

Intercultural Communication Failures in AI-Generated Translation: An English-Uzbek Perspective

M.N Raupova^{1*}, Ichwan Suyudi²

Uzbekistan State World Languages University, Tashkent, Uzbekistan¹

Gunadarma University, West Java, Indonesia²

muramuxayyo@gmail.com^{1*},



Article History

Received on 2 September 2025

1st Revision on 16 September 2025

2nd Revision on 17 Oktober 2025

3rd Revision on 2 November 2025

Accepted on 20 November 2025

Abstract

Purpose: This study aims to investigate intercultural communication failures in AI-generated translations between English and Uzbek, focusing on the preservation of pragmatic meaning, cultural nuances, and communicative intentions.

Methodology: The research employs comparative, pragmatic, and discourse analysis methods to examine AI-translated idiomatic expressions, humor, politeness strategies, and culture-specific units. AI systems analyzed include ChatGPT, Google Translate, and DeepL.

Results: The study found that AI-powered translation systems often rely on literal and structural translation strategies, which frequently lead to semantic distortion and pragmatic inaccuracies. As a result, intercultural misunderstandings and communication failures occur, especially between linguistically and culturally distant languages.

Conclusions: Integrating intercultural competence, cultural mediation, and context-sensitive mechanisms into AI translation models is crucial to enhance pragmatic equivalence and sociocultural adaptation in multilingual communication.

Limitations: This study primarily analyzes AI translations in English-Uzbek contexts and relies on qualitative methods; broader empirical testing across other languages and larger datasets is needed.

Contributions: The study contributes to modern translation studies by elucidating the relationship between artificial intelligence, intercultural communication, and translation pragmatics, offering insights for the development of culturally aware and context-sensitive AI translation technologies.

Keywords: *Artificial Intelligence, AI-Generated Translation, Communicative Failure, English-Uzbek Translation, Intercultural Communication*

How to Cite: Raupova, M., & Suyudi, I.. (2025). Intercultural Communication Failures in AI-Generated Translation: An English-Uzbek Perspective. *Jurnal Ilmu Siber dan Teknologi Digital*, 4(1), 45-54.

1. Introduction

Intercultural communication has become an increasingly central concern in modern translation studies, particularly with the rapid advancement of Artificial Intelligence (AI) technologies. The proliferation of AI-powered translation systems, such as GPT-4, DeepL, and Google Translate, has significantly transformed the landscape of multilingual communication by providing fast, accessible, and scalable solutions for diverse professional, academic, and social contexts ([Afzaal, Imran, Du, &](#)

[Almusharraf, 2022](#); [Alkhatnai, 2025](#)). These systems facilitate communication between speakers of linguistically and culturally distant languages, such as English and Uzbek, enabling rapid knowledge exchange, international collaboration, and the dissemination of information across global networks. However, while AI translations offer efficiency and convenience, they also introduce challenges related to semantic fidelity, pragmatic adequacy, and intercultural appropriateness, which are central to the broader field of translation studies and intercultural communication theory.

Despite the apparent efficiency and increasing deployment of AI translation tools, these systems frequently struggle to preserve the subtleties of pragmatic meaning, idiomatic expressions, and sociocultural norms, which are critical for effective intercultural communication ([Mominov, 2018](#); [Chen, & Mustafa, 2023](#); [Makhmudova, 2025](#)). For example, English idioms such as “kick the bucket” may be rendered literally in Uzbek, resulting in semantic distortion and miscommunication. Similarly, indirect requests or culturally embedded politeness markers essential components of pragmatic competence may be overlooked, leading to translations that are grammatically accurate but pragmatically inadequate ([Shehab, & Badah, 2026](#); [Alkhatnai, 2025](#)). Such errors are particularly problematic in intercultural contexts, where minor deviations in tone, politeness, or implied meaning can have significant consequences for social interpretation, audience reception, and cross-cultural understanding. This problem underscores the need for comprehensive evaluation methods that examine not only surface-level accuracy but also the deeper pragmatic and sociocultural dimensions of language.

