ease of understanding.

3.5 Questionnaire

Two online questionnaires were distributed to participants using Qualtrics:¹⁷ one translated to Catalan and the other to Dutch. The participants first read the information brochure and consent form and if they decided to participate they were taken to the following sections, sequentially:

1. Demographics and reading habits: This section contained thirteen items on demographics and reading habits (e.g., “What genre do you usually read?” and “How often have you read in Catalan/Dutch in the last 12 months?”).
2. Comprehension questions: After reading the text, the participants answered ten multiple-choice questions to ensure basic comprehension of the story. The participants could only continue if five questions were answered correctly. Even though it would have been interesting to analyze participants’ comprehension (1 to 10) according to the condition, this section was intended to act as a filter for spurious responses as the questionnaire had been posted on a variety of online reading sites (see Section 3.3).
3. Narrative engagement: Participants were presented with a narrative engagement scale originally designed for media studies that contains twelve items (Busselle and Bilandzic 2009) and a seven-point Likert-type response option for each question. The section focused on four categories: narrative understanding (e.g., “At points, I had a hard time making sense of what was going on in the story”), attentional focus (e.g., “While reading, I found myself thinking about other things”), narrative presence (e.g., “The story created a new world, and then that world suddenly disappeared when the story ended”), and emotional engagement (e.g., “I felt sorry for some of the characters in the story”).
4. Visual imagery: We added three questions for mental imagery (e.g., “When I was reading the story I had an image of the main characters in mind”) borrowed from the story world absorption scale (Kuijpers et al. 2014) to complement the more literary items captured in the narrative engagement scale.
5. Enjoyment: The participants were asked to answer three questions related to enjoyment: “How much did you enjoy reading the text?”, “Do you think this text is an example of good literature?”, and “Would you recommend this text

17. The questionnaire and the anonymized data are available at <https://github.com/AnaGuerberof/CREAMT>.

to a friend?” These questions were borrowed from experimental research on the effects of foregrounding in relation to specific text qualities (Dixon et al. 1993; Hakemulder 2004).

6. Translation reception: To our knowledge, there is no existing scale to measure translation reception. Therefore, we devised a scale with eleven items and a seven-point Likert-type response option for each item (e.g., “How easy was the text to understand?”, “I thought the text was very well written”, and “I found words, sentences or paragraphs that were difficult to understand”).
7. Debriefing and payment questions: At the end of the questionnaire, the participants were debriefed about the nature of the research. Only then were they informed about the author and the nature of the text. Following this, they were asked how much money they would be willing to pay for this translation or the original piece and, if their condition was MT, to rate the quality of the engine.

The experimental workflow is summarized in Figure 1.

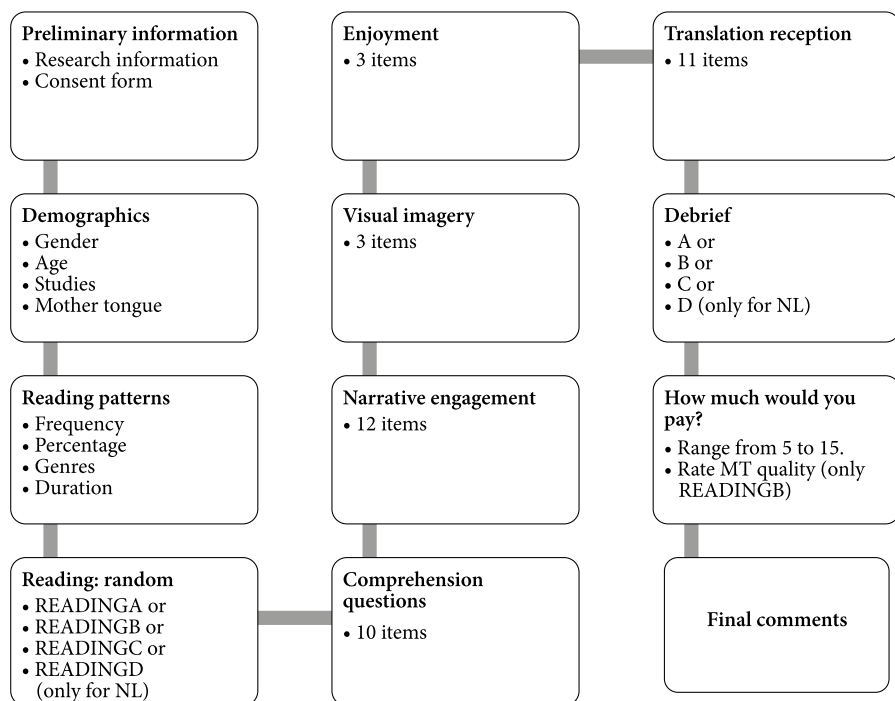


Figure 1. Experimental workflow

4. Analysis¹⁸

4.1 Participants

Figure 2 shows the distribution of participants according to their provenance ($n=223$).

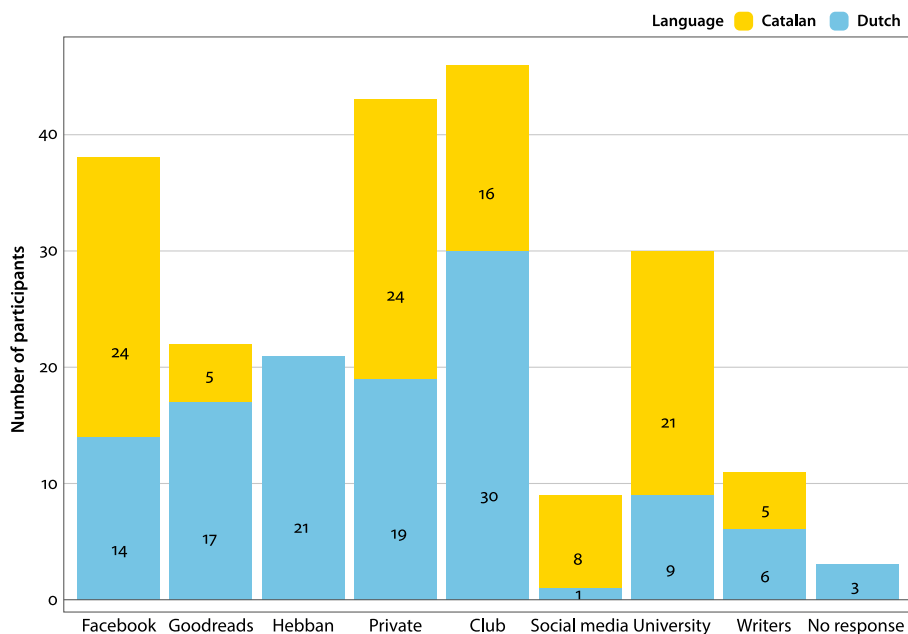


Figure 2. Distribution of participants according to recruitment avenue and mother tongue

The questionnaire was completed by 103 Catalan- and 120 Dutch-speaking participants. The latter were from the Netherlands and Belgium as we advertised on websites that cover both countries because books translated into Dutch are read in both countries. Table 1 presents the participants' demographic information.

