

Dear Ellen,

Hereby we submit a revised version of our paper entitled "The use of MT by undergraduate translation students for different learning tasks", which we originally submitted to the *Journal of Data Mining and Digital Humanities* on 31 January 2022.

We would like to thank the two reviewers for their valuable input. In the new version, we have tried to incorporate all of the comments made by the two reviewers. As a result, we feel that the quality of the paper has improved substantially. Below, we discuss for each comment (that required revision) which actions were taken to address the reviewer's concerns.

First, we will address the comments of Reviewer 1:

Comment 1.1

Is the manuscript technically sound and does it provide convincing evidence that support the conclusions?

Your opinion

0.8 - Provide a precise technical evidence

Comment

Hypothesis, methodology, results and discussion are clearly presented. The validity of the results is a little bit limited by the fact that different cohorts are being compared (it would probably have been more interesting to follow one cohort and have them take the test at the end of the first, then of the second, then of the third year).

We agree with the reviewer that the design of this particular study does not allow us to draw conclusions about changes in the use of MT at the level of the individual student or cohort. Therefore, we support the suggestion that comparing the result of a single cohort across study years would be an interesting follow-up to the current research. We have added this suggestion to the discussion section. For the exact phrasing, we refer to the change marked with **Comment 1.1** in the revised version of the manuscript (page 8).

Comment 1.2

Does the manuscript adhere to standards in this field for data availability?

Your opinion

0 - No information is provided concerning data availability

Comment

I didn't find information on the availability of the data.

We plan to make the data and SPSS output files available via an online repository. This information has also been added to the text (see the change marked with **Comment 1.2** on page 5). For the purpose of the review we have attached a folder containing all the relevant files. Could you please advise us on how to make these available permanently for readers in a way that fits with standard procedures of the journal and how we should refer to these files in the article?

Comment 1.3

Is the manuscript presented in an intelligible fashion and written in standard English?

Your opinion

0.5 - Legible manuscript that would definitely benefit from proof-reading

Comment

The language is not always as idiomatic and fluent as would be comfortable for the reader. I have made a few suggestions in the annotated version attached, but it would be better if a native speaker did a thorough copyedit.

We thank the reviewer for this suggestion and for the specific linguistic issues that he or she pointed out in the original manuscript. We have tried to improve the fluency of the article in two ways: by incorporating the remarks made by the reviewer (all changes marked with **Comment 1.3a**) and by asking a native English speaker and experienced proofreader to perform a thorough check on the language used in the article; the changes that she suggested are marked with **Comment 1.3b**.

Secondly, we will address the comments of Reviewer 2:

Comment 2.1

Is the manuscript technically sound and does it provide convincing evidence that support the conclusions?

Your opinion

0.3 - Several aspects from the methodology are missing

Comment

Not completely. For example:

1. It is not specified which tools were used in this study.
2. The questions of the survey were not detailed, therefore the reader cannot fully understand the development of the analysis.
3. There were problems with the scale employed. The authors state, "the fourpoint scales on which the first-year students indicated their answers were transformed into seven-point scales following the procedure outlined in [Lewis and Sauro, 2020]" (page 4). However, they do not carry out the same procedure that is detailed in this citation. Lewis and Sauro used a numerical scale accompanied by the description of the two extreme values. Nevertheless, the authors did not use numbers but "(Never, Sometimes, Often or Very often)". Then, they automatically transformed it into numerical values without specifying the validation.
4. The authors compared two groups of no homogeneous. Each group conducted a different survey carried out in different conditions. That should have been taken into account in the analysis. They only comment on it in the limitations section.
5. They also commented on the limitations section the fact that only Dutch students were included. Therefore, it is recommended that the conclusions not be generalized.