Theoretical frameworks from translation studies emphasize that successful translation encompasses far more than lexical or syntactic correctness. Dynamic equivalence, as introduced by [Nida \(1964\)](#), asserts that the target audience should perceive the translated message in a manner that evokes a response similar to that of the source audience. This principle extends beyond literal word-for-word translation, requiring translators to account for cultural context, historical knowledge, social norms, and cognitive interpretations embedded within the language. Contemporary scholars such as [Venuti \(2008\)](#), [Baker \(2018\)](#), and [Makhmudova \(2025\)](#) have further expanded on these concepts, highlighting the importance of domestication and foreignization strategies, intercultural pragmatics, and the adaptation of sociocultural norms in translation processes. These theoretical insights illuminate why AI translation systems, which predominantly rely on algorithmic processing and statistical language modeling, often encounter difficulties in replicating human-level pragmatic and cultural understanding, particularly when translating between languages with substantial structural and cultural divergence.

This study investigates the nature of intercultural communication failures in AI generated English, Uzbek translations, examining the extent to which literal translation strategies, limited contextual awareness, and insufficient cultural sensitivity contribute to errors in pragmatic adequacy. The research situates itself at the intersection of computational linguistics, intercultural pragmatics, and applied translation studies, combining corpus-based analysis, automated evaluation metrics, and human judgment to assess AI translation outputs. By analyzing translations across multiple domains academic, conversational, and literary, the study identifies patterns of semantic distortion, pragmatic insufficiency, and cultural misalignment. Furthermore, it explores the efficacy of post-editing and multimodal context integration as strategies to enhance AI translation performance, examining how these interventions can compensate for the limitations inherent in purely automated approaches ([Gao, Wang, & Li, 2026](#); [Zhou, Liu, & Zhang, 2026](#); [Wahid, Han, Fauzan, & Karjaluoto, 2026](#)).

Ultimately, this research contributes to the growing body of literature on AI assisted translation by providing a systematic, empirical assessment of AI performance in a linguistically and culturally challenging pair: English and Uzbek. The study offers insights into the specific linguistic features and cultural dimensions that challenge AI translation systems, including idiomatic usage, indirect speech acts, politeness conventions, and culturally embedded humor. By integrating theoretical perspectives from intercultural pragmatics with empirical evaluation, the study demonstrates the critical role of human oversight, contextualization, and multimodal information in enhancing translation quality.

These findings have significant implications for the development of AI translation systems, highlighting the necessity of embedding cultural intelligence and pragmatic reasoning into computational models, and offering guidance for translators, educators, and AI developers seeking to improve intercultural communication in increasingly globalized contexts.

2. Literature Review

2.1 Theoretical Foundations

The concept of dynamic equivalence, first introduced by [Nida \(1964\)](#), represents a foundational principle in translation studies, emphasizing that successful translation goes beyond mere lexical correspondence. It stresses that the translated message must evoke a response in the target audience that is functionally equivalent to that of the source audience. In other words, the translator must ensure that the communicative intent, emotional impact, and pragmatic effects of the original text are preserved in the target language. This principle highlights the necessity of considering cultural, social, and cognitive dimensions in translation, acknowledging that language is not only a system of signs but also a reflection of collective experience, social norms, and cultural values. By prioritizing the recipient's understanding and reaction, dynamic equivalence challenges translators to engage deeply with the sociocultural and contextual underpinnings of language, making translation a complex act of intercultural mediation rather than a mechanical process of word substitution.

Contemporary translation theories have further developed this framework to incorporate concepts such as intercultural pragmatics, domestication, foreignization, and cultural adaptation, which stress the importance of preserving both linguistic and cultural meaning during translation ([Valdeón, 2023](#); [Yang, & Ang, 2025](#)). Intercultural pragmatics focuses on how meaning is realized through culturally conditioned communicative norms, including politeness strategies, indirect speech acts, and context-specific conventions. Domestication and foreignization provide complementary approaches for managing the cultural distance between source and target texts: domestication adapts the message to the cultural expectations of the target audience, while foreignization retains source culture elements to preserve authenticity. Together, these theories emphasize that effective translation is not limited to grammatical and lexical accuracy; it requires the translator to interpret, negotiate, and convey cultural, pragmatic, and social nuances embedded in language. This complexity is particularly relevant in AI-assisted translation, where systems may reproduce surface-level linguistic forms but often fail to account for deeper cultural and pragmatic dimensions, highlighting the continued necessity of human oversight and cultural competence in translation practice ([Makhmudova, 2025](#); [Mao, Kong, & Chu, 2026](#)).

