LINGUOPRAGMATIC CRITERIA FOR ASSESSING THE QUALITY OF ARTIFICIAL INTELLIGENCE-GENERATED TRANSLATIONS

Aytmuxamedova Durdona Jasur qizi
Tayanch doktorantura talabgori
Toshkent, O'zbekiston
E-mail: aytmuxamedovadurdona@gmail.com
Tel:(+99890)3294410

Abstract

The rapid advancement of artificial intelligence (AI) has significantly transformed contemporary translation practices by enabling fast and accessible multilingual communication. Neural machine translation systems such as ChatGPT, Google Translate, and DeepL have achieved remarkable progress in producing grammatically accurate and semantically coherent translations. Nevertheless, these systems still experience considerable difficulties in preserving pragmatic meaning, communicative intention, discourse coherence, and culture-specific elements. Consequently, evaluating AI-generated translations requires criteria that extend beyond lexical and grammatical accuracy. This paper examines linguopragmatic criteria for assessing the quality of artificial intelligence-generated translations. Special attention is paid to pragmatic equivalence, communicative intention, discourse coherence, cultural adaptation, and contextual appropriateness as essential indicators of translation quality. The study argues that linguopragmatic evaluation provides a more comprehensive framework for assessing AI-generated translations and contributes to improving post-editing practices and translation quality.

Keywords: artificial intelligence, translation quality assessment, linguopragmatics, pragmatic equivalence, post-editing, machine translation.

Introduction

Artificial intelligence has become one of the most influential technologies in modern translation studies. The development of neural machine translation (NMT) has significantly increased the efficiency of multilingual communication by enabling the rapid production of translations in numerous language pairs. Contemporary AI systems, including ChatGPT, Google Translate, DeepL, and Microsoft Translator, demonstrate high levels of grammatical accuracy and lexical fluency. As a result, machine translation has become an indispensable tool in education, business, science, tourism, and international communication.

Despite these achievements, evaluating the quality of AI-generated translations remains a complex issue. Traditional approaches to translation quality assessment primarily focus on lexical accuracy, grammatical correctness, and semantic equivalence. However, these criteria alone are insufficient for evaluating machine-generated translations because they often fail to identify pragmatic inaccuracies that influence communication. Artificial intelligence systems frequently produce translations that are linguistically acceptable but pragmatically inappropriate due to their limited understanding of context, speaker intention, discourse conventions, and cultural norms. Recent developments in translation studies emphasize that translation quality should be evaluated not only from a linguistic perspective but also from a communicative and functional perspective. Linguopragmatics provides an effective theoretical framework for assessing whether a translated text successfully conveys communicative intention, pragmatic meaning, interpersonal relations, and discourse functions. Consequently, linguopragmatic criteria have become increasingly important in evaluating the quality of artificial intelligence-generated translations.

The relevance of this study is determined by the growing application of artificial intelligence in professional translation and the increasing need for reliable evaluation criteria capable of identifying communicative and pragmatic deficiencies in machine-generated translations. While numerous studies have investigated the technical performance of AI translation systems, relatively limited attention has been devoted to linguopragmatic criteria for assessing translation quality, particularly in English–Uzbek translation. The aim of this study is to examine linguopragmatic criteria for assessing the quality of artificial intelligence-generated translations and to determine their significance in improving translation evaluation and post-editing practices.

Translation Quality Assessment in Contemporary Translation Studies

Translation quality assessment (TQA) has long been recognized as one of the central issues in translation studies. The concept of translation quality has evolved considerably, moving beyond the evaluation of grammatical accuracy and lexical correspondence toward a more comprehensive analysis of communicative effectiveness, functional equivalence, and pragmatic appropriateness. Contemporary translation scholars argue that the quality of a translation should be assessed according to its ability to fulfil the communicative purpose of the source text while meeting the linguistic and cultural expectations of the target audience (House, 2015; Baker, 2018).

