https://doi.org/10.5007/2175-7968.2022.e80607

MOORKENS, John; CASTILHO, Sheila; GASPARI, Federico & DOHERTY, Stephen (eds.). *Translation Quality Assessment: From Principles to Practice*. Cham: Springer, 2018. 292 p.

Mirella Mota Cavalcante da Silva[1]
Walter Carlos Costa[1]
[1]Universidade Federal do Ceará, Fortaleza, Ceará, Brasil

There is a growing need to adapt to the digital cultural world in different sectors of our technological society. The acquisition of inter and intra-linguistic knowledge and its interface with Digital Information and Communication Technologies (DICTs) reinforces the importance of research in Translation Studies.

Machine Translation (MT) is a sub-area that includes translation and interpretation strategies. About this, Caseli (1783) points out that "although it emerged as a research area more than 70 years ago, MT still presents challenges, such as dealing with slang and colloquialisms […]".

Even so, the investigative nature of this subfield of knowledge and its relationship with human translation (HT) incites intrinsic approaches to translation practices. Integrating the heuristic scenario in MT and its collaborative interfaces, *Translation Quality Assessment: From Principles to Practice,* edited by Joss Moorkens, Sheila Castilho, Federico Gaspari and Stephen Doherty, brings important considerations on Translation Quality Assessment (TQA) to the

current state of the art. MT has been used on virtual platforms by billions of internet users.

The book presents researches that encompass different aspects of MT. The book is organised into twelve chapters, divided into three parts; the first two parts with four chapters, and the third part with three. The first part brings a contextualisation of the TQA scenarios; the second contemplates the application developments of TQA, and the third characterises TQA in practice.

The "Introduction", written by the organisers, presents briefly and clearly, each of the parts and each of the chapters. Starting part one of the book, the chapter "Approaches to Human and Machine Translation Quality Assessment" written by Sheila Castilho, Stephen Doherty, Federico Gaspari, and Joss Moorkens, exposes different approaches about the definition and measurement of quality in human and machine translations in workflows, in educational research settings, and industry. The authors discuss fundamental issues for the consolidation of TQA, through research, teaching and professional practice. They underline the centrality of the translation process and observe that evaluating the quality of its results questions the very concept of translation quality.

In this introductory chapter, the authors present a history of TQA in the field of Translation Studies, from the initial approaches of theorists such as Nida (1964), Catford (1965), Reiss (1971), Reiss and Vermeer (1984), Baker (1992), House (1997), Schäffner (1997), Secară (2005), Lauscher (2000), Pym (2010), Fields et al. (2014), among others, to support the analyses submitted. In addition, the chapter relates the theoretical conceptions to the types of software and language service providers in the face of the TQA process. The authors cite the *Localization Industry Standards Association* (LISA) as a model in TQA, discussing its use in various adaptations, although its operations ceased in 2011. The *Translation Automation User Society* (TAUS), a *think tank* of the translation

industry, has been trying to develop effective QAT indicators. In parallel, the QTLaunchPad, a collaborative research project funded by the European Union (EU). Regarding TQA realised by humans, the authors list the key issues in reference to evaluators, such as adequacy and fluency; readability and comprehensibility; acceptability; ranking, usability, and performance. As for the automatisation of TQA, the main goal of state of the art is to compare the output of an MT system for one or several reference translations, which are assumed to be good since they come from humans. The chapter presents the collaborative relationship between the human and machine translations and the effort in post-editing situations. Among the tools for data collection, there are *Post- Editing Tool* (PET, Aziz et al. 2012), *Translog-II* (Carl 2012), *Cognitive Analysis and Statistical Methods for Advanced Computer-Aided Translation* (Casmacat, Alabau et al. 2013).