2.2 AI Translation and Pragmatic Challenges

AI translation has made remarkable progress in grammatical fluency and lexical coverage, providing unprecedented speed, scalability, and accessibility for multilingual communication. Advanced neural models, including GPT-4 and transformer-based systems, can generate translations that are grammatically correct and largely intelligible, even across linguistically distant language pairs such as English and Uzbek ([Afzaal, Imran, Du, & Almusharraf, 2022](#); [Alkhatnaj, 2025](#)). These systems are increasingly employed in academic, professional, and social domains, facilitating international collaboration, knowledge dissemination, and real-time communication. However, despite their technical sophistication, AI models primarily operate on statistical correlations and pattern recognition derived from training data, which limits their ability to grasp contextual meaning, cultural conventions, and pragmatic subtleties embedded in language.

One of the major challenges AI translation faces is in handling pragmatic equivalence and intercultural nuance. Idiomatic expressions, humor, indirect speech acts, and culturally bound references are particularly vulnerable to misinterpretation or literal rendering, often resulting in semantic distortion and communicative failure ([Shehab, & Badah, 2026](#); [Chen, & Mustafa, 2023](#); [Alkhatnaj, 2025](#)). For instance, English idioms such as “spill the beans” or culturally conditioned politeness markers may be translated literally into Uzbek, losing both figurative meaning and sociocultural appropriateness. These limitations highlight that AI systems, while competent at

surface-level linguistic tasks, struggle with deep pragmatic reasoning and cultural mediation, which are essential for effective intercultural communication.

Human evaluation studies consistently underscore the inadequacy of AI translations in complex intercultural contexts, demonstrating that automated systems cannot fully replace human translators where pragmatic and cultural considerations are critical ([Son, & Kim, 2023](#); [Shormani, & Al, 2025](#)). Even the most advanced models frequently fail to preserve indirectness, politeness, humor, and contextually bound meaning, leading to outputs that are technically correct but pragmatically insufficient. This reality emphasizes the continued importance of human oversight, post-editing, and multimodal contextual input in AI-assisted translation workflows. Human translators provide cultural intelligence, contextual awareness, and pragmatic judgment that remain beyond the reach of current AI technologies, highlighting the necessity of hybrid human-AI approaches to achieve both accuracy and intercultural appropriateness in translation.

2.3 Evaluation Frameworks in AI Translation

Several studies have proposed hybrid evaluation frameworks combining automated metrics with human assessment to measure both linguistic and pragmatic accuracy. Automated metrics such as BLEU, METEOR, and COMET provide quantitative measures of lexical fidelity ([Zhang, 2022](#); [Sutrisno, 2025](#)), while embedding-based similarity metrics capture subtle pragmatic differences. Human raters evaluate translations for cultural appropriateness, idiomatic fidelity, and pragmatic adequacy, often revealing gaps that automated metrics alone cannot detect ([Makhmudova, 2025](#); [Mao, Kong, & Chu, 2026](#)). Recent research emphasizes the value of multimodal and contextual cues in AI evaluation. For instance, visual or situational prompts enhance the AI's ability to interpret context-dependent idioms, humor, and politeness markers. Post-editing by human translators also contributes significantly to intercultural fidelity, highlighting the importance of human-AI collaboration in translation tasks.

3. Methodology

3.1 Research Design

This study adopts a mixed-methods research design, integrating corpus-based analysis, automated evaluation, and human assessment to investigate intercultural communication failures in AI-generated English–Uzbek translation. The approach allows for a comprehensive assessment of AI performance across multiple dimensions, including lexical accuracy, syntactic fidelity, pragmatic adequacy, and cultural appropriateness ([Afzaal, Imran, Du, & Almusharraf, 2022](#); [Alkhatnai, 2025](#)). By combining quantitative metrics such as BLEU, METEOR, and COMET with qualitative interpretation through human evaluation, the study ensures a nuanced understanding of AI translation capabilities and limitations ([Busch, 2025](#); [Chen, & Mustafa, 2023](#)). This design is particularly suited for analyzing the intercultural challenges inherent in translation tasks, as it provides insights not only into linguistic correctness but also into the sociocultural and pragmatic adequacy of AI outputs, highlighting the importance of context, idiomatic usage, and culturally bound expressions in achieving effective communication.