18. All our statistical analyses are available at <https://github.com/AnaGuerberof/CREAMT>.

Table 1. Participants' demographic information

Gender	Categories				Total
	Female	Male	Non-binary	Prefer not to say	
	168	48	4	3	223
Age	18–34	35–54	55–74	75 or older	Total
	94	63	58	8	223
Education	Secondary school	University (not completed)	University degree	Professional	Total
	30	27	151	15	223
Mother tongue	Catalan	Catalan/Spanish	Dutch	Dutch/English	Total
	50	53	118	2	223
Profession	Language-related		Other	Unanswered	Total
	53		165	5	218

4.2 Similarities and differences between groups

Although the research advertisements were published on similar sites, some characteristics varied within the group of final participants. In both languages, most participants are women (68 out of 104 in Catalan and 100 out of 120 in Dutch) with a university degree (151 out of 223). However, the Dutch group has a higher percentage of participants with a university degree than the Catalan group (72.5% and 62%, respectively). Regarding age, the Catalan group has a higher percentage of participants in the age bracket 35–54, while the Dutch group has a higher percentage in the age bracket 55–74.

The Dutch and Catalan participants' reading habits are also quite different. The following questions about their reading habits were posed to the participants and they had to rank them on a five-point Likert scale: "How often have you read a book in the last two years?",¹⁹ "How much do you like reading?",²⁰ "How often

19. 1=Never, 2=Once every three months, 3=Once a month, 4=Once or twice per week, 5=Daily.

20. 1=Dislike a great deal, 2=Dislike somewhat, 3=Neither like or dislike, 4=Like somewhat, 5=Like a lot.

have you read a book in Catalan/Dutch in the last two years?”²¹ and “How long did you spend reading on these occasions?”²²

The variable *Reading_Patterns* represents the average given to these four questions. Figure 3 shows the results for the two languages ($n=223$).

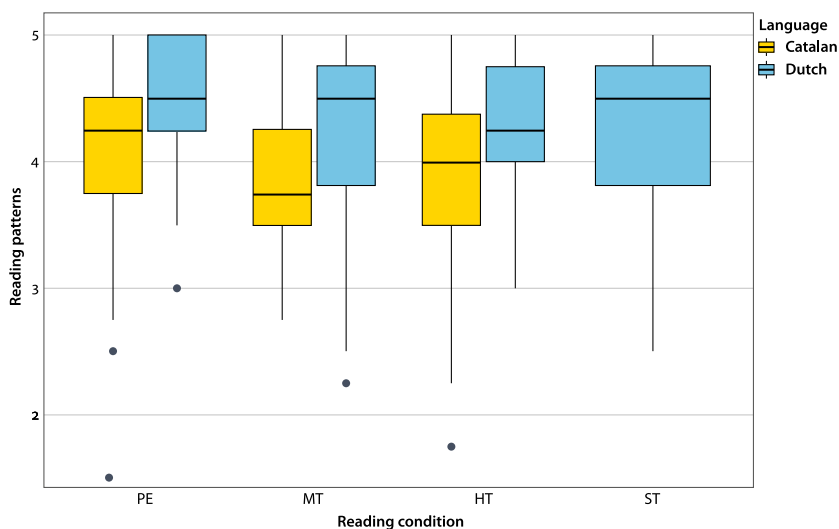


Figure 3. Reading habits by language and condition

Dutch participants have higher *Reading_Patterns* scores than the Catalan readers in all conditions. Since the data are not normally distributed, a Wilcoxon rank sum test was used to determine if there are significant differences between the two groups. The measured median *Reading_Patterns* score for Catalan (4) was significantly lower than that for Dutch (4.5) with a moderate effect ($p < .001$, effect size $r = 0.33$). The reasons for this could be the number of people who responded to our survey according to their age and their provenance (different avenues of recruitment). For example, Figure 2 shows that the Dutch group has more participants from reading clubs and reading sites (such as Hebban and Senia), while the Catalan group has a higher number from universities and Facebook groups. There were no statistically significant differences between the variable *Reading_Patterns* and the three translation conditions, so participants with higher or lower reading patterns were distributed evenly across the conditions.

21. 1=Never, 2=Once every three months, 3=Once a month, 4=Once or twice per week, 5=Daily.

22. 1=I don't read, 2=Less than 15 minutes, 3=Between 15 and 30 minutes, 4=Between 31 and 60 minutes, 5=More than 60 minutes.

In conclusion, most readers in the sample are women with a high level of education. The Dutch readers are slightly more educated, older, and declare reading more frequently than the Catalan group.

4.3 Comprehension questions

All participants had to answer at least five out of ten comprehension questions correctly to be able to continue to subsequent sections of the questionnaire. The mean value for all participants is above 7 for all reading conditions. The highest value is for the ST condition ($M=8.3$), followed by HT ($M=8.2$), PE ($M=7.9$), and MT ($M=7.8$). There are no significant differences between the conditions overall. However, a Kruskal-Wallis H test for non-parametric data²³ on the Catalan responses indicates that there is a significant difference between conditions ($H(2)=6.33, p<.05$). Post-hoc comparisons using the Conover test with the Holm-Bonferroni correction show statistically significant differences between PE and HT ($Z=-2.46, p<.05$). There are no significant differences among the translation conditions for the Dutch readers. Therefore, the comprehension questions are impacted by the reading condition only for Catalan, and the HT readers had significantly higher comprehension scores. To further explore the impact of comprehension, it would have been interesting to let participants continue regardless of the number of correct questions, but this would have jeopardized the validity of the experiment since the data were collected online.

4.4 Narrative engagement

Figure 4 shows the average ratings for the two languages of the twelve items related to narrative engagement, using a seven-point Likert scale ($n=223$). The Cronbach's alpha reliability²⁴ coefficient (α) is 0.85 for all the items in the scale, which is considered a reliable score.

The Catalan participants show a higher engagement than the Dutch participants in all conditions. Among the Catalan group, HT has the highest score, but this is not replicated in Dutch where ST, followed by PE, have the highest engagement scores. This is odd since the reviewers indicated that the HT was a translation of higher creativity than the PE in both languages, and reviewers were

23. This test is used to determine if there are statistically significant differences between two or more groups within the independent variable (reading condition) when the scale uses rank-based nonparametric values.

24. Cronbach's alpha measures the internal consistency of a scale. It gives an idea of how the items are interrelated and measure similar concepts.

