

LINGUOPRAGMATIC FEATURES OF PROCESSING ONOMATOPOEIC UNITS IN AI-BASED MACHINE TRANSLATION: A COMPARATIVE STUDY OF ENGLISH AND UZBEK

Kayumova Ozoda Baxrom qizi
ORCID: 0009-0008-2955-2817
bakhramovnaozoda@gmail.com
O'ZDJTU

Annotatsiya: Sun'iy intellekt rivoji zamonaviy mashina tarjimasida texnologiyalariga sezilarli ta'sir ko'rsatdi. Biroq onomatopik birliklar kabi ekspressiv til birliklarini tarjima qilish ularning semantik va pragmatik xususiyatlari sababli murakkab masala bo'lib qolmoqda. Mazkur maqolada ingliz va o'zbek tillaridagi onomatopik birliklarning SI asosidagi mashina tarjimasida qayta ishlanishining lingvopragmatik xususiyatlari tadqiq etiladi. Tadqiqotning maqsadi SI tarjima tizimlarining onomatopik birliklarning semantik, pragmatik va stilistik funksiyalarini saqlash imkoniyatlarini aniqlashdan iborat.

Kalit so'zlar: onomatopiya, sun'iy intellekt, mashina tarjimasida, lingvopragmatika, ingliz tili, o'zbek tili, tarjima ekvivalentligi

LINGUOPRAGMATIC FEATURES OF PROCESSING ONOMATOPOEIC UNITS IN AI-BASED MACHINE TRANSLATION: A COMPARATIVE STUDY OF ENGLISH AND UZBEK

Kayumova is the daughter of Ozoda Bahrom
ORCID: 0009-0008-2955-2817
bakhramovnaozoda@gmail.com
UZSUWL

Abstract: The development of artificial intelligence has significantly influenced modern translation technologies, particularly through the advancement of neural machine translation systems. However, the accurate translation of expressive lexical units remains a challenging issue due to their semantic complexity and pragmatic dependence. This study investigates the linguopragmatic features of processing onomatopoeic units in AI-based machine translation, focusing on English and Uzbek. The research aims to examine how AI translation systems preserve the semantic, pragmatic, and stylistic functions of onomatopoeic expressions. The study employs comparative, descriptive, and linguopragmatic analysis methods based on examples extracted from literary texts, media discourse, and AI-generated translations. The

findings reveal that while AI systems demonstrate relatively high accuracy in conveying basic meanings, they often encounter difficulties in maintaining emotional expressiveness, cultural specificity, and contextual interpretation.

Keywords: onomatopoeia, artificial intelligence, machine translation, linguopragmatics, English, Uzbek, translation equivalence, expressive units

In recent years, artificial intelligence (AI) has become one of the most influential technologies in the development of modern language processing and translation systems. The advancement of neural machine translation and large language models has significantly enhanced the ability of automated systems to analyse context, recognize linguistic patterns, and generate more natural translations. Nevertheless, despite these technological achievements, AI-based translation systems still encounter difficulties when dealing with expressive and context-dependent language units. Onomatopoeic expressions represent a particularly complex category because they are not limited to simple sound imitation; they also convey emotional attitudes, stylistic meanings, and pragmatic intentions. Their interpretation is determined by various factors, including linguistic structure, cultural background, communicative situation, and discourse context. The translation of onomatopoeic units between English and Uzbek is especially challenging due to differences in phonological systems, lexical patterns, and cultural representations of sounds. Although AI systems can often identify and translate the basic meanings of sound-symbolic words, they may fail to maintain their expressive force, emotional impact, and pragmatic functions in the target language.

Despite significant improvements in artificial intelligence-based translation technologies, the accurate processing of culturally and linguistically complex expressions remains a challenging issue. Onomatopoeic units are among the most problematic lexical categories because their meanings are closely associated with sound symbolism, emotional expression, and contextual interpretation. The primary problem addressed in this research is the limited ability of current AI-based translation systems to preserve the full linguopragmatic meaning of English and Uzbek onomatopoeic units. While these systems may achieve lexical and semantic accuracy, they often fail to reproduce the expressive and communicative effects embedded in such units.