The rapid development of artificial intelligence has introduced new challenges to translation quality assessment. Neural machine translation (NMT) systems have significantly improved translation fluency and lexical accuracy through deep learning technologies and large language models. Nevertheless, these systems continue to demonstrate limitations in interpreting contextual information, speaker intention, discourse conventions, and culture-dependent meanings. Consequently, conventional quality assessment models that primarily

evaluate lexical and grammatical accuracy are insufficient for measuring the communicative effectiveness of AI-generated translations.

Within contemporary translation studies, quality assessment is generally based on several fundamental principles, including **accuracy, adequacy, fluency, acceptability, coherence, and functional equivalence**. Accuracy refers to the correct transfer of semantic content from the source text to the target text, whereas adequacy evaluates the extent to which the translation fulfils the communicative purpose of the original message. Fluency concerns the naturalness and readability of the target text, while acceptability reflects its conformity with the linguistic norms of the target language. However, these criteria alone cannot fully explain pragmatic deficiencies frequently observed in AI-generated translations.

House (2015) argues that translation quality should be assessed through functional-pragmatic equivalence, which requires preserving not only semantic meaning but also interpersonal relationships, communicative intention, register, and discourse function. Likewise, Baker emphasizes that successful translation depends on achieving equivalence at multiple linguistic levels, including pragmatic equivalence, where implied meanings, presuppositions, and contextual interpretations play a decisive role. These theoretical perspectives provide a solid foundation for evaluating artificial intelligence-generated translations from a linguopragmatic perspective. The emergence of large language models has further increased the need for comprehensive quality assessment frameworks. Although AI systems frequently produce grammatically accurate and coherent texts, they may fail to recognize indirect speech acts, politeness strategies, figurative language, cultural references, and contextual implications. Consequently, translations that appear linguistically correct may still fail to achieve communicative equivalence. This phenomenon demonstrates that translation quality should be evaluated not only through linguistic criteria

but also through linguopragmatic analysis, which considers language as a tool of communication rather than merely a system of grammatical structures.

Therefore, contemporary translation quality assessment increasingly integrates linguistic, functional, pragmatic, and intercultural dimensions into a unified evaluation framework. Such an approach is particularly important for assessing artificial intelligence-generated translations because it enables researchers and professional translators to identify communicative deficiencies that cannot be detected through purely lexical or grammatical analysis. As artificial intelligence continues to develop, linguopragmatic evaluation is expected to become one of the principal methodologies for measuring translation quality in both academic research and professional translation practice.

Linguopragmatic Criteria for Assessing AI-Generated Translations

The increasing application of artificial intelligence in translation has highlighted the necessity of developing comprehensive evaluation criteria capable of measuring not only linguistic accuracy but also communicative effectiveness. Since translation is fundamentally a communicative activity, the quality of AI-generated translations should be assessed according to their ability to preserve the pragmatic functions of the source text. Linguopragmatic evaluation therefore represents a multidimensional approach that considers contextual meaning, speaker intention, discourse organization, cultural adaptation, and communicative impact alongside traditional linguistic criteria.

One of the primary linguopragmatic criteria is **pragmatic equivalence**. Baker defines pragmatic equivalence as the successful transfer of implied meaning, communicative intention, and contextual interpretation from the source language into the target language. Unlike lexical equivalence, pragmatic equivalence requires translators to consider how the target audience interprets utterances within specific communicative situations. Artificial intelligence systems frequently produce literal translations that preserve lexical meaning

while failing to reproduce pragmatic functions. Consequently, target texts may become grammatically accurate but communicatively ineffective.

Another essential criterion is the preservation of **communicative intention**. Every text is created with a specific purpose, whether to inform, persuade, request, warn, criticize, or entertain. According to Speech Act Theory developed by Austin (1962) and Searle (1969), utterances perform communicative actions rather than simply conveying information. AI-generated translations occasionally misinterpret indirect requests, recommendations, suggestions, or polite expressions because they prioritize lexical correspondence over communicative function. Therefore, evaluating whether the target text accurately reflects the speaker's intention constitutes an important aspect of linguopragmatic assessment.

