NEURAL MACHINE TRANSLATION OF PHRASAL VERBS IN LITERARY PARALLEL CORPORA: A CONTEXTUAL AND SEMANTIC ANALYSIS OF TRANSFORMER-BASED MODELS

Otamurodova Salomat Akhmedovna
Tashkent State Agrarian University Foreign language department teacher
E-mail: soloahmedovna@gmail.com

Annotatsiya. Ushbu maqola badiiy matnlarda frazali fe’llarni Neyron Mashina Tarjimasi (NMT) yordamida tarjima qilishdagi muammolar va takomillashtirish imkoniyatlarini ochib beradi. Bugungi kunda dunyo tobora ko‘proq o‘zaro bog‘langanligi sababli, ham to‘g‘ri, ham kontekstga asoslangan tarjimalarga bo‘lgan ehtiyoj ortib bormoqda. Transformer modellariga asoslangan zamonaviy NMT tizimlari qoidaga asoslangan statistik usullarga qaraganda ancha samarali hisoblanadi. Ushbu tizimlarning muhim afzalliklaridan biri —kontekstdan ma’no chiqarib olishi mumkinligidir. Bu esa ularni frazali fe’llar kabi murakkab ifodalarni tarjima qilishda juda foydali hisoblanadi. Shu sababli, bunday tarjimalarni qanday amalga oshirilishini o‘rganish va baholash uchun aniq va sodda metod ishlab chiqish muhim ahamiyatga ega.

Kalit so‘zlar: Neyron Mashina Tarjimasi (NMT), frazali fe’llar, adabiy tarjima, parallel korpuslar, transformer modellari, e’tibor (attention) mexanizmlari, idiomatik ifodalar, ko‘p so‘zli ifodalar (MWEs), kontekstual semantika, korpusga asoslangan tahlil.

Annotation. This study reveals at the problems and improvements in translating phrasal verbs in literary texts using Neural Machine Translation (NMT). Today, because the world is more connected, there is a higher need for translations that are both correct and sensitive to context. Modern NMT systems, especially those based on transformer models, work better than older rule-based and statistical methods. One important advantage of these systems is that they can understand meaning from context by using attention mechanisms. This makes them useful for translating difficult expressions like phrasal verbs. For this reason, it is important to have a clear and simple method to study and evaluate how these translations are done.

Keywords: Neural Machine Translation (NMT), phrasal verbs, literary translation, parallel corpora, transformer models, attention mechanisms, idiomatic expressions, multiword expressions (MWEs), contextual semantics, corpus-based analysis.

Аннотация. Данное исследование раскрывает проблемы и пути совершенствования перевода фразовых глаголов в литературных текстах с использованием нейронного машинного перевода (NMT). В современном мире, который становится всё более взаимосвязанным, возрастает потребность в переводах, которые являются не только точными, но и к контексту. Современные системы нейронного машинного перевода, особенно основанные на моделях-трансформерах, работают эффективнее по сравнению с более ранними методами, основанными на правилах и статистике. Одним из важных преимуществ таких систем является способность извлекать смысл из контекста. Это делает их особенно полезными для перевода сложных выражений, таких как фразовые

DOI: 10.5281/zenodo.19873874

глаголы. По этой причине важно разработать чёткий и простой метод для изучения и оценки того, как выполняются такие переводы.

Ключевые слова: *нейронный машинный перевод (NMT), фразовые глаголы, литературный перевод, параллельные корпусы, трансформерные модели, механизмы внимания, идиоматические выражения, многословные выражения (MWEs), контекстуальная семантика, корпусный анализ.*

Introduction. Today, global communication and interaction have reached an unprecedented level due to rapid advances in technology and transportation. People from different countries can easily share information, ideas and culture in real time. This high level of connectivity has significantly increased the need for effective and accurate translation across language. Because of this, people need fast and accurate translation[14]. Machine translation (MT) is a part of computer science that helps change text from one language to another automatically[15]. In the past, people used rule-based and statistical methods. Now, most systems use neural machine translation (NMT), which uses deep learning and gives better and more natural translations[16]. One important reason for this improvement is that NMT systems understand context better. They use something called attention mechanisms to connect words and phrases in the source and target languages. This helps them translate sentences more correctly. This is very important for phrasal verbs, because these expressions often have special meanings that are not literal[17]. For example, in the sentence *“She broke down in tears,”* the phrase *broke down* does not mean something physically broke. It means she became very emotional and started crying. A good NMT system can understand this meaning by looking at the whole sentence[18]. Modern NMT systems use transformer models and large parallel corpora. Parallel corpora are collections of texts with translations in two languages. These are very useful because they help the system learn how to translate difficult expressions like phrasal verbs[19]. Literary texts are especially useful for this purpose. They contain many phrasal verbs, both literal and figurative, and also include emotions and different writing styles. When NMT systems are trained on these texts, they can better understand meaning, style, and context, and produce more natural translations.