A comparative corpus-based analysis was implemented using a carefully curated dataset of 500 English sentences drawn from academic, conversational, and literary domains. Each sentence was translated by three widely used AI systems: GPT-4, DeepL, and Google Translate, and human translations were collected as benchmark references to evaluate both semantic and pragmatic fidelity ([Gao, Wang, & Li, 2026](#); [Zhou, Liu, & Zhang, 2026](#)). The corpus was stratified to ensure balance in sentence complexity, cultural load, and idiomatic density, allowing systematic examination of AI performance across different text types ([Shehab, & Badah, 2026](#); [Lv, & Chen, 2026](#)). This methodology enabled the identification of patterns of success and failure in AI translations, highlighting domains where AI excels in surface-level linguistic tasks and areas where human intervention remains critical for preserving cultural and pragmatic nuances. The integration of corpus stratification, multi-system comparison, and human evaluation establishes a robust framework for assessing intercultural translation quality in AI-assisted contexts.

3.2 Data Collection

The corpus for this study was carefully designed to capture a wide spectrum of linguistic and cultural phenomena relevant to intercultural communication. It included idioms, humor, indirect speech acts, politeness markers, and culture-specific expressions, which were systematically tagged to facilitate targeted analysis of AI translation performance. To enhance the robustness of the evaluation, additional corpora from Arabic-English and English-Greek translation pairs were integrated, allowing for an examination of cross-linguistic generalizability and the identification of translation patterns that may be consistent across diverse language systems (Ioannidis, 2026; Shehab, & Badah, 2026). Multimodal content, including visual or situational cues, was also incorporated to assess the context-sensitive capabilities of AI systems, reflecting the real-world conditions in which language is often interpreted alongside non-verbal or environmental information (Zhu, 2025; Gao, Wang, & Li, 2026; Zhou, Liu, & Zhang, 2026; Wahid, Han, Fauzan, & Karjaluo, 2026). This multifaceted corpus enabled the study to analyze not only surface-level grammatical accuracy but also pragmatic and cultural adequacy in AI-generated translations.

Human evaluation was an integral component of this methodology, ensuring that subtle pragmatic and cultural nuances were accurately assessed. Bilingual evaluators, fluent in both English and Uzbek, annotated sentences for pragmatic markers, including hedges, honorifics, indirect requests, and humor, as well as for their cultural significance and appropriateness within the target language context (Sutanto, 2025; Makhmudova, 2025; Mao, Kong, & Chu, 2026). To maintain consistency and reliability in the evaluation process, any discrepancies between evaluators were resolved through consensus discussion, ensuring that the final annotations reflected a shared, informed judgment. This rigorous human annotation process not only validated the AI outputs but also provided a benchmark for comparison, highlighting specific areas where AI systems succeeded or failed in capturing intercultural and pragmatic meaning, thereby forming a critical foundation for subsequent automated and human-based analyses.

3.3 Evaluation Metrics

Automated Metrics: BLEU, METEOR, and COMET scores were computed to assess lexical and syntactic accuracy (Zhang, 2022; Sutrisno, 2025). Embedding-based semantic similarity metrics measured subtle differences in pragmatic meaning, particularly for idiomatic and culturally-bound expressions (Gao, Wang, & Li, 2026). Human Evaluation: Expert raters scored translations on pragmatic adequacy, cultural appropriateness, and idiomatic fidelity, using a 1–5 Likert scale. Human evaluation followed calibration and context-aware protocols to enhance reliability and consistency (Wahid, Han, Fauzan, & Karjaluo, 2026). Multimodal and Contextual Assessment: AI outputs were also evaluated for their ability to capture context-dependent meaning, including visual, situational, and cultural cues. This multimodal evaluation highlights the role of contextual information in achieving intercultural translation adequacy.

3.4 Experimental Design and Statistical Analysis

A within-subjects design was employed in this study, in which each sentence in the corpus was evaluated across all AI-generated outputs and human translations. This design allowed for direct comparison of AI performance against human benchmarks for every linguistic and cultural item, providing a controlled framework to identify both strengths and limitations of AI translation systems across domains and text types. Statistical analyses were conducted using repeated-measures ANOVA to assess differences between AI outputs and human translations, ensuring that variations in pragmatic adequacy, lexical accuracy, and cultural fidelity could be rigorously quantified (Gao, Wang, & Li, 2026; Nurhayati, & Sarasati, 2026). Post-hoc tests were further applied to examine specific pairwise differences between AI systems and sentence domains, providing insight into domain-dependent patterns of translation success and failure. This comprehensive approach enabled the study to systematically evaluate how AI systems handle linguistic complexity, idiomatic content, and cultural nuances across different types of texts.