Previous investigations of onomatopoeia have mainly concentrated on its phonological structure, semantic features, cognitive aspects, and stylistic functions. Similarly, studies on machine translation have largely focused on general translation quality, accuracy, and computational performance. However, the linguopragmatic processing of onomatopoeic expressions within AI-based translation systems remains insufficiently explored, particularly in the context of English and Uzbek. The current

study aims to fill this gap by examining AI-generated translations through a comparative and pragmatic approach, with particular attention to meaning preservation and communicative effectiveness.

The purpose of this research is to explore the linguopragmatic characteristics of English and Uzbek onomatopoeic units in AI-based machine translation systems and to evaluate how effectively these systems preserve their semantic, pragmatic, and expressive meanings.

To achieve the aim of the study, this research focuses on several interconnected objectives. First, it aims to classify English and Uzbek onomatopoeic units according to their structural features and functional characteristics in different communicative contexts. Secondly, the study seeks to examine the semantic, pragmatic, and communicative functions of these units in order to determine how they contribute to meaning formation and emotional expression. Furthermore, the research analyses the translation strategies employed by AI-based machine translation systems when processing onomatopoeic expressions. Particular attention is given to identifying how artificial intelligence handles sound symbolism, contextual meanings, and expressive features. Another important objective is to compare AI-generated translations with human-produced equivalents in order to evaluate the level of semantic accuracy and pragmatic adequacy. In addition, the study aims to identify the major linguopragmatic challenges that arise during the translation of English and Uzbek onomatopoeic units. Based on the findings, the research provides recommendations for improving the effectiveness of AI-based machine translation systems, particularly in relation to the processing of expressive and culturally specific linguistic units.

Onomatopoeia is considered one of the most distinctive lexical phenomena that demonstrates the relationship between sound and meaning in language. It refers to words or expressions that imitate natural sounds, human actions, animal voices, environmental phenomena, or mechanical noises. Unlike conventional lexical units, onomatopoeic expressions are characterized by a closer connection between their phonetic form and semantic content, which makes them an important object of study in phonetics, lexicology, semantics, and cognitive linguistics. Researchers have approached onomatopoeia from different theoretical perspectives, emphasizing its role in representing sound experiences through linguistic forms. Although onomatopoeic units are often associated with direct sound imitation, their meanings may extend beyond simple reproduction of sounds and acquire additional emotional, expressive, and contextual meanings. For this reason, onomatopoeia cannot be interpreted only as a phonetic phenomenon; it also functions as a meaningful linguistic resource in communication and discourse. Onomatopoeic units can be classified according to their semantic and functional characteristics. One common classification includes imitative expressions representing natural sounds, animal sounds, human sounds, and sounds

produced by objects or mechanical processes. In addition, onomatopoeic units may perform different linguistic functions, such as creating vivid descriptions, intensifying emotions, expressing attitudes, and contributing to stylistic effects in literary and conversational discourse.

The linguopragmatic approach examines language units as elements of communication that acquire meaning through interaction between linguistic forms, speakers' intentions, and situational contexts. Within this framework, meaning is understood not only as a dictionary definition but also as a product of communicative purposes, social relations, and contextual interpretation. Pragmatics plays a central role in analysing onomatopoeic expressions because their meanings often depend on the situation in which they are used. The same onomatopoeic unit may express different emotions, attitudes, or communicative intentions depending on the speaker, audience, and discourse environment. Therefore, a pragmatic analysis allows researchers to identify how such expressions function beyond their literal sound imitation.

The development of artificial intelligence has brought significant changes to the field of machine translation. Modern translation technologies have moved from rule-based systems toward data-driven approaches, particularly Neural Machine Translation (NMT), which uses deep learning algorithms to analyse linguistic patterns and generate translations based on large-scale language data. Neural Machine Translation systems have improved translation quality by considering broader contextual relationships rather than translating individual words separately. Unlike traditional approaches, NMT models are capable of capturing complex linguistic structures and producing more natural translations. However, despite these advantages, they still face difficulties in processing expressions that contain cultural, emotional, or pragmatic information. The emergence of Large Language Models (LLMs) has further expanded the capabilities of AI-based language processing. These models can analyse context, generate coherent texts, and perform multilingual translation tasks with increasing accuracy. Nevertheless, their ability to interpret subtle pragmatic meanings, figurative expressions, and culture-specific linguistic elements remains a subject of ongoing research.