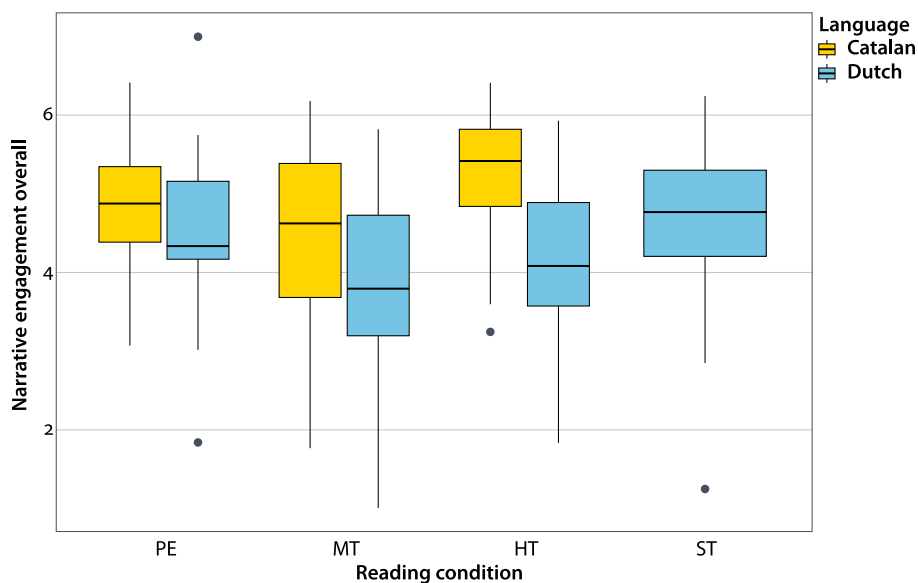


Figure 4. Narrative engagement overall per language and condition

unanimous in their assessment of the quality of these two conditions (Guerberof-Arenas and Toral 2022). It therefore appears that the Dutch readers do not share the same opinion as the professional reviewers.

Since the data does not meet the assumptions for a linear regression model that would allow us to investigate the interaction between engagement, condition, and languages, and given that the groups have different reading habits and the participants read different texts, we analyzed the two language groups separately.

The variable *Narrative_Engagement* was explored according to the different reading conditions using the Kruskal-Wallis H test for non-parametric data. In the Catalan group, there is a statistically significant difference between conditions ($H(2)=11.90$, $p<.001$) with a mean rank score of 65.64 for HT, 48.28 for PE, and 41.68 for MT. Post-hoc comparisons using the Conover-Iman test with the Holm-Bonferroni correction show statistically significant differences between PE and HT ($Z=-2.54$, $p=.01$) and between MT and HT ($Z=-3.51$, $p=.00$). For the Dutch group, there is a statistically significant difference between conditions ($H(3)=10.12$, $p=.02$) with a mean rank score of 72.88 for ST, 65.57 for PE, 58.27 for HT, and 45.52 for MT. Post-hoc comparisons show statistically significant differences only between MT and ST ($Z=-3.15$, $p=.02$).²⁵ This is indeed surprising.

25. The p -values presented here are not adjusted. In the Conover-Iman test, the null hypothesis is rejected if $p \leq \alpha/2$.

It appears that in the Dutch group, there is no significant difference between the translation conditions.

To clarify these findings and following the same statistical methods, each category in the narrative engagement scale was analyzed per language. Table 2 shows the results together with the outcome of the statistical testing.

Table 2. Narrative engagement results per category and language

Category	Catalan	Dutch
Narrative understanding 3 items	Significant MT ≠ HT ($Z = -3.26, p = .01$)	Significant MT ≠ ST ($Z = -4.11, p = 0$)
Attentional focus 3 items	Significant MT ≠ HT, $Z = -2.65, p = .02$	Not significant
Narrative presence 3 items	Not significant	Not significant
Emotional engagement 3 items	Not significant	Not significant
Overall narrative engagement 12 items	Significant PE ≠ HT ($Z = -2.54, p = .01$) MT ≠ HT ($Z = -3.51, p = .001$)	Significant MT ≠ ST ($Z = -3.15, p = .01$)

As far as engagement is concerned, the main difficulty with the texts translated using MT seems to be with narrative understanding, which relates to the ease of comprehension of a story. The participants' answers to these questions included the following: "At points, I had a hard time making sense of what was going on in the story," "My understanding of the characters is unclear," and "I had a hard time recognizing the thread of the story." Therefore, and even though these readers responded satisfactorily to the comprehension questions, they did not perceive this activity as an easy one in the MT condition.

In the Catalan group, there is also an issue with attentional focus, the state of being engaged and not distracted. The participants reacted to these statements with the following: "I found my mind wandering while reading the story," "While reading, I found myself thinking about other things," and "I had a hard time keeping my mind on the story." We had anticipated that the MT readers would find it more difficult to be transported to the story because they would find elements of distraction that would prevent this. This seems to be the case only for the Catalan readers.

This is partially in line with our previous finding (Guerberof-Arenas and Toral 2020) where narrative understanding was the dimension of narrative engagement most affected by the reading condition. It seems that when reading

a translated text, narrative presence (the feeling that one has entered the world of the story) and emotional engagement (feeling for and with the characters) are dimensions more strongly linked to the story, and, hence, the world created by the writer, than to the translation.

4.5 Visual imagery

This section of the questionnaire was added to investigate if, depending on the condition, readers experience a higher level of difficulty when imagining the characters, the situations, and the world depicted in the story. We found minor differences (none significant) in these scores between conditions: the Dutch readers, as in the rest of the survey, scored these statements lower than the Catalan readers.

4.6 Enjoyment

Figure 5 shows the results of the average scores the participants gave, using a seven-point Likert scale for the two languages, for the three items related to enjoyment ($\alpha = 0.87$, $n = 223$).