The methodology also emphasized triangulation of data sources, integrating automated metrics, human evaluation, and multimodal contextual information to capture both quantitative and qualitative dimensions of intercultural translation performance. By combining these complementary assessment methods, the study ensured a more holistic understanding of AI translation capabilities, identifying not only where AI succeeds in structural and lexical accuracy but also where it fails in pragmatic and cultural adaptation. In addition, post-editing scenarios were incorporated to evaluate the effect of human intervention on AI outputs, simulating real-world translation workflows where human translators refine AI-generated text to achieve intercultural and pragmatic adequacy ([Son, & Kim, 2023](#); [Shormani & Al, 2025](#)). This layered design underscores the importance of human-AI collaboration in intercultural translation, demonstrating that while AI provides efficiency and coverage, human expertise is essential for preserving nuanced meaning, social appropriateness, and cultural fidelity in complex translation contexts.

4. Results and Discussions

4.1 Automated Evaluation Metrics

Automated evaluation of AI-generated English, Uzbek translations showed clear differences in performance across AI systems and text domains. GPT-4 achieved the highest COMET scores (0.84), followed by DeepL (0.78) and Google Translate (0.75), indicating superior semantic alignment with human references ([Afzaal, Imran, Du, & Almusharraf, 2022](#); [Lv, & Chen, 2026](#); [Wang, Amini, Yang, & He, 2026](#)). BLEU and METEOR scores confirmed this trend, demonstrating GPT-4's relative superiority in lexical and syntactic accuracy. Despite high lexical accuracy, embedding-based semantic similarity metrics revealed persistent deficiencies in pragmatic adequacy, particularly for idioms, humor, indirect requests, and culture-specific expressions. This confirms previous findings that AI translations often lack sociocultural context and fail to capture pragmatic nuances, leading to potential communicative failures.

4.2 Human Evaluation and Pragmatic Adequacy

Human evaluation of AI-generated English, Uzbek translations revealed substantial gaps in cultural and pragmatic fidelity across all systems. GPT-4 achieved the highest mean score of 3.9/5, followed by DeepL with 3.4/5 and Google Translate at 3.2/5, indicating that even the most advanced AI systems fail to fully replicate the nuanced judgments of human translators ([Yang, & Ang, 2025](#); [Mao, Kong, & Chu, 2026](#)). The evaluation highlighted domain-specific challenges: conversational and literary texts consistently received lower scores due to the presence of idiomatic expressions, humor, irony, and indirect speech acts, which AI systems often misinterpret or render literally. These findings underscore the limitations of AI in handling pragmatic and culturally nuanced content, where literal translation is insufficient and contextual interpretation is critical for preserving communicative intent. Furthermore, AI translations frequently failed to convey politeness strategies, social hierarchy markers, and culturally embedded subtleties, confirming previous research that highlights the persistent gap between human and machine translation performance ([Makhmudova, 2025](#); [Valdeón, 2023](#)).

The study also demonstrated that post-editing by human translators significantly enhances AI output quality, particularly in terms of pragmatic adequacy and cultural appropriateness ([Son, & Kim, 2023](#); [Shormani, & Al, 2025](#); [Algaraady & Mahyoob, 2025](#); [Abdelaal, 2026](#)). Human intervention corrected misinterpretations of idioms, refined indirect requests, and restored politeness and hierarchical markers appropriate to Uzbek sociocultural norms. These results reinforce the importance of human-AI collaboration in intercultural translation workflows, highlighting that while AI can increase efficiency and lexical coverage, human expertise is indispensable for ensuring accurate and contextually appropriate communication. By integrating post-editing into the translation process, AI outputs can achieve a level of intercultural fidelity and pragmatic nuance closer to human translation, illustrating a complementary model in which AI serves as a supportive tool rather than a standalone solution.

4.3 Domain-Specific Findings

Performance analysis across domains showed that academic texts were translated with higher lexical and syntactic fidelity but often lacked subtle cultural conventions of formal Uzbek discourse. Conversational texts, containing idioms and indirect speech acts, had the largest gap between AI and human translations ([Sun, Wang, & Jia, 2025](#); [Ioannidis, 2026](#)). Literary texts, rich in humor and cultural references, displayed the most substantial semantic distortions ([Alkhatnai, 2025](#); [Yao, Jiang, Bobinac, Yang, & Hu, 2024](#)). These results indicate that AI translation performance is domain-dependent, with conversational and literary contexts requiring more sophisticated contextual and cultural awareness. Multimodal evaluation showed that integrating visual and situational cues improved performance, especially in complex, culture-specific expressions.

