
Translation Quality Assessment: From Principles to Practice offers a broad coverage of a number of approaches to human and machine translation quality assessment (TQA) based on a wide range of practices from research, industry and academia. The book appears in Volume 1 of the Machine Translation: Technologies and Application series, edited by Andy Way, whose aim is to reflect current state-of-the-art practices and developments in machine translation. With a focus on “the product, rather than the process of, translation” (1), the book under review is divided into three thematic areas. The first part presents human and machine TQA scenarios, with a focus on different types of evaluation metrics and issues. The second part examines applications developed for TQA. The third part discusses practical implementations of quality assessment processes in three different contexts based on empirical studies.

In Part I, Castilho et al. provide a very useful overview of TQA which reflects TQA practice in both academia and the industry, by introducing readers to quality measures, automatic metrics, and the integration of assessment in translation workflows. This chapter could be suggested as study material for students, researchers and practitioners who need an overview of the topic. Drugan et al. provide an excellent account of translation quality evaluation processes in the European Commission’s Directorate-General for Translation (DGT). The chapter presents a very well-organised study that offers readers a comprehensive view of the topic within its practical settings. The analysis focuses on consistency of approach and consistency of quality with regard to the implementation of TQA at inter- and intraintitutional levels in the complex EU translation setting. The examples provided explain how relevant policies are implemented within DGT and offer a very clear account of the translation and TQA workflows. Jiménez-Crespo looks into crowdsourced translation workflows and how quality is guaranteed in such workflows. The author provides a necessary definition of translation crowdsourcing in order to discuss quality accordingly, and looks into the very interesting topic of fitness for purpose whereby “quality is fully conceptualised as a scalable commodity” (79). The impact of crowdsourcing on quality evaluation is emphasised, with the author highlighting the importance of considering the mechanics of the language industry and the responsibility shared among translation stakeholders. Following this insightful chapter, Doherty et al. focus on TQA training while acknowledging how neglected the topic of assessment is in academia and how this affects translation students in terms of employability,
as the industry seems to gradually become more dependent on quality metrics. The authors highlight the need for the development of training in TQA in specific translation and educational contexts, expanding the currently observed focus on standardised TQA metrics and models and the rather theoretical approaches to them.

Part II starts with a chapter by Lommel, who provides a detailed account of Multidimensional Quality Metrics and an overview of the Dynamic Quality Framework, making a useful contribution to the study of systematic quality evaluation. Popović’s chapter focuses on manual, automatic and semi-automatic machine translation error classification and analysis, while the author also discusses the case of unmatched patterns and the evaluation of specific linguistic features. Section 4.4 of this chapter on the challenges involved in further development of such approaches opens a fruitful discussion on the analysis of language-related phenomena that cannot be handled adequately by machine translation systems and shows a lot of potential for further analysis. With the complexity that characterises the topic and considering the blurred boundaries between human and machine translation in current workflows, content from Way’s chapter could be suggested as introductory reading for those with limited background in machine translation, before they embark on the close study of assessment in Part I. Way manages to examine the complex nature of machine translation and its usefulness for human translators and post-editors and to discuss future expectations, while highlighting the importance of the human factor in the evaluation process. In the last chapter of this Part, Doherty and Kruger cover a number of challenges involved in the evaluation of human- and machine-captioned/subtitled material, also considering the domination of new media and the lack of empirical guidelines as opposed to standardised ones, making an interdisciplinary approach necessary in order to approach issues of quality in audiovisual translation.

Part III includes three empirical studies on different applications of TQA. Specia and Shah provide an overview of machine translation quality estimation, describe possible applications, and offer results from a comparison between random and quality estimation-based samples. The outcomes provide useful insights into ways in which quality estimation can help predict machine translation quality, while they also indicate that this area requires further investigation, considering how inconclusive results can be when the criteria used are more generic, such as quality levels. O’Brien et al. provide a comprehensive account of a study conducted in order to evaluate the usefulness of machine translation for academic writing and how it can affect the quality of the produced material. The presentation of the study is followed by interesting remarks indicating positive outcomes and the potential to
expand this study to different languages. The authors also hint the potential of extending the current sentence-level, specific scientific terminology and discourse approaches in the analysis of outcomes to an evaluation of the faithfulness of the translation output with regard to the genres of the source texts using a variety of evaluation modes. Finally, Toral and Way compare the output of a neural machine translation system with that of a dominant statistical phrase-based machine translation system when translating English literary texts into Catalan. This study showed significantly more fluent translation output with neural machine translation, based on both automatic and human evaluation.

With coverage of a wide range of interesting topics, this book is a very useful guide for students, researchers and practitioners in the field, including very interesting contributions that can inspire the reader in various directions.

Emmanouela Patiniotaki
University College London
E-mail: e.patiniotaki@ucl.ac.uk