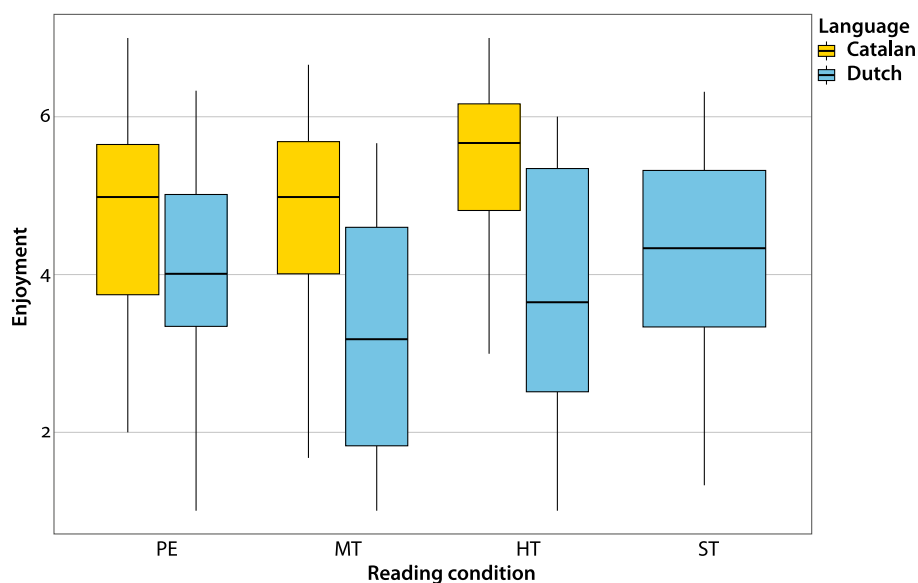


Figure 5. Enjoyment per language and condition

The Dutch readers enjoyed the story less than the Catalan readers. Similar patterns as for narrative engagement are found: the Catalan readers had the highest enjoyment in HT and the Dutch readers in ST. As before, a Kruskal-Wallis H test was carried out to compare enjoyment between the conditions per language group. In the Catalan group, there was evidence of a difference ($H(2) = 6.65$, $p < .05$) with a mean rank score of 62.63 for HT, 46.68 for PE, and 46.49 for MT. Pairwise comparisons, however, reveal no statistically significant differences between the conditions. In the Dutch group, there was no evidence of a significant difference. The median rank scores are 71 for ST, 64.52 for PE, 58.61 for HT, and 47.97 for MT. These results differ from our previous experiment (Guerberof-Arenas and Toral 2020) where MT scored significantly lower in enjoyment than PE and HT.

4.7 Translation reception

Figure 6 shows the results of the average ratings given for the eight quantifiable items ($\alpha = 0.79$) using a seven-point Likert scale for the two languages ($n = 193$ since the ST condition is not included).

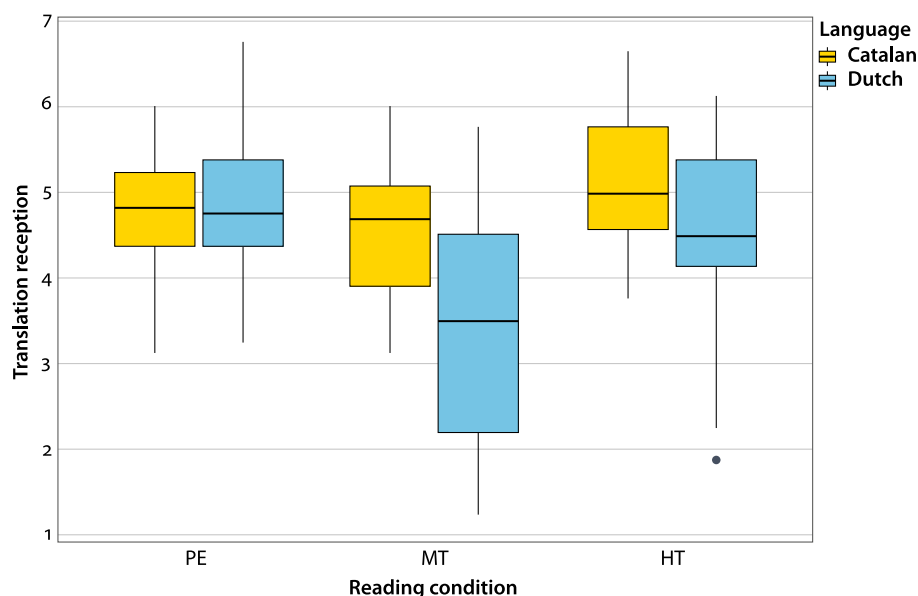


Figure 6. Translation reception per language and condition

The Dutch participants gave lower scores to MT and HT on average, but similar scores to the Catalan participants when rating PE. As with previous variables, in Catalan the reception is highest for HT, but, unexpectedly, the difference in reception between PE and MT does not seem as pronounced. For the Dutch participants, the MT scores are indeed lower than the PE and HT scores. A Kruskal-Wallis H test shows a statistically significant difference between conditions in Catalan ($H(2)=8.66, p=.01$) with a mean score rank of 62.6 for HT, 51.62 for PE, and 41.47 for MT. Post-hoc comparisons show statistically significant differences only between MT and HT ($Z=-3.05, p=.01$). For Dutch, there is a statistically significant difference between conditions ($H(2)=16.68, p=0$) with a mean rank score of 55.47 for PE, 51.39 for HT, and 29.78 for MT. Post-hoc comparisons show statistically significant differences between MT and HT ($Z=-3.55, p<.001$), and between PE and MT ($Z=4.15, p<.001$). These results, which appear to favour the Catalan MT system, could be an indication of the quality of this system. However, the Dutch readers tended to give lower scores overall, and they reported higher scores on the measures for reading habits. The results could therefore be an indication of a more demanding reader.

4.8 Comparison between translation and original text

Participants who read the original English text were also presented with the following questions/statements: “How easy was the text to understand?”, “I thought the text was very well written,” “I found words, sentences or paragraphs that were difficult to understand,” “I found words, sentences or paragraphs that I especially liked,” and “Would you like to read a text by the same author?” Since the first four were also posed to participants in the translation conditions, the responses were compared to see if, as with other variables, readers of the original text gave a higher score than those that read the translations. Figure 7 shows the average scores for each condition including the ST.

The ST did not obtain higher scores than the translation conditions, and the Catalan HT condition had the highest scores. Indeed, a Kruskal-Wallis H test shows a statistically significant difference between conditions in Catalan ($H(2)=15.92, p=0$) with a mean rank score of 65.4 for HT, 53.93 for PE, and 36.68 for MT. Post-hoc comparisons reveal a statistically significant difference only between MT and HT ($Z=-4.27, p<.001$) and between PE and MT ($Z=2.58, p=.01$). For Dutch, there is a statistically significant difference between conditions ($H(3)=15.48, p=0$) with a mean rank score of 71.82 for PE, 66.83 for ST, 64.15 for HT, and 39.45 for MT. Post-hoc comparisons show statistically significant differences between MT and HT ($Z=-2.94, p<.01$), between PE and MT ($Z=3.79, p<.001$), and between MT and ST ($Z=-3.24, p<.001$). Again, here we can see that