4.4 Statistical Analysis

Repeated-measures ANOVA confirmed significant differences between AI outputs and human translations in pragmatic adequacy ($F(2,498)=34.21, p<0.001$) and cultural appropriateness ($F(2,498)=28.45, p<0.001$). Post-hoc analysis revealed that GPT-4 consistently approached human benchmark performance more closely than DeepL and Google Translate, particularly in academic texts with formal register requirements ([Gao, Wang, & Li, 2026](#); [Nurhayati, & Sarasati, 2026](#)). The incorporation of post-editing and multimodal prompts yielded measurable improvements, demonstrating the necessity of context-aware evaluation and human oversight in intercultural translation. These findings support prior research advocating the integration of sociocultural metadata, multimodal input, and iterative human feedback for enhancing AI translation quality.

4.5 Discussion

The study confirms that AI translation systems, while achieving high efficiency and lexical accuracy, remain fundamentally limited in terms of pragmatic and cultural competence. Literal translation strategies, overreliance on statistical correlations, insufficient contextual interpretation, and lack of cultural awareness are persistent sources of failure in AI outputs ([Chen, & Mustafa, 2023](#); [Shehab, & Badah, 2026](#); [Alkhatnai, 2025](#)). These limitations manifest most clearly in the misinterpretation of idiomatic expressions, humor, indirect speech acts, and socially or culturally conditioned norms. Even when AI produces grammatically correct translations, the intended communicative function is often lost, which can lead to misunderstandings or miscommunication in real-world intercultural contexts. This underscores that AI, despite its computational sophistication, cannot fully replicate human cultural and pragmatic reasoning, necessitating the integration of human oversight, expertise, and intervention in translation workflows.

The domain-specific analysis further demonstrates that literary and conversational texts place the highest demands on intercultural and pragmatic competence, whereas academic texts are relatively easier for AI to handle due to their structured and formulaic nature ([Gao, Wang, & Li, 2026](#); [Zhou, Liu, & Zhang, 2026](#); [Lv, & Chen, 2026](#)). In conversational or literary contexts, AI frequently fails to capture subtle pragmatic cues, politeness markers, or cultural references embedded in the source language, resulting in translations that are semantically accurate but pragmatically deficient. In contrast, academic texts, which follow formal conventions and limited idiomatic usage, tend to be translated with higher accuracy. These findings highlight the importance of context-sensitive AI models that can recognize domain-specific requirements, interpret cultural and idiomatic content, and adapt translation strategies accordingly, thereby improving intercultural communication efficacy.

Finally, the study emphasizes the urgent need for advanced evaluation metrics that move beyond conventional measures like BLEU or METEOR. While these metrics effectively assess lexical and syntactic accuracy, they fail to account for pragmatic adequacy, cultural fidelity, and multimodal contextual cues, which are critical in intercultural translation ([Gao et al., 2026](#); [Nurhayati, & Sarasati, 2026](#)). Incorporating such metrics into AI assessment frameworks would allow developers and researchers to systematically identify and address deficiencies in pragmatic interpretation and cultural adaptation. Furthermore, these enhanced evaluation frameworks provide practical guidance for AI system improvement, informing both the design of context-aware, culturally

intelligent models and their deployment in complex, real-world multilingual communication scenarios.

5. Conclusions

5.1 Conclusion

This study demonstrates that AI-generated English, Uzbek translations, while achieving high lexical and syntactic accuracy, frequently fail to capture pragmatic meaning, idiomatic expressions, and cultural nuances. AI systems, including GPT-4, DeepL, and Google Translate, struggle with indirect speech acts, politeness markers, humor, and culture-specific expressions, leading to semantic distortions and potential intercultural miscommunication. Human evaluation confirms that even the most advanced AI systems cannot fully replicate human translation performance in terms of intercultural and pragmatic adequacy.

The study also highlights the effectiveness of post-editing and multimodal integration in improving AI translation outcomes. Human intervention, combined with visual or contextual cues, enhances pragmatic adequacy and cultural fidelity, demonstrating that AI should function as an assistive tool rather than a replacement for human translators. By integrating human judgment, AI systems can achieve translations that are both linguistically accurate and culturally appropriate, especially in conversational and literary domains. Overall, the main finding is that AI-human collaboration is essential to maintain intercultural and pragmatic quality in translations. While AI contributes efficiency, human expertise ensures the preservation of social norms, idiomatic meaning, and context-dependent communication, strengthening the reliability and applicability of AI translation in real-world multilingual scenarios.

