JOURNAL OF INTERNATIONAL SCIENTIFIC RESEARCH
Volume 4, Issue 6, July- August 2026 Online ISSN: 3030-3508
https://spaceknowladge.com

AN ERROR TYPOLOGY FOR NEURAL MACHINE TRANSLATION IN ENGLISH-TO-UZBEK ACADEMIC DISCOURSE

Ismoilov Ibrokhimjon Bekpulat ugli
Student of Karshi International University
+998 95 066 08 12 ismoilovibrohimjon18@gmail.com
ORCID: https://orcid.org/0009-0000-8523-9050
Scientific Supervisor: **Rakhmonova Mokhinur Zohidovna**
Lecturer of Philology and english language at Karshi International University

MAQOLA MALUMOTI

MAQOLA TARIXI:
Received:01.07.2026
Revised: 02.08.2026
Accepted:03.08.2026

KALIT SO'ZLAR:
Neural Machine Translation, translation errors, academic text, english-uzbek translation, computer linguistics, sentence structure, MQM model

ANNOTATSIYA:

Neural Machine Translation (NMT) systems help to share knowledge between different languages very fast. However, translating academic texts between very different languages, like english and uzbek, is still difficult. english uses Subject-Verb-Object (SVO) word order, while uzbek uses Subject-Object-Verb (SOV) order. This research looks at translation errors in NMT outputs like Google Translate and ChatGPT-4 when they translate english humanities texts into uzbek. We used the Multidimensional Quality Metrics (MQM) system to check 150 academic paragraphs. We checked them for accuracy, fluency, terms, and syntax. The results show that AI tools choose words well. However, they still have problems with long sentences, passive verbs, and academic style in uzbek.

Introduction

Neural Machine Translation (NMT) systems have improved automated translation a lot. They work very well for main world languages. However, translation quality drops when these systems translate difficult academic texts into less common languages. english academic texts usually have long sentences, passive voice, precise words, and a formal tone. On the other hand, uzbek has a strict Subject-Object-Verb (SOV) sentence order and uses suffixes to connect words.

When NMT systems translate english academic texts into uzbek word-for-word, they make grammatical errors. The translation often sounds unnatural or loses its main meaning. Therefore, we need to find these systematic errors. This helps to create better editing steps for researchers, students, and translators. This study checks machine translation errors in english-

to-uzbek academic texts. It creates a simple error list to show where human editing is still necessary.

Methods and Materials

This study used 150 english academic texts from social science and humanities journals.

1. Translation Process: We translated the text samples using two popular tools: Google Translate and ChatGPT-4.
2. Evaluation Framework: We checked and categorized the errors using a simplified Multidimensional Quality Metrics (MQM) model. We focused on four main areas:

Terminology Errors**:** Wrong use of academic terms or using different words for the same idea.

Syntactic Distortion: Incorrect word order, wrong case endings, or errors in long sentences.

Morphological Mismatch: Wrong verb forms, incorrect active-passive conversion, or wrong suffixes on nouns.

Style and Register: Using everyday uzbek expressions instead of formal academic language.

Results and Discussion

The quantitative categorization of translation errors across the evaluated NMT outputs shows a clear concentration of flaws in structural and stylistic dimensions:

Error Category	Frequency Share (%)	Primary Linguistic Cause	Typical NMT Failure Pattern
Syntactic & Word Order	36.4%	Failure in the transition from analytic Subject-Verb-Object (SVO) to agglutinative Subject-Object-Verb (SOV) structures.	Main verb placed prematurely; awkward relative clauses
Morphological / Voice	29.1%	Incorrect passivization and case assignment.	Misuse of *-il-/-ish-* suffixes; broken agent markers
Style & Academic Register	19.5%	Lack of domain-specific stylistic fine-tuning	Using informal conversational

			connectors in Uzbek
Terminology Discrepancy	15.0%	Polysemy in academic vocabulary	Translating abstract terms with everyday literal meanings

Empirical evaluation shows that Syntactic and Word Order Errors (36.4%) are the main problem for Neural Machine Translation (NMT) quality. english academic sentences often have long connected parts. NMT tools often fail to change these sentences into correct uzbek word order. Because of this, the translated sentences sound incomplete and broken.

Morphological Mismatch (29.1%) also happens often when translating english passive sentences. english academic writing uses passive voice a lot to stay objective. However, direct NMT translations into uzbek sound awkward or unclear. On the other hand, lexical terminology errors are low (15.0%). This shows that modern NMT systems have good dictionary memory. Still, these tools cannot always choose the right word for formal academic style.

Conclusion

Neural Machine Translation is a good tool to make first drafts in english-uzbek academic translation. However, current systems cannot make final texts without human editing. The main problems of NMT tools come from bad sentence structure and wrong word forms, not from small vocabulary. To fix these problems, we need to build larger academic text databases for low-resource languages like english-uzbek. Also, we need to train AI models for formal academic grammar. Human translators should focus on fixing sentence structures and making the style sound natural.

References:

1. Vaswani, A., et al. (2017). Attention is all you need. Advances in Neural Information Processing Systems, 30.
2. Lommel, A., Uszkoreit, H., & Burchardt, A. (2014). Multidimensional Quality Metrics (MQM): A framework for declaring and evaluating translation quality. Tradumàtica, (12), 455-463.
3. Koehn, P. (2020). Neural Machine Translation. Cambridge University Press.