The materials selected for this study consist of various textual sources containing English and Uzbek onomatopoeic units. The research corpus includes examples taken from literary works, children's literature, media texts, and digital discourse, where onomatopoeic expressions frequently perform expressive, descriptive, and communicative functions. Literary texts were selected due to their rich use of sound-symbolic expressions for creating imagery, emotional impact, and stylistic effects. Children's literature was included because onomatopoeic units play an important role in early language development and are widely used to imitate sounds and convey emotions. In addition, media and digital discourse materials were analysed to examine

the contemporary use of onomatopoeic expressions in informal communication, online platforms, and modern forms of written interaction. AI-generated translations of selected English and Uzbek onomatopoeic units were also included as research materials in order to evaluate the performance of machine translation systems and identify possible semantic and pragmatic shifts.

To examine the effectiveness of artificial intelligence in processing onomatopoeic expressions, several AI-based translation systems are selected for analysis. The study focuses on widely used neural machine translation platforms, including ChatGPT, Google Translate, and DeepL. These systems represent different approaches to AI-based language processing and provide an opportunity for comparative evaluation. The selected AI tools are analysed according to their ability to recognize, interpret, and translate English and Uzbek onomatopoeic units. Particular attention is given to how effectively these systems preserve the original semantic meaning, emotional colouring, stylistic features, and pragmatic functions of onomatopoeic expressions. The comparison of AI-generated outputs allows for identifying strengths and limitations of current machine translation technologies in handling expressive linguistic units.

The analysis of the selected linguistic materials demonstrates that onomatopoeic units in English and Uzbek represent a diverse lexical category with different semantic and functional characteristics. Based on the collected examples, these units can be classified into several major groups according to the sources of the sounds they imitate. The main categories include animal sounds, human sounds, natural phenomena, and sounds produced by objects or mechanical processes. Animal-related onomatopoeic expressions constitute one of the most common groups in both languages. These units are frequently used in literary texts, children's literature, and everyday communication to represent animal voices and actions. For example, English expressions such as *meow* and *buzz* correspond to Uzbek forms such as *miyov* and *viz-viz*. Although these units perform similar communicative functions, their phonetic realization differs due to the specific sound systems and cultural conventions of each language. Human sound expressions form another important category. Words such as *cough* and *ha-ha* in English and their Uzbek equivalents represent physical reactions, emotions, or interpersonal communication. These expressions often carry additional pragmatic meanings, including humour, surprise, agreement, or emotional evaluation. Therefore, their interpretation depends not only on their literal meaning but also on the communicative context. Natural and environmental sounds represent another significant group of onomatopoeic units. Expressions such as *splash* and *boom* in English and *shalop* and *gumbur* in Uzbek demonstrate how languages represent sounds produced by natural events or physical actions. The comparison shows that while the motivation behind these expressions is universal, their linguistic forms may vary considerably across languages.

No	Category	English Examples	Uzbek Examples
1	Animal sounds	<i>meow, buzz, woof</i>	<i>miyov, viz-viz, vov</i>
2	Human sounds	<i>cough, ha-ha, sigh</i>	<i>yo'tal, haha, uf</i>
3	Natural sounds	<i>splash, boom, crack</i>	<i>shalop, gumbur, qars</i>
4	Object sounds	<i>click, bang, tick-tock</i>	<i>chik, taraqlash, tik-tak</i>

The findings indicate that onomatopoeic units serve multiple functions beyond simple sound imitation. They contribute to emotional expression, create vivid imagery, strengthen stylistic effects, and influence the interpretation of discourse. This multifunctionality explains why their translation requires consideration of semantic, pragmatic, and cultural factors.

The analysis of AI-generated translations reveals that modern artificial intelligence systems demonstrate considerable ability to recognize and translate the basic meanings of onomatopoeic units. However, the preservation of pragmatic and stylistic features remains a challenging issue. AI systems often provide direct lexical equivalents but may not fully reproduce the emotional intensity, cultural associations, or contextual meanings expressed in the original text. The comparison between AI-generated translations and human translations shows several differences. Human translators usually consider the broader discourse context, the author's intention, and the expected effect on the reader. In contrast, AI systems primarily rely on learned linguistic patterns and statistical associations, which may lead to partial loss of expressive meaning.

