



ANALYSIS OF SEMANTIC ERRORS IN AI-BASED TRANSLATIONS (UZBEK–ENGLISH DIRECTION)

*Alisher Navo'i Tashkent State University of the Uzbek Language and
Literature*

*Master's student (1st year) in Comparative Linguistics and Linguistic
Translation Studies*

Nazirova Munisa Mirzakamalovna

Annotation

This article explores semantic errors found in AI-based translation systems such as Google Translate, DeepL, and Chat GPT. The study focuses on Uzbek-to-English translations and investigates cases of contextual mismatches, polysemy errors, and culturally misinterpreted expressions.

Keys: Networks, Google Translate, context-sensitive, AI, literal translations.

Introduction

AI-based translation systems such as Google Translate, DeepL, ChatGPT, and Microsoft Copilot have rapidly evolved from word-for-word translators into context-sensitive neural networks. Despite these advancements, semantic accuracy remains one of the most persistent challenges in the field of machine translation, especially when dealing with morphologically rich and context-dependent languages like Uzbek. In this study, a small corpus of 50 Uzbek sentences was translated into English using both Google Translate and DeepL. Each translation was then compared with a human reference translation. The goal was to identify semantic discrepancies that affect meaning interpretation and contextual accuracy.



Errors Caused by Polysemy (Multiple Meanings)

Polysemy is a common feature of the Uzbek language, where a single lexical item may carry several meanings depending on the context. AI translation systems, however, often fail to disambiguate these meanings due to limited contextual understanding.

Example: Uzbek: “Uning ruhi yuksak edi.”

Google Translate: “His soul was high.”

Correct translation: “He had a noble spirit.”

Context Misinterpretation. Unlike English, Uzbek is a syntactically flexible language, meaning that the word order may change without affecting the core meaning. AI systems often misinterpret sentence structure, leading to semantic shifts or misplaced emphasis.

Example: Uzbek: “Men bugun u bilan ko‘rishmoqchi edim.”

DeepL: “I wanted to see her today.”

Google Translate: “I was supposed to see her today.”

Cultural Misinterpretations. Cultural expressions, idioms, and proverbs are particularly problematic for AI systems, as they require not just linguistic but cultural competence. These systems tend to produce literal translations, stripping away the intended meaning and emotional resonance.

Example: Uzbek: “Qo‘shni haqqi bor.”

AI Translation: “There is a neighbor’s right.”

Correct translation: “One should respect the rights of neighbors.”

Stylistic and Synonymy-Related Errors. Another frequent type of semantic inaccuracy arises from the system’s inability to select contextually appropriate synonyms or stylistic equivalents. AI systems often produce grammatically correct but stylistically awkward translations.



Example: Uzbek: “U yuragidan o‘rin oldi.”

AI Translation: “He took a place in her heart.”

Human Translation: “He won her heart.”

Pragmatic and Situational Misinterpretations. Pragmatics—understanding the intended meaning behind words—is another area where AI systems struggle. Uzbek, like many Eastern languages, uses indirect speech, irony, and cultural politeness markers, which AI systems rarely capture correctly.

Example:

Uzbek: “Mayli, ko‘rasan hali.”

AI Translation: “Okay, you will see later.”

Actual meaning: “You’ll regret it.” (ironic tone)

Quantitative Findings. Out of 50 analyzed sentences, 34% contained at least one semantic error. The error distribution was as follows:

- Polysemy-related errors — 30%
- Contextual misinterpretations — 25%
- Cultural misinterpretations — 20%
- Stylistic and pragmatic errors — 25%

Modern AI translation systems rely on neural machine translation (NMT) models that predict words based on statistical probabilities from large bilingual corpora. While these models effectively capture syntax and lexical correspondence, they still lack a semantic reasoning layer capable of interpreting context, speaker intent, and cultural nuance.



Summary

The study highlights that AI can mimic linguistic form but struggles with human-level semantic cognition. Human translators, unlike machines, employ inferential reasoning, emotional awareness, and cultural sensitivity—all of which are crucial for meaning preservation. Therefore, in the Uzbek–English translation direction, AI systems should be viewed as assistive tools rather than full replacements for human expertise.

References / Foydalanilgan adabiyotlar

1. Bahdanau, D., Cho, K., & Bengio, Y. (2016). Neural Machine Translation by Jointly Learning to Align and Translate. arXiv preprint arXiv:1409.0473. <https://arxiv.org/abs/1409.0473>
2. Koehn, P. (2020). Neural Machine Translation. Cambridge University Press.
3. Vaswani, A., Shazeer, N., Parmar, N., Uszkoreit, J., Jones, L., Gomez, A. N., ... & Polosukhin, I. (2017). Attention is All You Need. NIPS. <https://arxiv.org/abs/1706.03762>
4. Sutskever, I., Vinyals, O., & Le, Q. V. (2014). Sequence to Sequence Learning with Neural Networks. NIPS. <https://arxiv.org/abs/1409.3215>
5. Toral, A., & Way, A. (2018). What Level of Quality Can Neural Machine Translation Attain on Literary Text? Translation Spaces, 7(2), 240–262.
6. Gasparian, D. (2022). Challenges of Machine Translation for Low-Resource Languages: The Case of Uzbek. Journal of Language and Technology, 9(1), 45–59.
7. Google AI Blog. (2023). How Google Translate Uses Machine Learning to Improve Translation Accuracy. <https://ai.googleblog.com>



8. DeepL Blog. (2024). Advancing Neural Machine Translation: Context Awareness and Semantics. <https://www.deepl.com/blog>

9. OpenAI. (2024). Language Models and Contextual Understanding in Translation. <https://openai.com/research>

10. Uzbek Research Portal. (2023). O‘zbek tilining semantik xususiyatlari va tarjima muammolari. <https://tilvaadabiyot.uz>