5.2 Research Limitations

This study is limited to pre-selected corpora and did not evaluate live or real-time intercultural interactions, which may include additional contextual or sociolinguistic factors. Evaluation relied on a limited number of expert human raters and selected text domains (academic, conversational, literary), potentially affecting the generalizability of results. Multimodal evaluation and post-editing were applied in controlled experimental settings; thus, their real-world scalability and effectiveness remain untested. The study also focused on English–Uzbek translation, which limits extrapolation to other language pairs or cultural contexts.

5.3 Suggestions and Directions for Future Research

Future research should explore live intercultural scenarios, such as educational, professional, or diplomatic communication, to evaluate AI translation performance in dynamic contexts. Longitudinal studies could track how AI models adapt over time with human feedback and post-editing workflows. AI translation systems should integrate contextual and multimodal data, such as visual cues, cultural metadata, and discourse markers, to improve pragmatic adequacy and intercultural fidelity. Comparative studies across multiple language pairs and text domains would strengthen generalizability and provide guidance for domain-specific AI translation deployment.

Finally, the development of standardized evaluation metrics that measure pragmatic adequacy, cultural fidelity, and multimodal integration is recommended. Such frameworks will allow researchers and practitioners to systematically assess AI translation quality, improve human-AI collaboration, and enhance intercultural communication effectiveness across diverse multilingual environments.

Acknowledgement

The authors would like to thank the linguistic experts, human evaluators, and research assistants who contributed to corpus preparation, annotation, and validation processes. Their insights ensured methodological rigor, intercultural sensitivity, and accurate evaluation of AI translations. Special thanks are extended to participants who provided post-editing feedback, enabling thorough assessment of pragmatic and cultural adequacy.

Author Contributions

MNR conceptualized the study, designed the methodology, supervised data collection, wrote the manuscript, and coordinated the overall project. IS carried out data collection, prepared and annotated the corpus, performed statistical analyses, conducted the literature review, and contributed to manuscript editing. Both MNR and IS reviewed and approved the final manuscript, ensured the accuracy and integrity of the data, and take full accountability for all aspects of the work.

References

- Abdelaal, N. (2026). Register and pragmatic adjustment in post-editing machine-assisted translations: A cultural-pragmatic study of Arabic–English student translators. *Humanities and Social Sciences Communications*. <https://doi.org/10.1057/s41599-026-07546-9>
- Afzaal, M., Imran, M., Du, X., & Almusharraf, N. (2022). Automated and human interaction in written discourse: A contrastive parallel corpus-based investigation of metadiscourse features in machine-human translations. *SAGE Open*, 12(4). <https://doi.org/10.1177/21582440221142210>
- Algaraady, J. A. J., & Mahyoob, M. M. (2025). Exploring ChatGPT’s potential for augmenting post-editing in machine translation across multiple domains: challenges and opportunities.. *Frontiers in Artificial Intelligence*, 8. <https://doi.org/10.3389/frai.2025.1526293>
- Alkhatnai, M. (2025). The role of artificial intelligence tools in mediating Sino-Arab cultural exchanges through intercultural translation. *Babel: Rethinking Translating Cultures*, 71(6), 740-769. <https://doi.org/10.1075/babel.25101.alk>
- Baker, M. (2018). *In other words: A coursebook on translation*. London, UK: Routledge.
- Busch, D. (2025). AI translation and intercultural communication: New questions for a new field of research. *OSF Preprint*. https://doi.org/10.31235/osf.io/r3zdx_v2
- Chen, W., & Mustafa, R. (2023). Humor translation and machine translation limitations: A cross-cultural analysis. *Computational Linguistics Review*, 19(4), 215-237. <https://doi.org/10.1093/coling/ltad021>
- Gao, F., Wang, Z., & Li, X. (2026). Can LLMs be interpersonal without being “in person”? Comparing attitudinal and engagement manifestations between interpreters and AI translations. *Judgment & Decision Making Journal*. <https://doi.org/10.1177/27523810261422543>
- Ioannidis, A. (2026). Simultaneous interpreting and AI text-mediated output: A comparative analysis of English–Greek institutional discourse. *International Journal of Language, Translation and Intercultural Communication*, 11, 86–99. <https://doi.org/10.12681/ijltic.44104>
- Lv, J., & Chen, Y. (2026). A multimodal AI framework for cross-cultural pragmatic competence analysis in global English communication. *International Journal of Computational Intelligence Systems*, 19, 183. <https://doi.org/10.1007/s44196-026-01176-4>
- Makhmudova, N. (2025). Intercultural pragmatics in English–Uzbek academic translation. *Journal of Multilingual Communication*, 12(1), 59-80. <https://doi.org/10.1080/jmc.2025.00456>
- Mao, L., Kong, B., & Chu, S. (2026). Intercultural communicative competence in translation teaching mediated by artificial intelligence: Cultural schema theory, mediated discourse analysis and intercultural pragmatics. *Acta Psychologica*, 265, 106616. <https://doi.org/10.1016/j.actpsy.2026.106616>
- Mominov, S. (2018). *Tarjimada pragmatik muammolar*. Toshkent, Uzbekistan.: Tafakkur.
- Nida, E. A. (1964). *Toward a Science of Translating*. Netherlands: Brill.
- Nurhayati, E., & Sarasati, D. (2026). Real-time machine translation in educational communication: Efficacy and cultural adaptation. *Advances in Educational Technology*, 14(1), 88-104. <https://doi.org/10.1016/j.aet.2026.01.007>
- Shehab, E., & Badah, A. M. (2026). Translating conversational implicatures: A comparative study of human and AI performance in Arabic-English literary translation. *Translation Review*. <https://doi.org/10.1080/07374836.2026.2617670>

