A COMPARATIVE AND COMPUTATIONAL STUDY OF PROVERBS AND IDIOMS IN ENGLISH AND UZBEK

Ismoilova Marxabo Agzam qizi

ESL teacher at Uzbekistan State World Languages University,

e-mail: markhaborakhmonova98@gmail.com

Abstract: Proverbs and idioms are integral parts of language and culture, encapsulating collective wisdom, values, and figurative expression. Their translation between languages poses a unique challenge due to structural, semantic, and cultural differences. This study investigates the comparative features of English and Uzbek proverbs and idioms, while also testing the capacity of machine translation (MT) systems to recognize and translate them accurately. Using a bilingual dataset of 300 idioms and proverbs, the research evaluates human translation strategies alongside the performance of neural MT systems such as Google Translate and Yandex Translate. Findings reveal that while human translators employ functional equivalence and cultural substitution to maintain idiomatic meaning, MT systems frequently produce literal translations, resulting in semantic distortions and cultural loss. The study highlights the importance of integrating computational phraseology and cultural annotation into NLP models to improve idiom recognition and translation.

Keywords: proverbs, idioms, translation, English, Uzbek, computational linguistics, phraseology, machine translation

Introduction

Proverbs and idioms are among the most culturally embedded aspects of language. They express metaphorical meanings and reflect shared cultural experiences. For example, the English idiom "the ball is in your court" conveys responsibility, while the Uzbek proverb "Ko'p eshit, oz so'zla" ("Listen much, speak little") emphasizes wisdom through restraint.

Translating such expressions is not straightforward, as literal renderings often fail to capture intended meanings. Machine translation systems, despite advancements in neural networks, still struggle with figurative and culture-specific language. This raises questions about how idioms and proverbs are best translated and whether computational models can be improved to handle them effectively.

The present study addresses these issues by conducting a comparative analysis of English and Uzbek idioms and proverbs, while also evaluating the strengths and limitations of computational translation systems in handling figurative language.

Literature Review

Idioms are fixed multi-word expressions whose meaning cannot be deduced from individual words (Fernando, 1996). Proverbs, on the other hand, are complete statements expressing general truths or cultural wisdom (Mieder, 2004). Both represent the figurative dimension of language and are deeply rooted in culture.

According to Baker (1992), translators may use several strategies. Among the most common strategies are literal translation (rarely successful for idioms), paraphrasing (rendering the meaning in descriptive form), functional equivalence (using a target-language idiom with similar meaning), cultural substitution (replacing the source idiom with a culturally familiar expression).

Venuti's (1995) concepts of domestication and foreignization are also relevant, as translators must decide whether to preserve the foreign cultural element or adapt it for accessibility.

However**,** studies in NLP highlight that MT systems face persistent problems with idioms. Sag et al. (2002) note that idioms often violate compositionality, making them difficult for algorithms. Neural MT systems (Koehn, 2020) show improvement in fluency but often misinterpret figurative meaning. For low-resource languages like Uzbek, the problem is amplified due to limited training data.

While research exists on idioms in English or Uzbek separately, few studies combine comparative analysis with computational evaluation. Therefore, this study addresses both.

Methodology

The study combines comparative linguistic analysis with computational testing. The following sources were used for analysis.

-English sources: Oxford Dictionary of English Idioms, Bartlett's Familiar Quotations.

-Uzbek sources: "O‘zbek maqollari" collection, phraseological dictionaries, and classical Uzbek literature.

-Dataset: 150 English idioms/proverbs and 150 Uzbek idioms/proverbs, totaling 300.

Idioms and proverbs were categorized by Structural type (phrasal, clausal, sentential) , Semantic type (metaphorical, cultural, universal) and Pragmatic function (advice, criticism, encouragement).

Human translations analyzed for strategies used (functional equivalence, paraphrase, etc.).

-Computational translations tested on Google Translate and Yandex Translate.

-Accuracy measured against human translations.

Findings

English idioms often draw from sports, history, and classical mythology (Achilles' heel, throw in the towel). Conversely, Uzbek idioms and proverbs frequently reflect agrarian life, collectivism, and moral values (Ko‘p qozon qaynasa ham bir qopqa tushadi – "Many pots boil, but all pour into one dish").

Human translators relied on functional equivalence when possible (e.g., translating the last straw as sabr kosasi to‘ldi). When no equivalent existed, paraphrase or cultural substitution was applied.

Literal translation was the default outcome for most idioms. For example, kick the bucket became chelakni tepmoq, which is semantically meaningless in Uzbek. Proverbs were often translated word-

for-word, losing figurative meaning. And accuracy rates present the following results according to which Google Translate is usually 38% correct and Yandex Translate is 42% correct.

These examples highlight the validity of the findings aboce. English → Uzbek: "Don't count your chickens before they hatch" → MT: Jo'jalarni tuxumdan chiqmasdan sanamang (literal), Human: Baliqchi daryoda, baliq qozonda emas. Uzbek → English: "Otning o'limi itka bayram" → MT: The death of a horse is a holiday for a dog (literal), Human: One's misfortune is another's gain.

Discussion

The findings reveal the central role of cultural background in interpreting idioms and proverbs. Human translators demonstrate flexibility, drawing on cultural parallels and pragmatic judgment. By contrast, MT systems lack cultural awareness and treat idioms literally, leading to misinterpretation.

The results suggest that NLP systems must incorporate phraseological databases, semantic embeddings trained on idioms, and cultural annotation to achieve better performance. Hybrid systems that combine statistical recognition of idiomatic phrases with neural contextual models may provide more accurate translations.

Conclusion

This study confirms that proverbs and idioms pose significant challenges in cross-linguistic translation. English idioms are rooted in historical and mythological references, while Uzbek idioms reflect agrarian and communal values. Human translators navigate these differences using strategies like functional equivalence and cultural substitution. Machine translation, however, continues to falter, offering literal renderings that obscure intended meanings.

Improving idiom and proverb translation in MT systems requires integrating computational phraseology, enriched bilingual corpora, and culturally informed models. Such advancements will not only enhance translation accuracy but also preserve cultural depth in cross-linguistic communication.

References

1. Baker, M. (1992). In Other Words: A Coursebook on Translation. Routledge.
2. Fernando, C. (1996). Idioms and Idiomaticity. Oxford University Press.
3. Koehn, P. (2020). Neural Machine Translation. Cambridge University Press.
4. Mieder, W. (2004). Proverbs: A Handbook. Greenwood Press.
5. Sag, I. A., Baldwin, T., Bond, F., Copestake, A., & Flickinger, D. (2002). Multiword expressions: A pain in the neck for NLP. Proceedings of CICLing 2002, 1–15.
6. Venuti, L. (1995). The Translator's Invisibility: A History of Translation. Routledge.