Contextual appropriateness also plays a crucial role in evaluating translation quality. The interpretation of meaning depends not only on linguistic structures but also on situational context, shared background knowledge, and discourse environment. Machine translation systems frequently process sentences in isolation, which may result in the incorrect interpretation of ambiguous expressions, personal references, or context-dependent meanings. Linguopragmatic evaluation therefore examines whether the translated text remains appropriate within its communicative context and accurately reflects the relationships between participants.

Another important criterion involves **cultural adaptation**. Translation is not limited to transferring words from one language into another; it also involves transferring cultural meanings and social conventions. English and Uzbek differ considerably in politeness strategies, forms of address, idiomatic expressions, and discourse conventions. AI systems often translate such elements literally without adapting them to the sociocultural expectations of the target audience. Consequently, linguopragmatic assessment should evaluate the degree to which

culture-specific expressions are rendered naturally and appropriately in the target language.

Discourse coherence represents another fundamental criterion. High-quality translations should demonstrate logical organization, cohesive relationships, and consistent information flow throughout the text. Although artificial intelligence systems generally produce grammatically well-formed sentences, they sometimes fail to maintain coherence across larger discourse units. Inconsistent terminology, abrupt thematic shifts, inappropriate connective devices, and inaccurate reference tracking may reduce the communicative quality of AI-generated translations. Therefore, discourse-level analysis constitutes an essential component of linguopragmatic evaluation.

Finally, the assessment of AI-generated translations should include the analysis of **presupposition** and **conversational implicature**. According to Grice (1975), successful communication depends on the recognition of implied meanings that are not explicitly expressed in linguistic form. Similarly, Yule (1996) emphasizes that presuppositions reflect shared background knowledge between interlocutors. Since artificial intelligence systems primarily rely on statistical prediction rather than genuine pragmatic reasoning, they frequently overlook these implicit aspects of communication. As a result, translations may accurately reproduce explicit information while distorting or omitting implicit meaning.

Taken together, these linguopragmatic criteria provide a comprehensive framework for evaluating artificial intelligence-generated translations. Unlike purely linguistic assessment models, this approach considers translation as an act of intercultural communication in which contextual meaning, communicative intention, discourse coherence, and cultural appropriateness are equally important. The integration of these criteria contributes to more objective translation quality assessment and provides a reliable theoretical basis for linguopragmatic post-editing in contemporary translation practice.

The Role of Human Evaluation in AI Translation Quality Assessment

Although artificial intelligence has significantly enhanced the efficiency of machine translation, human evaluation remains an indispensable component of translation quality assessment. Modern AI systems are capable of generating fluent and grammatically acceptable translations; however, they are still unable to fully interpret contextual meaning, communicative intention, cultural references, and pragmatic nuances. Consequently, the assessment of AI-generated translations cannot rely exclusively on automatic evaluation metrics or linguistic accuracy but requires expert human judgment based on linguopragmatic principles.

Traditional automatic evaluation metrics, including BLEU, METEOR, TER, and COMET, primarily measure lexical similarity and structural correspondence between machine-generated translations and reference translations. Although these metrics provide valuable quantitative information, they cannot adequately assess pragmatic equivalence, discourse coherence, politeness strategies, or communicative effectiveness. Therefore, recent studies in translation quality assessment emphasize the importance of combining automatic evaluation with human-centered qualitative analysis. From a linguopragmatic perspective, human evaluators examine whether the translated text successfully fulfils the communicative function of the source text. This process involves assessing the preservation of communicative intention, pragmatic meaning, contextual appropriateness, and cultural acceptability. Unlike artificial intelligence systems, professional translators are capable of recognizing implicit meanings, interpreting presuppositions, identifying conversational implicatures, and adapting discourse to the sociocultural norms of the target audience. These competencies enable human evaluators to identify translation deficiencies that remain undetected by automatic evaluation systems.