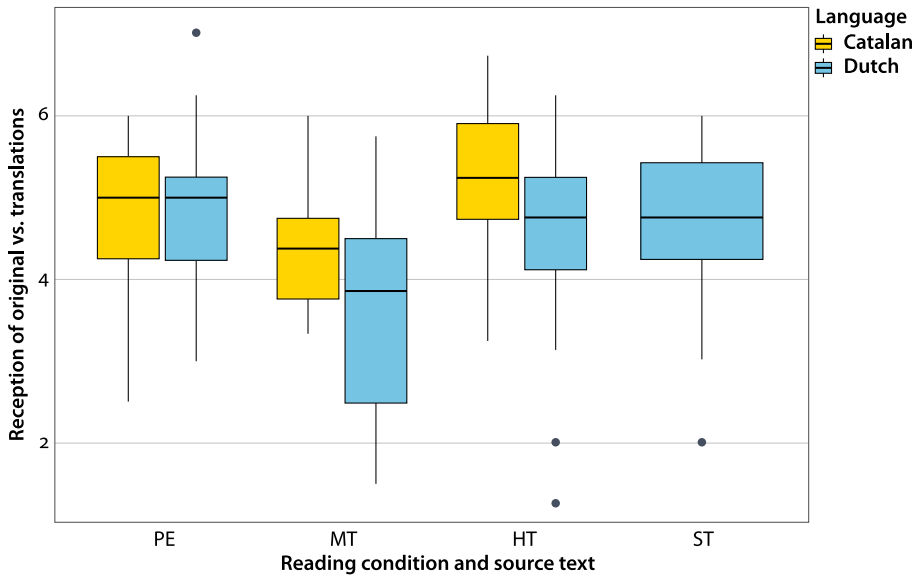


Figure 7. Reception of the translation vs. reception of the source text

the quality of the Catalan and Dutch MT systems plays a role in reception, while there are no significant differences between PE and HT.

4.9 How much would you pay?

After debriefing the participants about the author and reading condition, they were asked to indicate the amount of money they would be willing to pay for the text that they read, using the following answer options: 1=Less than a euro, 2=Between 1 and 5 euros, 3=Between 5 and 10 euros, 4=Between 10 and 15 euros, and 5=More than 15 euros. Figure 8 shows the results for this question.

Some participants were willing to pay more than 15 euros when they realized that the translation was done by professional literary translators. They were willing to pay less for PE, (up to 10 Euros, with some being willing to pay less than 1 Euro), and even less for MT. The amount that the Dutch participants were willing to pay for the ST is remarkably low, possibly because the author is no longer alive and they might be aware that the text is free of copyright. We wonder if participants would be willing to pay the same amount if we had asked them before the debriefing. Still, the results show that participants do attach value to hand-crafted translations, even if they are not willing to pay the same price for the original work.

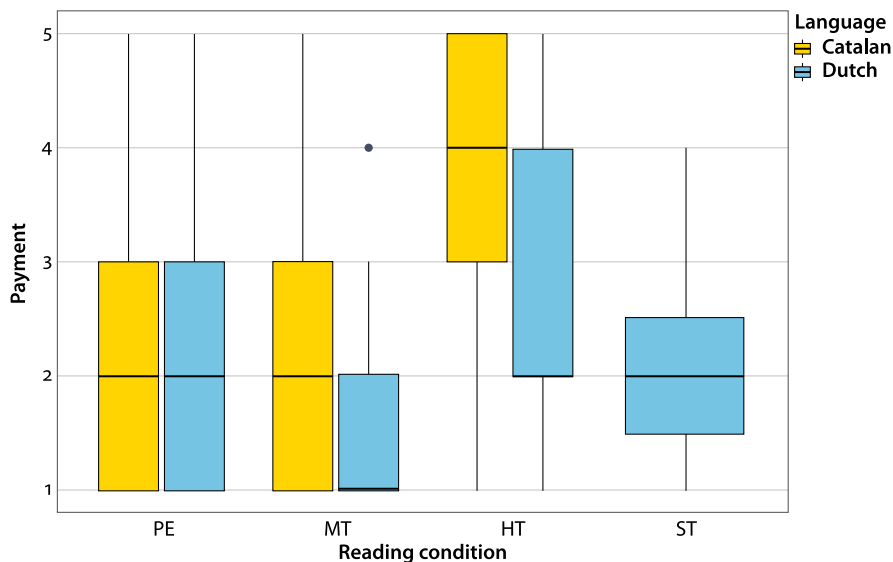


Figure 8. Payment per language and condition

A Kruskal-Wallis H test shows that there is a statistically significant difference between conditions in the Catalan group ($H(2)=30.24$, $p=0$) with a mean score rank of 73.25 for HT, 44.40 for PE, and 37.74 for MT. Post-hoc comparisons show a statistically significant difference only between PE and HT ($Z=-5.22$, $p=0$), and between MT and HT ($Z=-6.10$; $p=0$). For the Dutch group, there is a statistically significant difference between conditions ($H(3)=17.31$, $p=0$) with a mean rank score of 71.17 for ST, 70.89 for HT, 57.52 for PE, and 41.98 for MT. Post-hoc comparisons show statistically significant differences between MT and HT ($Z=-4.63$, $p<.001$), but not between the other pairs of conditions. The participants are therefore willing to pay more for the service of professional translators.

4.10 Quality of the MT systems according to the readers

The participants who were randomly assigned to the MT condition were asked, after the debrief, to indicate their opinion of the quality of the system on a seven-point scale, where 1 = Extremely poor and 7 = Extremely good. Figure 9 shows the results per language for the variable MT quality ($n=64$).

The Catalan participants ranked the MT quality noticeably higher than the Dutch participants. Since the data is not normally distributed, a Wilcoxon rank sum test was used and this showed that the difference is significant with a large effect ($p<.001$, effect size $r=0.62$). The reasons for this could be that the Dutch participants were more demanding of quality since their scores on the self-

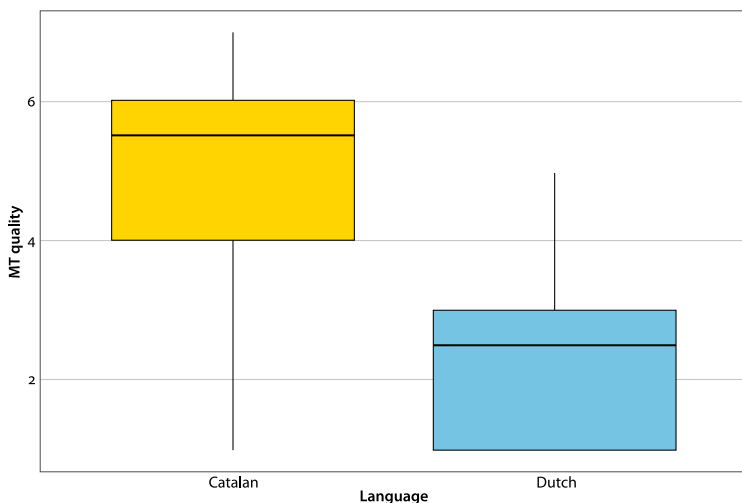


Figure 9. Reported MT quality per language

reported reading habits suggest that they are more avid readers, but it could also indicate differences between the quality of the systems. The Dutch translators had also reported a lower perception of quality than the Catalan translators (Guerberof-Arenas and Toral 2022).

4.11 Summary of findings

Table 3 gives an overview of the quantitative results from the survey, distributed per language.

