Translation Quality Assessment in Translation Teaching

Xiangyan Zhou, Yaru Cai

Abstract
Translation quality assessment, an indispensable part of translation teaching, has drawn much attention in translation studies. To get an overview of the topic, this study systematically reviews previous research on translation quality assessment to summarize achievements and point out avenues for future research. Bibliometric analysis was used to analyze relevant journal articles from the Web of Science, a publisher-independent global citation database. The results demonstrated that existing studies mainly discussed translation quality assessment procedures, assessment methods, and assessment tools. The study can provide a reference for quality assessment in translation teaching and research.

Key Words: translation quality assessment; procedures; methods; tools; translation teaching

1. Introduction
Nowadays, the number of universities with translator-training programs is increasing in China, reflecting greater demand for skilled translators. Besides translator-training programs, translation courses are also offered to students admitted to non-translation programs in tertiary institutions. Moreover, translation has become an integral part of the College English Test (Band 4 and Band 8), a national test designed to check the English proficiency of undergraduate students in China (Liu & Zheng, 2022). Such phenomena have demonstrated the importance of translation teaching for universities. Translation teaching in China has always been criticized for its derailment with the market; the trained translation students must be sufficiently qualified for the industry (Li & Xiao, 2020). This highlights the role of translation quality assessment (TQA), the process of judging a translator’s work and giving relevant recommendations. However, TQA remains a thorny problem partly due to philosophical issues concerning quality. Other possible reasons include the complexity of human languages, the diversity of the tasks considered, and the situations in which they are undertaken. In this regard, translation quality assessors are supposed to possess a high degree of analytic ability and discernment. Despite its importance and challenging nature, a systematic review of relevant research needs to be done. Under such a backdrop, the current study aims to investigate the topic using Bibliometrics to provide a reference for future research.

2. Research Method and Data Source
This study analyzes plenty of TQA articles through Bibliometrics, intending to have a holistic understanding of the topic. Bibliometric analysis is a method that synthesizes a large amount of literature to provide a general overview of a topic. Relevant articles are obtained from the Web of Science. With such keywords as “translation quality” AND “quality assessment” entered into the search bar, the search engine produced 72 search results. After carefully screening the articles by reading their abstracts, the study kept 25 articles to investigate the topic.

3. Results and Discussion
Details of the 25 articles are listed in Table 1 below. These articles were classified into three groups according to their significant content: assessment methods, tools, and procedures. Of them, the number of articles dealing with TQA methods ranked first.

<table>
<thead>
<tr>
<th>S. N.</th>
<th>Themes</th>
<th>No. of articles</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>Methods</td>
<td>15</td>
</tr>
<tr>
<td>2</td>
<td>Tools</td>
<td>5</td>
</tr>
<tr>
<td>3</td>
<td>Procedures</td>
<td>5</td>
</tr>
</tbody>
</table>

3.1 TQA Methods
So far, translation researchers have proposed a wide range of TQA methods, most of which are subjective and rely heavily on human judgment. Subjective TQA approaches have perplexed many students, resulting in a big gap between high-quality translations and those produced by students. As a branch of translation studies, TQA has been discussed from different perspectives for a long time.
However, the presuppositions of various models conflict with each other, and their internal theoretical structures have apparent defects, which are inconsistent with the actual situation of translation assessment (Liu, 2018). TQA methods are discussed in detail in the following paragraphs.

First, the error analysis approach is the method that has drawn the most scholarly attention and has been widely mentioned in educational and professional settings over the past decades. The creation and use of error analysis mark an essential milestone in TQA, as it represents a programmatic effort to pursue objective and reliable assessment (Han, 2020). This approach aims to provide assessors with systematic and consistent classification models to reduce the subjective nature of the assessment process (Martikainen, 2017). Error analysis is objective, systematic, and reliable because of its clear quality standards, detailed error types, and quantification of evaluation results. Besides, it is open and transparent because the evaluation process is predetermined and known to all stakeholders. Therefore, error analysis remains a preferred method in the translation industry.

There is a fundamental dichotomy between error-based assessment and holistic assessment. An anecdotal experiment (Koby & Lacruz, 2017) points to a whole set of issues in TQA: how to score individual errors, how to group them into an overall grade, how to set a threshold for acceptability, and whether to break down any score into separate sub-scores. Even though the weight of anecdotal evidence is light, the above issues are well worth our attention. Therefore, the error-based assessment may be more suitable for high-stakes examinations for licensed translators where accuracy is deemed essential. If total accuracy means near-zero errors, then an argument can be made for preferring error-based assessment over holistic assessment.

The holistic method is regarded as being accurate and objective. However, its flexibility and subjectivity are limited (Bahameed, 2016), depending on the raters’ judgment and the kind and the number of translation errors the candidate produces, which makes it different from the error analysis approach. It cannot quickly tell the diligent top respondents. It can give many students the same high score, which may eliminate the differences in grades and thus reduce student motivation toward translation. This approach is so lenient that there need to be more opportunities to see individual differences among these top students. This method needs to clarify how many minor errors can be overlooked and forgiven by the assessors. Suppose the holistic approach is accused of being too lenient. In that case, it remains solid only because it keeps the failure rate within a reasonable range and is applied impartially to all students without distinction.