We agree with the reviewer that this article could benefit from a methodological upgrade and have tried to incorporate all suggestions made by this reviewer. We will address these issues below:

1. We have made the references to the two different tools used for conducting the surveys in this study (Microsoft Forms and Questback) more explicit. Please see the changes marked with **Comment 2.1.1**. on page 4 and 5). The reference to the tool used for data analysis (IBM SPSS version 28) did not need further clarification in our opinion.
2. We agree that it would be easier for a reader to follow the logic of the analysis if the questions of the survey were detailed *within* the article. However, the prescribed maximum length did not allow us to go beyond what we have described in section 2.2. We are planning however to

upload an export of the questions (and data) of each survey to an online repository; this should allow readers to gain more insight into the exact questions and the nature of the data. It should be noted, however, that only part of the surveys for year 2 and 3 pertained to the use of MT by translation students (year 2: slide 36 to slide 92; year 3: slide 42 to 98), and that the original material was in Dutch. Please see the revision marked with **Comment 2.1.2** on page 5.

3. The fact that the first-year students used a different answering scale than the second- and third-year students is indeed an issue that merits attention. We agree with the reviewer that Lewis and Sauro only explained the procedure for a semantic differential scale, and not for an ordinal scale with labels. However, the same procedure has also been applied to Likert scales that were accompanied by labels (IBM, 2020). Therefore, we have added the IBM reference to the text (see the change marked with **Comment 2.1.3** on page 5). Perhaps more importantly, our decision to use non-parametric tests for testing the hypothesis also implied that the exact nature of the transformation would have no effect on the outcomes of the statistical tests, as the rank order of the values would remain the same. We explicitly verified this by converting the 7-point scales to 4-point scales (the exact opposite of the procedure described in the article) and repeating the analysis, which did indeed lead to the same outcomes. Although we realize that there is no ideal solution for converting scales in order to make them comparable, we are confident that any other procedure would have led to the same results.
4. We agree that the groups that were compared were not homogeneous and that the circumstances in which they participated in the research differed. That is one of the reasons why we used non-parametric statistical tests, as many of the assumptions underlying parametric tests were violated in our data. In the new version of the article, we have added the methodological differences between the two survey versions to the list of reasons for using non-parametric tests (see the change marked with **Comment 2.1.4** on page 6).
5. We now mention the lack of generalizability explicitly in the discussion section (see the change marked with **Comment 2.1.5** on page 8).

Comment 2.2

Does the manuscript provide sufficient discussion with respect to related work?

Your opinion

0.8 - Good description of the background, a couple of references may be added

Comment

A fairly complete exposition of related work, but the details of the studies cited are not sufficiently explained. As for the hypothesis and objectives, I think they should be in the same section. On the other hand, the hypotheses are not very clear in my opinion (“The first hypothesis is derived from the observation that students with a lower proficiency in a foreign language tend to resort to MT more often”). I recommend a new wording of the first hypothesis. In addition, the second hypothesis is not presented. That is, there is not a “the second hypothesis is...”

We agree with the reviewer that some detail could be added to the description of previous findings. Consequently, we have added more detailed information about the frequency of use of MT by language learners in different studies to the literature review. Please see the changes marked with **Comment 2.2.1** on page 2 and 3 of the revised manuscript.

Additionally, we have reformulated both the first and the second hypothesis, in order to set them apart more clearly. Please see the revisions marked with **Comment 2.2.2** on page 3.

Comment 2.3

Does the manuscript adhere to standards in this field for data availability?

Your opinion

0.3 - A hint is given, but no precise procedure is given

Comment

As I have previously explained, there are no details regarding some survey data, such as the questions and the scale neither the platform used for the survey. The authors state: "The students could access the survey via a web link, which redirected them to an online survey environment created using Microsoft Forms." Why Microsoft Forms? Other elements should have been explained, such as informed consent or data protection.

We hope that the additional information provided in response to comments 1.2 and 2.1.2., i.e., the link with access to the data, the SPSS output and the survey questions, resolves the first issue mentioned above about the data, the questions and the scale being used.

We have also tried to explain why two different survey tools were used. Please see the changes marked with **Comment 2.3.1** on page 4.

Finally, we have added specific information about how we have dealt with informed consent and data protection in both surveys. Please see the addition marked with **Comment 2.3.2** on page 4 of the revised manuscript.

We hope that this letter provides sufficient clarification of the modifications we have made in response to the reviewers' comments. If you have any further questions or comments, do not hesitate to contact us.

Yours sincerely,

Joop Bindels
Mark Pluymaekers