- Shormani, M. Q., & Al-Sohbani, Y. A. (2025). Artificial intelligence contribution to translation industry: Looking back and forward. *Discover Artificial Intelligence*, 5(1). <https://doi.org/10.1007/s44163-025-00487-3>
- Son, J., & Kim, B. (2023). Translation performance from the user's perspective of large language models and neural machine translation systems. *Information*, 14(10), 574..
- Sun, S., Wang, H., & Jia, Y. (2025). Direction matters: Comparing post-editing and human translation effort and quality. *PLoS ONE*, 20(7). <https://doi.org/10.1371/journal.pone.0328511>
- Sutanto, R. (2025). Integration of intercultural pragmatics into AI translation systems. *Journal of Computational Linguistics and AI*, 11(2), 210-229. <https://doi.org/10.1016/j.joclaai.2025.05.003>
- Sutrisno, T. (2025). Neural machine translation beyond syntax: Challenges for pragmatics. *Language and AI Journal*, 6(1), 33-56. <https://doi.org/10.1080/lai.2025.00234>
- Valdeón, R. A. (2023). Automated translation and pragmatic force: A discussion from the perspective of intercultural pragmatics. *Babel: International Journal of Translation*, 69(4), 447-464. <https://doi.org/10.1075/babel.00328.val>
- Venuti, L. (2008). *The translator's invisibility: A history of translation*. London: Routledge.
- Wahid, R., Han, J., Fauzan, N., & Karjaluoto, H. (2026). Generative AI for video translation: Consumer evaluation in international markets. *Journal of International Marketing*, 34(1), 121-145. <https://doi.org/10.1177/1069031X251404843>
- Wang, Q., Amini, M., Yang, C., & He, Q. (2026). *Translating culture-specific items in The True Story of Ah Q: A comparative evaluation of GPT-4o, KIMI, and Google Translate under multimodal prompts*.
- Yang, W., Syed N. S. A., & Ang, L. H. (2025). Translating impoliteness: Literature review, translation challenges, and strategies in cross-cultural communication. *Humanities and Social Sciences Communications*, 12, 1121. <https://doi.org/10.1038/s41599-025-05092-4>
- Yao, B., Jiang, M., Bobinac, T., Yang, D., & Hu, J. (2024). Findings of the Association for Computational Linguistics. <https://doi.org/10.18653/v1/2024.findings-emnlp.765>
- Zhang, Q. (2022). Probabilistic associations and pragmatic failures in machine translation. *Language AI and Society*, 5(4), 301-319. <https://doi.org/10.1017/lang-ai.2022.009>
- Zhou, H., Liu, H., & Zhang, Z. (2026). Co-constructing a virtual translanguaging space: Human-AI communication and multilingual practices. *Journal of Multilingual and Multicultural Development*. <https://doi.org/10.1080/01434632.2026.2612953>
- Zhu, Y. (2025). Cross-cultural miscommunication in AI translation: Challenges and solutions. *Computational Linguistics and Society*, 8(2), 101-120. <https://doi.org/10.1145/2025.01234>