Table 3. Summary of findings

Category	Catalan	Dutch
Comprehension	<p>Highest values in HT</p> <p>Significant difference between conditions</p> <p>PE ≠ HT ($Z = -2.46, p < .05$).</p>	<p>Highest values in ST</p> <p>No significant difference between conditions</p>
Reading habits	<p>No significant difference between conditions</p>	<p>No significant difference between conditions</p> <p>Significantly higher than the Catalan group</p> <p>$p = .00$, effect size $r = 0.33$</p>

Table 3. (continued)

Category	Catalan	Dutch
Narrative engagement	Highest values in HT Significant difference between conditions PE \neq HT ($Z = -2.54, p = .01$) MT \neq HT ($Z = -3.51, p = .00$)	Highest values in ST Significant difference between conditions MT \neq ST ($Z = -3.15, p = .01$)
Imagery	No significant difference between conditions	Lower scores than Catalan, but no significant difference between conditions
Enjoyment	Highest values in HT Significant difference between conditions However, pairwise comparisons show no significant differences between modalities	Highest values in ST No significant difference between conditions
Translation reception	Highest values in HT Significant difference between conditions MT \neq HT ($Z = -3.05, p = .01$)	Highest values in PE Significant difference between conditions MT \neq HT ($Z = -3.55, p = .00$) PE \neq MT ($Z = 4.15, p = .00$)
Translation vs. original	Highest values in HT Significant difference between conditions MT \neq HT ($Z = -4.27, p = .00$) PE \neq MT ($Z = 2.58, p = .01$)	Highest values in PE Significant difference between conditions MT \neq HT ($Z = -2.94, p = .01$) PE \neq MT ($Z = 3.79, p = .00$) MT \neq ST ($Z = -3.24, p = .00$)
Payment	Highest values in HT Significant difference between conditions PE \neq HT ($Z = -5.22, p = .00$) MT \neq HT ($Z = 6.10, p = .00$)	Highest values in HT Significant difference between conditions MT \neq HT ($Z = -4.63, p = .00$)
MT quality	Mean value 5 (Slightly good)	Mean value 2.5 (Poor)

4.12 Comments from the readers

To complement the quantitative analysis, the readers were asked to comment on those parts of the text they found difficult, those parts they liked, how they realized they were reading a translation, and for final comments at the end of the survey (before the debriefing).

4.12.1 *Difficult parts*

There were no comments from the readers of the original English, while readers of the other conditions did make comments, with MT receiving the most.

In the HT version, both the Catalan and Dutch readers found problems related to the story (“it was difficult to know what was going on at the beginning,” “the character descriptions,” “the phone numbers,” and “it was a terribly boring piece”), and certain words in Catalan such as *trigèmins* ‘triplets,’ “the use of a very high register mixed with a low register,”²⁶ lack of commas, or too much use of the conjunction *ni* ‘either/neither’.

In the PE version, the Catalan readers commented on some words that confused them such as the name of one of the characters, Affleck,²⁷ the use of unknown vocabulary, the sobriquets²⁸ (e.g., the Dutch *kattenbak* for the English ‘catbox’), and the word *zelador* ‘caretaker’ in Catalan. Some comments referred to the story: a reader thought the first part was confusing, one reader complained about the text being too long, and a Dutch reader commented on bizarre and out-dated parts of the text.

In the MT version, there were comments about incorrect sentence structures (“Some sentences were built up a bit oddly and referred to each other in unclear terms”); odd punctuation; lack of coherence; incorrect use of verb tenses; the use of the wrong vocabulary such as *foto* (the Dutch word for ‘photography’ instead of ‘painting’) and the translation *kamer voor twee* [room for two] for ‘space for two’; and words left in English in the Catalan version such as “daub, sobriquet, sheepdip [sic], drupelets [sic].” The participants who read the MT made similar comments about the story as those who read the HT and PE: a confusing beginning, or the use of neologisms. Only one participant referred to MT: “It seems like a sort of machine translation; there were words that did not correspond to the meanings that we could get from the context. For example, there was the word *jackrabbīt*.”

As demonstrated in comments from readers for previous, similar experiments, when commenting about difficulties, the HT and PE readers tend to comment more on intrinsic characteristics of the story, with a few comments on the

26. This could be partially due to Vonnegut’s ironic style.

27. This name was translated into Catalan to reflect the same image evoked by the name ‘Wehling’ in English: someone who is suffering or in pain. When compiling the PE version that was post-edited by two translators, some decisions were made to unify the version, and the name ‘Affleck’ was left in this version. The Catalan readers were not aware of the original name, so it is surprising that they remarked on this name.

28. Vonnegut creates a series of new terms to refer to the death system based on cultural references. The translators adapted these to the target language.

language, while the MT readers tend to refer mainly to the characteristics of the text, the translation, the structure, punctuation, out of context vocabulary, and to a sense of general ‘oddity’ that confuses them.

4.12.2 Preferred sections

With regard to the parts that the readers liked, there were no comments on the original English, but comments were made on all translation conditions in both languages. The number of comments on the MT was the highest, which is surprising.

In the HT version, participants referred to parts of the story they liked: “The painter describing the mural,” “The father’s moral conflict when deciding on the life or death of his children and the irony of the painter,” “The end has undoubtedly been the best part of the story,” and “I found the paragraph in which the father took his decision to kill Hitz, Leora and himself quite impressive.” There were some references to the translation itself: “*El número de telèfon CONOC (ser o no ser)*” [The telephone number 2BR02B (to be or not to be)], “The synonyms/names of Suicide Studio,” and “I thought the names for the gassing in connection with birth control were funny, like *kattenbak*.”

In the PE version, readers also referred to certain moments in the text: “The atmosphere was set by the phrases that for every child born, someone else must die,” “The description of the painting,” “The predominant colour purple,” “The bit about the future father taking fate into his hands and making sure his three children are all allowed to live,” and “The paragraph where he describes the Garden of Eden.” Many also referred to the ending of the story, its moral dilemma, and the painter’s reflections. As before, they also refer to the translations: “The nicknames for the Federal Bureau of Termination, such as *kattenbak*,” “The sentence with the flowery descriptions of the body that regulates killing *dompelaar, kattenbak*,” “I’ve been amused by some of the names used for the gas chambers. Also, the phone number, which when you read it is ‘To be or not to be.’”

For the MT version, the readers mentioned certain sections again such as the description of the Garden of Eden, the father’s moral dilemma, reflections about getting older, the life cycles and general reflections of the characters, their characterization (sincerity, frankness), and the impressive ending. They also mentioned certain translations, for example, a Dutch reader mentioned a particular line of the song “*Ik ga van deze oude planeet af, laat een baby ijn [sic] plaats innemen*” [I am leaving this old planet, let a baby take my place]. A Catalan reader commented: “I really enjoyed reading words I haven’t heard in a long time. One of the reasons I like to read science fiction in Catalan. Also to support the language and writers,” “I liked the way the writer/translator was able to divide the scenes using the language,” and “*La teva ciutat t’ho agraeix però més t’ho agraeixen les futures*

generacions” [Your city is grateful to you, but future generations are more grateful to you]. Also, in this version, the readers referred to word choices for the gas chambers and the names of the characters that in this case were left in the original English.