The results demonstrate that AI systems are more successful when translating isolated onomatopoeic units with clear meanings. However, when these expressions appear in literary or emotional contexts, translation accuracy decreases due to the difficulty of interpreting implicit meanings and pragmatic functions. The study also shows that AI-based translation tools differ in their approaches to expressive lexical units. Some systems tend to preserve the original form through direct translation, while others attempt contextual adaptation. Nevertheless, none of the analysed systems consistently achieves complete pragmatic equivalence without human correction.

The findings confirm that onomatopoeic expressions represent a complex area for AI-based machine translation because their meanings extend beyond simple lexical correspondence. The translation of such units requires understanding the relationship between sound form, semantic content, speaker intention, and cultural context. One of the main issues identified is semantic equivalence. AI systems generally demonstrate satisfactory performance in transferring the basic meaning of onomatopoeic expressions. However, semantic similarity does not always guarantee full translation adequacy because many onomatopoeic units contain additional expressive meanings that are not explicitly represented in the lexical form. Another important aspect is

pragmatic adequacy. Onomatopoeic units frequently function as emotional and stylistic devices. They may express excitement, fear, humour, surprise, or emphasis depending on the communicative situation. The study shows that AI systems often struggle to preserve these pragmatic functions because they require deeper contextual interpretation. The issue of cultural adaptation is also significant in English-Uzbek translation. Sound representation is influenced by cultural perception and linguistic traditions. A sound that is represented in one language by a particular form may have a different equivalent in another language. Therefore, effective translation requires not only linguistic knowledge but also cultural awareness. The preservation of expressive meaning remains one of the greatest challenges for AI translation systems. Literary texts, children's literature, and digital communication often rely on onomatopoeic expressions to create emotional and visual effects. When these functions are lost during translation, the overall communicative impact of the text may be reduced. The results suggest that various translation strategies, including adaptation, contextual substitution, paraphrasing, and human post-editing, remain important for achieving higher levels of translation quality. Integrating linguopragmatic information into AI models could improve their ability to process expressive lexical units more effectively.

This study demonstrates that onomatopoeic units present a significant challenge for AI-based machine translation due to their complex interaction between sound, meaning, emotion, and context. Although modern artificial intelligence systems have achieved considerable progress in processing multilingual texts, they still face limitations in reproducing the pragmatic and stylistic functions of expressive lexical units. The analysis of English and Uzbek onomatopoeic expressions shows that AI systems are generally capable of transferring basic semantic information but often fail to maintain emotional colouring, cultural specificity, and contextual interpretation. These limitations are especially noticeable in literary and expressive discourse, where onomatopoeic units perform important communicative and stylistic roles. The findings emphasize the importance of incorporating linguopragmatic principles into future AI translation models. A deeper understanding of communicative intention, cultural context, and expressive meaning can contribute to improving machine translation quality and achieving more natural and accurate translations of sound-symbolic expressions. This research contributes to the fields of computational linguistics, translation studies, and pragmatics by highlighting the need for a closer interaction between artificial intelligence technologies and linguistic knowledge.

References

1. Bahdanau, D., Cho, K., & Bengio, Y. (2015). Neural machine translation by jointly learning to align and translate. *International Conference on Learning Representations (ICLR)*.

2. Brown, T. B., Mann, B., Ryder, N., Subbiah, M., Kaplan, J. D., Dhariwal, P., ... Amodei, D. (2020). Language models are few-shot learners. *Advances in Neural Information Processing Systems*, 33, 1877-1901.
3. Cho, K., Van Merriënboer, B., Gulcehre, C., Bahdanau, D., Bougares, F., Schwenk, H., & Bengio, Y. (2014). Learning phrase representations using RNN encoder-decoder for statistical machine translation. *Proceedings of EMNLP*, 1724-1734.
4. Goodfellow, I., Bengio, Y., & Courville, A. (2016). *Deep Learning*. MIT Press.
5. Koehn, P. (2020). *Neural Machine Translation*. Cambridge University Press.
6. Vaswani, A., Shazeer, N., Parmar, N., Uszkoreit, J., Jones, L., Gomez, A. N., Kaiser, Ł., & Polosukhin, I. (2017). Attention is all you need. *Advances in Neural Information Processing Systems*, 30, 5998-6008.