Methodology. Based on these ideas, this study uses a corpus-based method and focuses on teaching. It can be used in classes like translation studies, corpus linguistics, and applied linguistics. The methods in this study can be turned into simple teaching tools and learning activities for students. The corpus-based approach serves as a powerful foundation for data-driven learning (DDL), allowing learners to explore authentic language use through real examples20. By using parallel corpora of literary texts and their human translations, students can independently explore how phrasal verbs function in authentic contexts21. This aligns with modern approaches in Corpus Linguistics, where learners actively discover linguistic patterns rather than passively receiving rules. Students can be guided to identify phrasal verbs, analyze their meanings,

[14] Munday J. Introducing Translation Studies. Routledge.2016.p.8.
[15] Koehn,P. Neural Machine Translation. Cambridge University Press.2020.p.3
[16] Bengio Y., Goodfellow I. & Courville A. Deep Learning. MIT Press. 2016.p.210.
[17] Greeraerts D. Theories of Lexical Semantics. Oxford University Press. 2010. P.125.
[18] Goodfellow I. et al. Deep Learning. MIT Press. 2016.p.412.
[19] Baker M. In Other words: A Course book on Translation. Routledge. 1995. P. 230.
[20] Abduraxmonova N. Korpus lingvistikasi va zamonaviy til texnologiyalari.Toshkent. 2021.p.88.
[21] Rahmatullayev Sh. O‘zbek tilining izohli lug‘ati. Toshkent. P.95.

and compare translation equivalents across languages. The integration of Neural Machine Translation (NMT) systems introduces a technological dimension to the classroom. Tools based on Neural Machine Translation can be used as didactic instruments rather than mere translation tools. Machine translations can be compared with human translations to see differences in accuracy, fluency, and meaning22. This helps students notice common mistakes, especially with phrasal verbs, which often have non-literal meanings and depend on context. By studying these differences, students can find errors, changes in meaning, or loss of style in machine translations23. They can then suggest better translations by understanding the context more deeply. This process helps them understand phrasal verbs better and improves their translation skills. It also develops critical thinking and post-editing skills, which are important in modern translation learning.

Multiword expressions (MWEs) can be used as practical classroom activities. Students can work with real texts and mark MWEs in a corpus. This helps them recognize common phrases in real language. They can also group phrasal verbs by meaning, for example, separating literal and figurative uses. In addition, students can study how often these expressions appear in different texts and how their use changes in different contexts. These activities help students understand phrasal verbs more clearly and build their research and analysis skills (John Sinclair, 1991, p. 110).

Results and Discussion. The analysis reveals that transformer-based NMT systems demonstrate a high level of fluency and grammatical accuracy when translating literary texts. In many cases, these systems successfully capture the contextual meaning of phrasal verbs, particularly when the expressions are frequent and well-represented in training corpora. The use of attention mechanisms enables effective alignment between source and target segments, contributing to more coherent and natural translations.

However, the findings also indicate several persistent challenges. NMT systems tend to struggle with:

- ✓ Idiomatic and figurative phrasal verbs, especially in emotionally loaded or stylistically marked contexts
- ✓ Rare, novel, or creatively formed expressions, including neologisms derived from the “V+Adv” model
- ✓ Preservation of stylistic nuance, often replacing expressive phrasal verbs with neutral or paraphrased equivalents

For example, expressions like *“break down in tears”* can sometimes be translated word-for-word, which may lose part of the meaning and emotion. This study shows that language features such as having many meanings (polysemy), using metaphors, and depending on context are very important for correct translation. Phrasal verbs with particles like *up* and *out* can change meaning in complex ways, and this makes them difficult for machine systems to translate correctly.

1. ***“give up”*** – It does not mean physically “giving something up,” but means ***to stop trying***. *Example: He decided to give up smoking.*
2. ***“break out”*** – It does not mean breaking something physically, but means ***to start suddenly (e.g., war, fire, or escape)***. *Example: A fire broke out in the building.*

[22] Koehn,P. Neural Machine Translation. Cambridge University Press.2020.p.145.
[23] Pym A. Exploring Translation Theories. Routledge. 2010. P.112.

These examples show how particles like **up** and **out** change the meaning of phrasal verbs in a non-literal way. So good parallel corpora are very important, especially for literary texts. Human translations in these corpora give useful examples of how to translate difficult, idiomatic, and emotional expressions. They also show different translation strategies that machine systems try to follow, but not always successfully. The results suggest that adding more linguistic knowledge to NMT systems—such as sentence structure, meaning roles, and semantic patterns—can improve translation quality. Also, training systems on specific types of texts and using annotated corpora of phrasal verbs can help them handle complex expressions better.

Although modern NMT systems work well in general, they still have problems when translating phrasal verbs in literary texts. These problems are caused by complex meanings, different writing styles, and new word uses. To solve these issues, it is important to combine advanced machine learning methods with linguistic knowledge, especially in word formation, meaning, and cognitive linguistics.

Conclusion. In conclusion, Neural Machine Translation (NMT) systems have improved a lot in translating texts. They can make translations that are clear and grammatically correct, and they often understand the meaning of sentences. However, phrasal verbs in literary texts are still difficult for these systems. These expressions often have special meanings that depend on context. Because of this, machines sometimes make mistakes, especially with figurative meanings, style, and unusual expressions. The study also shows that parallel corpora with human translations are very helpful because they give good examples of correct language use and help improve translation quality.

In addition, this research shows that students and translators can use NMT as a learning tool, not only as a translation tool. It can help them compare different translation versions and better understand language use in real contexts.

REFERENCES:

1. Munday J. Introducing Translation Studies. Routledge. 2016. p.8.
2. Koehn, P. Neural Machine Translation. Cambridge University Press. 2020. p.3
3. Bengio Y., Goodfellow I. & Courville A. Deep Learning. MIT Press. 2016.p.210.
4. Greeraerts D. Theories of Lexical Semantics. Oxford University Press. 2010. P.125.
5. Goodfellow I. et al. Deep Learning. MIT Press. 2016.p.412.
6. Baker M. In Other words: A Course book on Translation. Routledge. 1995. P. 230.
7. Abduraxmonova N. Korpus lingvistikasi va zamonaviy til texnologiyalari. Toshkent. 2021. p.88.
8. Rahmatullayev Sh. O‘zbek tilining izohli lug‘ati. Toshkent. P.95.
9. Koehn, P. Neural Machine Translation. Cambridge University Press.2020. p.145.
10. Pym A. Exploring Translation Theories. Routledge. 2010. P.112.