When describing what they liked, the readers focused more on what they liked about the story, but also on parts of the translation, which is surprising for MT; it seems that the MT did, after all, capture a certain essence of the text. It seems that when it comes to expressing positive aspects of a story, the reading condition was not as determinative for the readers, even if the MT readers did give lower scores to this condition throughout the experiment.

4.12.3 *How did you realize it was a translation?*

Readers reported a higher score (meaning that the fact that the text was a translation was more salient) for the MT, than for the HT and PE. As a follow-up question, we asked them how they arrived at this conclusion because we wanted to know if this was due to errors or to the setting of the story.

In the HT version, most readers commented, as we expected, that characters’ names and the action that occurred in the United States or Chicago, or even the translated names of the offices, echoed names in the United States. Some mentioned that some words or structures were not very frequent in their language, or literal, and that the style resembled English with “little flourish.” In Dutch, the title ‘*2BR0zB*’ (to be or not to be) was left in English, and this was also mentioned. One Dutch reader mentioned: “I generally prefer not to read translations, but here it wasn’t distracting,” which reflects a tendency in the Netherlands to read literature from English-speaking authors in the original language, but also an appreciation for this translation by this reader.

In the PE version, readers commented that some expressions, constructions, and words read like translations, or that they were “clumsy.” They also commented that the translation was artificial and “cold” and, as before, also commented on the characters’ names, the story not being set in their country, but in Chicago, and the names of the offices in the Dutch translation. One Dutch reader had read the story in English before, so he knew it was a translation when he started reading.

For the MT version, the readers made more comments about the style, incoherent sentences, wrong sentence structure, words out of context, absence of certain pronouns, unnatural dialogues, expressions translated literally from English without corresponding to the concept (e.g., in Catalan *ningú a les meves sabates* [nobody in my shoes]), punctuation, grammatical errors, alternate use of formal and informal pronouns, terms left in English, and confusing storyline. One reader commented: “Phrases a bit unconnected, but since it was so surreal, I didn’t know if it was because of that or because of a bad translation.” As before,

they also commented on the style of narration with few adjectives and short sentences more typical of English, the names of the characters, names of places left in English, the setting in Chicago, and the use of the term ‘2BRo2B’.

As we have observed in previous experiments, the main giveaway that the MT is a translation is the quality of the language, while in the other conditions, the main giveaway is the setting of the story.

4.12.4 *Final comments*

The participants’ final comments were varied, but they did not differ noticeably by reading condition. The readers commented that they either did or did not like the story. Some found it “fascinating and intriguing” while others found it “unpleasant” and a “terribly dark text.” They also mentioned whether they like or dislike science fiction or dystopian literature. Many were intrigued by the experiment and wanted to know the results. One Dutch reader in the MT version commented, “I happen to know that this is a story by Kurt Vonnegut. But I can hardly imagine that someone has translated his story in the way presented before. This translation really does not do justice to his authorship,” which happens to touch upon an issue often discussed among translators and scholars about the role of MT in promoting literature in unusual language combinations. Can MT really do justice to authorship?

5. Conclusions

In the second part of the CREAMT project, the data obtained partially supports the conclusions from the pilot experiment (Guerberof-Arenas and Toral 2020). Catalan users reading a text produced via MT have different reading experiences, and they are more engaged when they read HT. In this experiment, the HT version scores significantly higher than PE in many categories. This is not surprising, as the HT version had scored higher in creative shifts and lower in errors when compared with the PE and MT versions during the review (Guerberof-Arenas and Toral 2022). Moreover, the Catalan reviewers praised many of the solutions found in the HT version (even if the PE version was post-edited by the same two translators).

However, for the Dutch group, the results differ. They returned higher values in most categories when reading the original ST (this condition was not present in Catalan) or the PE version than when reading MT or HT. For many categories, the reading condition does not show any significant difference even if PE tends to score higher. These results seem odd since all Dutch texts were rated by professional reviewers and they found that the PE version, when compared to the HT,

was too close to the English and seemed to be translated by amateurs or novel translators and, in addition, that it contained far more errors than the HT version. We hypothesize that the cause for these results could be the habituation of Dutch readers to reading English literary texts in the original language. Consequently they may favour a translation closer to the ST, in this case the PE, than a more creative alternative in Dutch. Of course, this hypothesis would need to be tested with a different ST and source language. Overall, the Dutch readers have a fuller reading experience when reading in the original language.

It must be said that, according to the reviewers, the Catalan translators performed better than their Dutch counterparts. This could also be the reason why the Dutch readers were not as impressed by the HT as the Catalan readers. In literary translation, the role that individual translators play appears to be decisive for readers. This would need to be further tested.

The only category where the Dutch readers rated HT significantly higher was in the payment category, but only after being debriefed as to the nature of the translation. This means that readers do place a monetary value on the work done by professional translators, even if this was not reflected in the other scales for the Dutch group. This could be a question of trust in what a professional translator can accomplish, but could also be based on their experiences when using publicly available MT engines.

Further, the Catalan participants rated the quality of the MT system significantly higher than the Dutch readers. It could be the case that the quality was indeed different even if automatic metrics did not reflect this – the Dutch translators also scored the quality lower than the Catalan translators, but this could also be related to the higher scores on reading habits measures for the Dutch readers, signalling that they are possibly more demanding readers.

We have also confirmed that, irrespective of the target language, when looking into narrative engagement, the category narrative understanding, and attentional focus (in part), are the ones most affected when using MT, while narrative presence, emotional engagement, and visual imagery are not significantly different in any condition. This is also reflected in the answers to the comprehension questions devised for this story. However, readers commented on several critical issues related to language when reading the MT that might negatively impact comprehension and their opinion of the author, if MT is used without editing.







Funding

This project has received funding from the European Union's Horizon 2020 research and innovation programme under the Marie Skłodowska-Curie grant agreement (Grant No. 890697). Open Access publication of this article was funded through a Transformative Agreement with University of Groningen.


Acknowledgements

We would like to thank the translators Carlota Gurt, Yannick Garcia Porres, Núria Molines Galarza, Josep Marco Borillo, Scheherezade Surià, Theo Schoemaker, Roos van de Wardt, Linda Broer, and Leen van de Broucke, and the annotators Tia Nutters and Gerrit Bayer-Hohenwarter for their crucial contribution to this study. We would also like to thank all readers who contributed to this project, many of whom waived the offered payment. Thank you for making this possible!