The next chapter, "Translation Quality, Quality Management and Agency: Principles and Practice in the European Union Institutions", by Joanna Drugan, Ingemar Strandvik and Erkka Vuorinen, shows how translation quality is managed in practice, identifying two guiding principles: consistency of approach and consistency of quality involving the EU project scenario. Then, the authors point out a brief overview of procedural developments in European translation. Thus, the inter-institutional operationalisation of the European Commission's Directorate-General for Translation (DGT) is highlighted to support quality management. Therefore, the Total Quality Management (TQM) approach is explained, with emphasis on valuing communication between suppliers, customers, employees and managers, to further support the functional performance of *freelance* translators. In fact, by pointing out the EU model and highlighting the exponential guidelines of TQM, the reading brings reflections on the importance of the translator figure.

The chapter "Crowdsourcing and Translation Quality: Novel Approaches in the Language Industry and Translation Studies" by

Miguel A. Jiménez-Crespo provides relevant information to understand the *crowdsourcing* organisations in innovative workflows. The author presents the scope of crowdsourcing in translation contexts, presenting investigations of its impact on the industry and translation studies. Thus, one-off questions are reported about the consolidation of process-based approaches and the perspective of these workflows to incite the management of desirable levels to quality. The author emphasises that each type of translation is associated with a different kind of process, with perceptive variables, such as crowdsourcing composition and the presence or revision of translations. Therefore, the adequacy of the purpose models and the diversification of the translation quality is related to post-editing quality. The author refers to classical approaches, in which the Language Service Provider (LSP), institutions, translators and qualified reviewers controlled the quality of translations. For him, due to the diverse nature of crowdsourcing, the responsibility for quality goes beyond Lsps, translation providers and translators themselves. So, the examples of Facebook and Twitter, where the quality of translations tends to be the responsibility of community managers who oversee the whole process of these collaborative flows, raise reflections on the relationships of *crowds* and their professional paradigms.

In sequence, in the chapter "On Education and Training in Translation Quality Assessment" by authors Stephen Doherty, Joss Moorkens, Federico Gaspari and Sheila Castilho, there is the parallel of the translators training and the translation services in the social sphere in the face of the interaction of theory and practice. The authors argue that education and training in TQA have been neglected by most interested in translation processes. As for academia, they point out that there is a lack of information for translation students, even at the graduate level, about the knowledge and skills needed to understand and use TQA models, with immediate consequences on their long employability run, on their professional practice.

Starting part two of the book, the chapter "Metrics for Evaluating Translation Quality: A Case for Standardising Error Typologies" by Arle Lommel provides an overview of three TQA systems: (1) the Multidimensional Quality Metrics (MQM) framework developed at the German Research Center for Artificial Intelligence, the DFKI, in Berlin; (2) Dynamic Quality Framework (DQF) error typology of the Amsterdam-based TAUS; (3) Harmonisation of the previous two, carried out as a collaborative effort between the DFKI and TAUS within the EU-funded QT21 project in 2014 and 2015. The author explains that these projects aim to improve the current state of the art for TQA, in addition to solving the lack of best practice approaches in this area. Thus, since 2015, the integrated typology of DQF/MQM errors has become the preferred method for MQM implementation, and its inclusion in the DQF raises the profile of TQA approaches. Although it is the central axis of this chapter, it is worth asking how far a thorough knowledge of error metrics and typologies is necessary for the professional practice of the translator.

Then, Maja Popović, in the chapter "Classification and analysis of errors for Evaluation of the quality of machine translation", presents approaches and tasks related to the classification and analysis of errors of MT results. For the author, error analysis is an active field of research, and its complexity and variety lead to the development of innovations. The chapter presents a survey of studies with results on data and MDM about the typologies of errors in MT and quality in post-editing. For Popović, error classification can occur manually, automatically, and in combination. Considering the praxis of post-editing, the reading registers, according to the data collected, that automatic error classification produces more reliable results when the post-edited MT output is used as a reference translation. Furthermore, according to the types, the technical effort for editing operations is strongly related to the estimated cognitive cost, regardless of the temporal action.

