GAMETRAPP project in progress: Designing a gamified environment for post-editing research abstracts

Cristina Toledo-Báez
Research Institute on Multilingual Language Technologies
University of Malaga
Spain
toledo@uma.es

Laura Noriega-Santiáñez
Research Institute on Multilingual Language Technologies
University of Malaga
Spain
laura.noriega@uma.es

Abstract

The «App for post-editing neural machine translation using gamification» (GAMETRAPP) project (TED2021-129789B-I00), funded by the Spanish Ministry of Science and Innovation (2022–2024), has been in progress for a year. Thus, this paper presents its main goals and the analysis of neural machine translation and post-editing errors of research abstracts carried out. This leads to the designing of the gamified environment, which is currently under construction.

1 Introduction

The era of artificial intelligence has undoubtedly shaped the evolution and refinement of language technologies, such as neural machine translation (NMT) systems, leading to the widespread adoption of post-editing (PE). Previous studies have explored the application of PE by scholars (Parra Escartín et al., 2017; Parra Escartín and Goulet, 2020; O'Brien et al., 2018) from a first language (L1) to a second language (L2) within academic contexts.

Against this background of scientific dissemination, the GAMETRAPP project aims to create a web application to bring closer the full PE of research abstracts following the IAMRaC structure,[1] in the Iberian Spanish→American English directionality, i.e., from a L1 to a L2. The potential users will be Spanish non-professional translators, specifically scholars.

Moreover, given the promising results of using gamification in non-professional contexts to enhance learning skills (Mahat et al., 2023), the GAMETRAPP app will feature a tailored-made gamified environment.

The GAMETRAPP team is made up by 23 scholars from both Spanish and American universities. Specifically, 7 universities from Spain (University of Malaga, University of Córdoba, University Pablo de Olavide, University of Alcalá, Complutense University of Madrid, University of Valladolid, and Valencia International University) and 2 from United States (Kent State University and Utah Valley University) (Toledo-Báez, 2023).

2 The GAMETRAPP project setup

After a year of its inception, the GAMETRAPP project is currently pursuing the following goals:

1. Analyze the NMT errors of research abstracts, specifically, those made by Google Translate in the Iberian Spanish→American English language combination.
2. Propose full PE guidelines to address the issues derived from the NMT output of research abstracts.
3. Study PE Literacy in the Spanish→English directionality in order to integrate these notions into the gamification.
4. Analyze the experience of Spanish scholars about NMT and PE knowledge and practice.
5. Create gamified activities to teach notions about PE of research abstracts presenting an IAMRaC structure to non-professional translators, specifically, scholars.

[1] This acronym stands for Introduction, Aims, Methodology, Results, and Conclusion. It is a variant of the IMRaD structure (i.e., Introduction, Methodology, Results, and Discussion).

6. Design an inclusive gamified environment to create an engaging, playful, and educational experience among users.
7. Develop a fully responsive web application to integrate the gamified environment.

3 NMT and PE analysis of research abstracts

To design the content of the gamified environment, four linguistic tasks have been carried out. First, the compilation of research abstracts that have been selected from 244 Spanish journals from the Q1 and Q2 in Scimago Journal & Country Rank in 2022. All abstracts have been manually analyzed and labelled considering their parts. Only 126 abstracts followed these three criteria: 1) published in 2023, 2) followed the IAMRaC structure, and 3) written by scholars affiliated with Spanish universities.

Second, Google Translate was chosen as the NMT system since it is widely used by scholars. Third, the selected abstracts were both human-translated into English by a professional translator having English as L1. They were also post-edited into English by a professional post-editor having English as L1.

Last, the fourth task consists of the linguistic analysis of the 126 machine-translated and post-edited research abstracts, using human translation as a gold standard. To carry out this task, the GAMETRAPP team in charge first focused on the resulting machine translation of the abstracts marking in bold red the NMT error, and then classified it based on the MQM (Multidimensional Quality Metrics)[2] errors typology (i.e., Terminology, Accuracy, Linguistic conventions, Style, Locale conventions, Audience appropriateness, Design and markup). Next, they focused on the post-editing of the abstracts marking in bold red the PE error and/or modifications in the segment under review. Once the PE error is identified, they assessed the PE based on the categories proposed by the Post-edit Me! Project[3] (i.e., value adding/successful edits, unnecessary edits, incomplete edits, or unsuccessful/error-introducing/missing edits). Finally, they added any key issues from the evaluation that they deemed pertinent.

This double analysis will not only help to determine several of the most frequent NMT errors for machine-translated research abstracts, but also to develop specific ES→EN PE guidelines of research abstracts and lay foundation for the gamified exercises.

4 The gamified environment

After carrying out the analysis, the gamified environment will be designed based on both the IAMRaC structure and the NMT and PE output. Thus, users will have to complete a series of activities within a gamified escape room experience divided into five levels. Each level represents each of the proposed IAMRaC parts of a research abstract. Users will play individually to unlock the levels, having three lives per level but not time limit.

The gamified activities are designed to train in PE through three stages: 1) the identification of NMT errors, i.e., those of Google Translate NMT system, 2) the practice of PE strategies in context, and 3) the pinpointing of PE errors. These activities are planned to be multiple choice (carried out by analyzing fuzzy matches different PE versions), gap-filling, and/or error correction. Lastly, a questionnaire will be created within the gamified activities to collect users' experiences, and thus test the usefulness of this technique. Thanks to the GAMETRAPP gamified environment, users should have learned full PE notions of research abstracts in the Iberian Spanish→American English language combination to help them post-edit their own academic productions.

Acknowledgments

This work has been carried out in the framework of several research projects: GAMETRAPP (TED2021-129789B-I00), NEUROTRAD (B1-2020_07), VIP II (PID2020-112818GB-I00), Proof-of-Concept VIP (PDC2021-121220-I00), RECOVER (ProyExcel_00540) and T2T (D5-2023_14).

References

Mahat, J., Alias N., and Yusop F. D. 2023. Systematic literature review on gamified professional training among employees. *Interactive Learning Environments*, 31(10):6747–6767.

O'Brien, S., Simard M., and Goulet M. 2018. Machine Translation and Self-Post-Editing for Academic Writing Support: Quality Explorations. In *Translation Quality Assessment. Machine Translation: Technologies and Applications* pages 237–262. Springer.

Parra Escartín, C., and Goulet M. J. 2020. When the Post-Editor is not a Translator: Can machine translation be post-edited by academics to prepare their publications in English? In *Translation Revision and Post-Editing* pages 89–106. Routledge.

Parra Escartín, Carla, Sharon O'Brien, Marie-Josée Goulet, and Michel Simard. 2017. Machine Translation as an Academic Writing Aid for Medical Practitioners. In *Proceedings of MT Summit XVI*, 254–267, Nagoya, Japan.

Toledo Báez, C. 2023. GAMETRAPP: Training app for post-editing neural machine translation using gamification in professional settings. In *Proceedings of the 24th*

[2] Available online at https://themqm.org/error-types-2/typology/.

[3] Available online at https://oer.uclouvain.be/jspui/handle/20.500.12279/829.

Annual Conference of the European Association for Machine Translation, pages 497–498, Tampere, Finland. European Association for Machine Translation.