Another important aspect of human evaluation concerns **post-editing**. According to ISO 18587:2017, post-editing aims not only to eliminate linguistic

errors but also to improve the overall communicative quality of machine-generated translations. Professional post-editors analyze lexical choice, grammatical accuracy, stylistic consistency, discourse organization, and pragmatic appropriateness before producing the final version of the translated text. Consequently, linguopragmatic evaluation serves as an essential methodological foundation for effective post-editing practice. The increasing integration of artificial intelligence into professional translation has also transformed the role of translators. Contemporary translators no longer function solely as language converters but also as evaluators, editors, and intercultural mediators. Their responsibilities include identifying pragmatic mismatches, correcting contextual inaccuracies, ensuring cultural adaptation, and preserving communicative effectiveness. These tasks require advanced linguistic knowledge, intercultural competence, critical thinking, and professional experience that artificial intelligence systems cannot yet fully replicate.

Furthermore, the development of reliable linguopragmatic evaluation criteria contributes to improving both translator training and translation technology. Clearly defined assessment standards enable translators to evaluate AI-generated translations more objectively while simultaneously providing researchers and software developers with valuable information for enhancing machine translation systems. Consequently, linguopragmatic evaluation benefits not only professional translation practice but also the future development of artificial intelligence technologies.

Overall, the combination of artificial intelligence and human expertise represents the most effective approach to achieving high-quality translation. While AI systems considerably increase translation speed and productivity, human evaluators ensure communicative accuracy, pragmatic equivalence, and cultural appropriateness. Therefore, linguopragmatic assessment should be regarded as an essential component of contemporary translation quality

evaluation and a key direction for future research in artificial intelligence-based translation.

Conclusion

The findings of this study demonstrate that the evaluation of artificial intelligence-generated translations requires a multidimensional approach that extends beyond traditional linguistic assessment. Although contemporary AI-based translation systems have achieved substantial progress in lexical accuracy, grammatical correctness, and translation fluency, they continue to encounter significant challenges in conveying pragmatic meaning, communicative intention, discourse coherence, and culture-specific elements. These limitations confirm that linguistic accuracy alone cannot guarantee translation quality. The analysis has shown that linguopragmatic criteria provide a comprehensive theoretical framework for assessing the communicative effectiveness of AI-generated translations. Pragmatic equivalence, communicative intention, contextual appropriateness, cultural adaptation, discourse coherence, speech acts, presupposition, and conversational implicature represent essential indicators for evaluating translation quality. These criteria enable researchers and professional translators to identify communicative deficiencies that frequently remain undetected by automatic evaluation metrics.

Furthermore, the study emphasizes the indispensable role of human evaluation in artificial intelligence-based translation. Professional translators and post-editors possess linguistic, pragmatic, and intercultural competencies that enable them to recognize contextual meanings, adapt culture-specific expressions, and preserve the communicative purpose of the source text. Consequently, linguopragmatic assessment should be regarded as an essential component of post-editing and translation quality assurance.

In conclusion, the integration of linguopragmatic principles into translation quality assessment contributes significantly to improving the reliability and communicative adequacy of artificial intelligence-generated

translations. Future research should focus on developing standardized linguopragmatic evaluation models and practical assessment criteria applicable to English–Uzbek translation as well as other language pairs. Such studies will contribute to both the advancement of translation studies and the further development of artificial intelligence technologies in multilingual communication.

REFERENCES

1. Austin, J. L. (1962). *How to Do Things with Words*. Oxford University Press.

2. Baker, M. (2018). *In Other Words: A Coursebook on Translation* (3rd ed.). Routledge.

3. House, J. (2015). *Translation Quality Assessment: Past and Present*. Routledge.

4. Levinson, S. C. (1983). *Pragmatics*. Cambridge University Press.

5. Mahmudova, U. (2024). *Raqamli texnologiyalar diskursi tarjimasida sun'iy intellekt ChatGPT va Google Translate tarjimon dasturlarining farqi*. *Uzbekistan: Language and Culture*, 4(4).

6. Bekmurodova, F., Djumabayeva, J., & Madiyorova, V. (2024). *Pragmatic Equivalence in the Translation of Cultural References from Uzbek into English*. *Arab World English Journal for Translation & Literary Studies*, 8(2), 241–254.

7. Kenny, D. (2022). *Machine Translation for Everyone: Empowering Users in the Age of Artificial Intelligence*. Language Science Press.

8. International Organization for Standardization. (2017). *ISO 18587:2017. Translation Services—Post-editing of Machine Translation Output—Requirements*.