In “Quality Expectations of Machine Translation”, Andy Way discusses the usefulness of MT, describing how it has been deployed, how its production is evaluated, and how this could be improved. For Way, not everyone is convinced that MT can be useful, especially in relation to improving the productivity of human translators. The author assumes that quality metrics, such as Bilingual Evaluation Understudy (BLEU), are not discriminatory enough to reflect the translation quality of systems accurately - Neural Machine Translation (NMT), i.e., neural machine translations, character-based. Therefore, given the comments set out in this reading, we believe that the logistics of translation technologies tangents its complex structural character, but does not detract from the usefulness of MT as a computational system and strengthens the importance of collaborative integration between human and automatic translations.

Concluding part two of the book, the chapter “Assessing Quality in Human- and Machine-Generated Subtitles and Captions”, by Stephen Doherty and Jan-Louis Kruger, presents TQA in Audiovisual Translation (AVT) and Language Technologies, including computer-assisted translation tools, machine translation, automated subtitling and captioning software, and automatic speech recognition systems. Based on an informative range of new industry, and moving in the direction of cutting-edge research on the processing and reception of AVT products in various media and languages, authors provide a critical view of TQA issues in human and machine-generated subtitling and captioning. The chapter then describes industry parameters for accuracy, presentation and timing, so that subtitles and captions have good visual quality for viewers, regardless of language or accessibility needs.

The chapter “Machine Translation and Self-post-editing for Academic Writing Support: Quality Explorations”, by Sharon O’Brien, Michel Simard, and Marie-Josée Goulet, presents an exploratory study about the potential of AT and self-post-editing (PE) to support the academic writing process of authors who have English as

a Foreign Language (EFL). This study aimed to verify if MT is a valuable aid for academic writing and its impact on the quality of the written product. The chapter offers a literature review of the following themes: English as a *lingua franca* in academic writing and the consequences this may have for individual authors and subjects; second-language writing; and the use of MT as a second language writing aid. The results suggest that MT and self-post-editing had no negative impact on the text produced and indicate that the use of MT can be an aid to EFL scholars, offering advantages such as reduced cognitive load.

Finally, in part three and the last chapter of the book, "What level of quality can Neural Machine Translation achieve in a Literary Text? ", Antonio Toral and Andy Way examine the neural approach to machine translation - NMT, and the challenges of MT for literary texts. The authors research the novel genre and present an NMT system adapted to literature in translation from English to Catalan. Using MT, NMT and PBSMT on a large corpus and through the BLEU metric, the researchers included 12 widely known novels from the 1920s to the present day. They concluded that NMT is significantly better than PBSMT ($p < 0.01$) in all the novels considered. These authors point out, in the chapter, according to the data collected, that NMT results in a relative improvement of 11% concerning PBSMT.

Translation Quality Assessment: From Principles to Practice covers a wide range of issues in MT. It is a dense work that deepens the analysis of the relations between human and machine translation, clarifying the evaluative strategies in the extensive sub-area of TM in Translation Studies.

References

Caseli, Helena, de Madeiros. “Tradução Automática: estratégias e motivações”. *Domínios de Lingu@gem*, v. 11, n. 5, (2017): 1782-1796. Disponível em: https://doi.org/10.14393/DL32-v11n5a2017-21. Acesso em 20 de Mar. de 2020.

Moorkens, John; Castilho, Sheila; Gaspari, Federico & Doherty, Stephen (eds.). *Translation Quality Assessment: From Principles to Practice*. Cham: Springer, 2018. 292p.

Recebido em: 04/10/2021
Aceito em: 04/12/2021
Publicado em fevereiro de 2022

Mirella Mota Cavalcante da Silva Email: mirellamota10@gmail.com. http://orcid.org/0000-0003-3261-881X.
Walter Carlos Costa. Email: walter.costa@gmail.com. https://orcid.org/0000-0001-5853-0950.