After searching dozens of articles, most have mentioned the mixed-methods approach. As the error analysis approach and rubric scoring represent the micro and macro textual perspectives of TQA, the combination of the two makes TQA more objective and credible. The advantages of rubrics are as follows: They provide a reference framework for the assessor that facilitates his/her decision-making based on limited, well-known, and transparent criteria which limit the subjective burden inherent to any assessment process with human intervention (Martínez-Mateo et al., 2017). Besides, empirical evidence suggests that the mixed-methods approach could be more reliable and practicable than error analysis and rubric scoring used alone. Including multiple assessment methods can help assessors learn more about evaluation criteria, and they can, in turn, combine the advantages of each method.

The corpus-based translation evaluation is used as a benchmark against which assessors can compare and judge the appropriateness of lexical and phrasal choices in multiple translation versions. Nevertheless, there are three barriers to its widespread use. First, it requires technical expertise to compile a reference corpus, and assessors need the necessary training to use it correctly. Second, maintaining and expanding a universal corpus takes time and effort. Besides, each assessment method, including the corpus-based method, is limited by its inherent weaknesses. It cannot entirely automatically evaluate translation quality, which means that it requires statistical knowledge for assessors to analyze the results. Namely, it has succeeded in maximizing its reliability and practicality but may not stand up to the test of effectiveness and may prove less useful than it has claimed. Therefore, advocates of the corpus-based method argue that it cannot replace traditional evaluation methods but can still serve as a valuable tool in translation teachers’ toolbox to offer supplementary information in TQA (De Sutter et al., 2017). Therefore, to reduce the inconvenience brought by the corpus, it is necessary to establish a corpus free of charge that can be used universally and allow supervisors to supplement the corpus on a nationwide network. In addition, qualified universities can create their corpus to facilitate the use of corpus by their teachers and students.

Many existing scoring criteria for translation quality are based on experience and anecdotes (Colina, 2009). They tend to be developed for particular professional organizations or industries, such as ATA certification.
exams (Colina, 2008). In many cases, as we can find, they often need to be more relaxed about transferring to other environments: Marking scales are based on experience instead of on an explicit theoretical framework. Besides, some frustrating deficiencies make some TQA methods challenging to apply in professional and pedagogical situations; for example, they focus merely on partial aspects of quality, whereas assessment needs to include a wide range of issues.

3.2 TQA Tools

Regarding digital texts, the crowdsourced quality evaluation model needs to be mentioned. Many scholars have noted that developing this kind of novel evaluation tool could become a frequently used method since the production of digital content such as websites, software, or video games has increased exponentially (Jiménez-Crespo, 2011). Also, the crowdsourced quality evaluation model proves that technology’s impact could fundamentally change how translation is practiced in previously never imagined ways. Students can evaluate their translations without asking their teacher for help, which can save unnecessary human and material resources. At the same time, it could indeed offer the opportunity and arouse the desire of students to study independently. Therefore, this online marking method should be widely generalized and implemented in this era. Besides, it is significant to strengthen the supervision of this method to ensure the accuracy of software assessment.

Another study reveals that the scale method mainly used to assess machine-generated translations is also a valid tool to assess human translations. The measurement was accepted in Taiwan and applied in the 2007 public translation proficiency test (Lai, 2011). After comparing the scales with the error-analysis method, the researcher revealed that the scale method was reliable to prove the scales as a valid instrument for assessing translations in the exam.

Moreover, a new approach has been mentioned (Martínez-Mateo et al., 2017)—the Quality Assessment Tool (QAT), a prototype quantitative tool developed by the Directorate General for Translation (DGT). Using an assessment tool that combines qualitative and quantitative models allows raters to evaluate translated texts with a unified frame of reference. It provides a more balanced view of the texts. It discusses two necessary and complementary perspectives: The tool could be adapted to other language combinations or to accommodate other professional settings by considering its excellent benefit-cost ratios.

3.3 TQA Procedures

Evaluation includes changing the source text to suit the culture of the target readers, stylistically modifying the style of the target language, and presenting the layout of the translated text in a proofread and formatted manner. Assessment needs to be seen as a specialized task that requires highly qualified translators who are familiar with the particularities of the translation work. In addition, in principle, assessors shall be able to fully understand the source text and the translation process and navigate it with complete ease (Chakhachiro, 2005).

Students were asked to hand in the translations to their teacher, who would treat them as regular homework assignments. To minimize subjective bias in grading, the teacher can exchange translations, sort the texts into different groups, and focus on selected aspects. Regarding College English Text (Band 6), the teacher should combine coherence and cohesion (Li & He, 2015). The assessment should cover the following dimensions: content, coherence, syntax, grammar, and vocabulary. Besides, four levels can be distinguished, with the lowest level indicating poor performance or a lack of evaluative message, the highest level indicating excellent performance, and the other two indicating average performance and good performance.

4. Conclusion

Through Bibliometrics, most TQA studies paid attention to TQA methods and provided their insights. Meanwhile, some studies addressed assessment tools or procedures. The various assessment methods, tools, and procedures mentioned herein can enable TQA specialists to understand each method critically for better decisions in future assessment practices and, in turn, put all these methods to use better. Besides, the study can also help students improve their translations’ quality and understand the translation practice’s refinement. Given the lack of visualization tools and the limited number of raw data used in the current study, it is hoped that more attempts are made to enrich this line of research.

References