References


-  Busselle, Rick, and Helena Bilandzic. 2009. "Measuring Narrative Engagement." *Media Psychology* 12 (4): 321–347.
- Colman, Toon, Margot Fonteyne, Joke Daems, Nicolas Dirix, and Lieve Macken. 2022. "GECO-MT: The Ghent Eye-tracking Corpus of Machine Translation." In *13th Conference on Language Resources and Evaluation (LREC 2022)*, edited by Nicoletta Calzolari et al., 29–38. European Language Resources Association (ELRA).
-  Dixon, Peter, Marisa Bortolussi, Leslie C. Twilley, and Alice Leung. 1993. "Literary Processing and Interpretation: Towards Empirical Foundations." *Poetics* 22 (1–2): 5–33.
-  D'Ydewalle, Géry. 1984. "Processing TV Information and Eye Movements Research: Interfaces in the Field." In *Readings on Cognitive Ergonomics – Mind and Computers: Proceedings of the Second European Conference, Gmunden, Austria, September 10–14, 1984*, edited by Gerrit C. van der Veer, Michael J. Tauber, Thomas R. G. Green, and Peter Gorny, 200–204. Berlin: Springer.
-  D'Ydewalle, Géry, Johan van Rensbergen, and Joris Pollet. 1987. "Reading a Message When the Same Message Is Available Auditorily in Another Language: The Case of Subtitling." In *Eye Movements from Physiology to Cognition: Selected/Edited Proceedings of the Third European Conference on Eye Movements, Dourdan, France, September 1985*, edited by J.K. O'Regan and A. Levy-Schoen, 313–321. Amsterdam: Elsevier.
-  D'Ydewalle, Géry, and Johan van Rensbergen. 1989. "13 Developmental Studies of Text-Picture Interactions in the Perception of Animated Cartoons with Text." In *Knowledge Acquisition from Text and Pictures*, edited by Heinz Mandl and Joel R. Levin, special issue of *Advances in Psychology* 58: 233–248.
-  Guerberof-Arenas, Ana, and Antonio Toral. 2022. "Creativity in Translation: Machine Translation as a Constraint for Literary Texts." *Translation Spaces* 11 (2): 184–212.

- doi Guerberof-Arenas, Ana, and Antonio Toral. 2020. "The Impact of Post-Editing and Machine Translation on Creativity and Reading Experience." *Translation Spaces* 9 (2): 255–282.
- doi Hakemulder, Jemeljan F. 2004. "Foregrounding and Its Effect on Readers' Perception." *Discourse Processes* 38 (2): 193–218.
- Hu, Ke, Sharon O'Brien, and Dorothy Kenny. 2021. "A Reception Study of Machine Translated Subtitles for MOOCs." In *Mapping Contemporary Audiovisual Translation in East Asia*, edited by Dingkun Wang, Xiaochun Zhang, and Arista Szu-Yu Kuo, special issue of *Perspectives* 28 (4): 521–538.
- doi Kotze, Haidee, Berit Janssen, Corina Koolen, Luka van der Plas, and Gys-Walt van Egdom. 2021. "Norms, Affect and Evaluation in the Reception of Literary Translations in Multilingual Online Reading Communities: Deriving Cognitive-Evaluative Templates from Big Data." *Translation, Cognition & Behavior* 4 (2): 147–186.
- doi Kruger, Haidee. 2013. "Child and Adult Readers' Processing of Foreign Elements in Translated South African Picturebooks." *Target: International Journal of Translation Studies* 25 (2): 180–227.
- doi Kruger, Jan-Louis. 2018. "Eye Tracking in Audiovisual Translation Research." In *The Routledge Handbook of Audiovisual Translation*, edited by Luis Pérez-González, 350–366. Abingdon: Routledge.
- doi Kuijpers, Moniek M., Frank Hakemulder, Ed S. Tan, and Miruna M. Doicaru. 2014. "Exploring Absorbing Reading Experiences: Developing and Validating a Self-Report Scale to Measure Story World Absorption." *Scientific Study of Literature* 4 (1): 89–122.
- doi Mangen, Anne, and Don Kuiken. 2014. "Lost in an iPad: Narrative Engagement on Paper and Tablet." *Scientific Study of Literature* 4 (2): 150–177.
- NOS. 2022. "Jongeren lezen graag boeken, maar dan wel in het Engels [Young people like to read books, but then in English]." NOS, April 16. <https://nos.nl/nieuwsuur/artikel/2425380-jongeren-lezen-graag-boeken-maar-dan-wel-in-het-engels>
- Nuland, Sherwin B. 1995. *How We Die: Reflections of Life's Final Chapter, New Edition*. New York: Vintage.
- doi Orrego-Carmona, David. 2018. "Audiovisual Translation and Audience Reception." In *The Routledge Handbook of Audiovisual Translation*, edited by Luis Pérez-González, 367–382. Abingdon: Routledge.
- Ortiz Boix, Carla. 2016. *Implementing Machine Translation and Post-Editing to the Translation of Wildlife Documentaries Through Voice-over and Off-Screen Dubbing*. PhD diss. Universitat Autònoma de Barcelona. <http://www.tdx.cat/handle/10803/400020>
- Stasimioti, Maria, and Vilemini Sosoni. 2022. "Creative Texts Translation vs Post-Editing: A Qualitative Study of the Product Quality, the Translators' Perception and Audience's Reception." Presentation at the *Workshop on Creativity and Technology: Proceedings of the 1st NETTT Conference*. Rhodes: NETT.
- Vonnegut, Kurt. 1999. *Bagombo Snuff Box*. New York: G. P. Putnam's Sons.
- Walker, Callum. 2020. *An Eye-Tracking Study of Equivalent Effect in Translation: The Reader Experience of Literary Style*. Cham: Palgrave Macmillan.
- doi Walker, Callum. 2021. "Investigating How We Read Translations: A Call to Action for Experimental Studies of Translation Reception." *Cognitive Linguistic Studies* 8 (2): 482–512.


-  Whyatt, Bogusława, Olga Witczak, Ewa Tomczak-Łukaszewska, and Olha Lehka-Paul. 2023. "The Proof of the Translation Process Is in the Reading of the Target Text: An Eyetracking Reception Study." *Ampersand* 11, 100149.

Address for correspondence

Ana Guerberof-Arenas
Center for Language and Cognition
University of Groningen
Oude Kijk in 't Jatstraat 26
9712 EK GRONINGEN
The Netherlands

 a.guerberof.arenas@rug.nl
<https://orcid.org/0000-0001-9820-7074>

Co-author information

Antonio Toral
University of Groningen
a.toral.ruiz@rug.nl
 <https://orcid.org/0000-0003-2357-2960>

Publication history

Date received: 27 August 2022
Date accepted: 6 March 2024
Published online: 2 April 2024