

Transformation Of Pejorative Vocabulary In AI-Generated Texts: A Comparative Analysis Of Russian- And Uzbek-Language LLMS

Matluba Qulfiddinovna Zaripova

Termez State University

Associate Professor, Department of Russian Linguistics

Abstract

The article examines the transformation of pejorative vocabulary in AI-generated texts through a comparative analysis of Russian- and Uzbek-language large language models (LLMs). It explores how modern language models interpret, transfer, and neutralize pejorative units, as well as the influence of linguistic and cultural factors on the quality of generated texts. Particular attention is paid to the similarities and differences in the processing of emotionally evaluative vocabulary in Russian and Uzbek.

Keywords: artificial intelligence, large language models (LLMs), pejorative vocabulary, text generation, Russian language, Uzbek language, comparative analysis, natural language processing.

The rapid development of artificial intelligence (AI) and natural language processing (NLP) technologies has significantly expanded the capabilities of automated text generation. Modern large language models (LLMs) are capable of producing coherent texts, analyzing linguistic units, and accounting for the structural and semantic features of different languages. In this context, particular attention has been given to how these models process pejorative vocabulary, which carries negative emotional and evaluative connotations.

Pejorative units play an important role in expressing a speaker's attitude toward a person, their behavior, or various social phenomena. However, their meaning depends not only on their lexical semantics but also on context, stylistic features, and the national and cultural characteristics of a language. Consequently, the accurate processing of pejorative vocabulary remains one of the key challenges in contemporary computational linguistics.

The relevance of this study lies in the need for a comparative analysis of Russian- and Uzbek-language LLMS in generating texts containing pejorative vocabulary. The aim of the article is to identify the patterns of

transformation of pejorative units and to evaluate the ability of modern language models to convey negative evaluation in Russian and Uzbek.

Modern large language models (LLMs) have significantly expanded the possibilities of automated text processing and generation. They are capable of analyzing context, identifying the emotional and evaluative connotations of linguistic units, and selecting the most appropriate way to convey them. However, pejorative vocabulary remains one of the most challenging categories for computational processing, as its meaning depends not only on lexical semantics but also on the communicative situation, stylistic features of the text, and the national and cultural context.

During text generation, artificial intelligence may transform pejorative units in different ways. In some cases, the model preserves the negative evaluation when it constitutes an essential element of a literary or journalistic text. In other cases, LLMS tend to mitigate the expression by replacing harsh or offensive words with more neutral alternatives. For example, the Russian word *bolvan* ("blockhead") may be rendered as *an unreasonable person*, *idiot* as a

person who made a mistake, and scoundrel as a dishonest person. A similar tendency can be observed in Uzbek, where words such as *ahmoq*, *nodon*, or *tentak* may be replaced with expressions such as *an inexperienced person*, *a person with insufficient knowledge*, or *someone who made a mistake*. Such transformations reduce the degree of negative evaluation while preserving the overall meaning of the utterance.

A comparative analysis of Russian- and Uzbek-language LLMs demonstrates that the quality of pejorative vocabulary processing largely depends on the size of the training data and the availability of linguistic resources. Russian-language models generally recognize emotional nuances, stylistic features, and implicit meanings more accurately because they are trained on larger text corpora. In contrast, linguistic resources for Uzbek remain relatively limited, which may cause models to encounter difficulties when processing culturally specific vocabulary, colloquial expressions, and dialectal forms. The processing of pejorative idiomatic expressions is of particular importance. Their meaning cannot be determined through the literal interpretation of individual words, as it is conveyed by the expression as a whole. For example, the Russian idiom *bez tsarya v golove* ("without a tsar in one's head") describes a person as foolish or irresponsible, whereas the Uzbek expression *ilonning yog'ini yalagan* literally means "one who has licked a snake's fat" and refers to an extremely cunning and deceitful person. If a language model fails to recognize an idiom as a fixed expression, it may interpret it literally, leading to semantic distortion.

Additional challenges arise in the processing of irony, sarcasm, and implicit negative evaluation. For instance, the Russian sentence "Well, you're a genius!" may express either genuine praise or

sarcasm, depending on the context. A similar phenomenon exists in Uzbek, where the expression "*Juda aqlli ekansan!*" ("You're so smart!") is often used ironically. To interpret such utterances correctly, a language model must consider not only the lexical meaning of individual words but also the broader context, communicative intention, and pragmatic features of the discourse.

It should be emphasized that contemporary LLMs continue to improve through training on increasingly diverse and extensive text corpora. As a result, they have become more effective at recognizing emotionally charged vocabulary, distinguishing between stylistic registers, and accounting for cultural characteristics of language. Nevertheless, errors in processing pejorative vocabulary cannot yet be completely eliminated, particularly when dealing with culturally specific expressions, colloquial speech, or literary texts rich in metaphorical language and implicit evaluation.

In conclusion, it can be noted that modern Russian- and Uzbek-language Large Language Models (LLMs) demonstrate a relatively high level of proficiency in processing pejorative vocabulary. At the same time, the comparative analysis shows that the quality of its transformation directly depends on the context, the scope of linguistic resources, and the model's ability to take into account the cultural and pragmatic characteristics of each language. Further development of artificial intelligence technologies and the expansion of national language corpora will improve the accuracy of text generation and ensure a more appropriate rendering of emotionally evaluative vocabulary.

The conducted analysis has shown that modern LLMs possess extensive capabilities for processing and generating texts containing pejorative vocabulary. They are able to consider contextual

factors, recognize emotionally evaluative linguistic units, and, when necessary, transform them while preserving the overall meaning of the utterance.

A comparison of Russian- and Uzbek-language models demonstrated that the quality of pejorative vocabulary processing depends on the volume of linguistic resources, cultural characteristics, and the context of use. Despite the significant progress achieved, the accurate interpretation of nationally specific and phraseological pejorative units remains an important challenge. Further advancement of LLMs will contribute to a more precise representation of emotionally evaluative vocabulary in both languages.

References

- Aitchison, J. (2018). The articulate mammal: An introduction to psycholinguistics (6th ed.). Routledge.
- Fill, C. (2013). Marketing communications: Brands, experiences and participation (6th ed.). Pearson Education.
- Goddard, A. (2002). The language of advertising (2nd ed.). Routledge.
- Zaripova, M. K. (2022). Improving the process of preparing future teachers for innovative pedagogical activity based on an integrative approach (PhD dissertation). Termez State University, Termez, Uzbekistan.
- Zaripova, M. Q. (2021). The theory of using new information technologies in the professional activity of future teachers. Academic Research in Educational Sciences, 2(10), 98–987.
- Zaripova, M. K. (2023). The content and manifestations of pedagogical innovative activity of the future teacher. JournalNX, 7(10), 296–300.
- Zaripova, M. K. (2023). The content and manifestations of pedagogical innovative activity of the future teacher. JournalNX, 7(10), 296–300.
- Zaripova, M. K. (2025). Definition of the concept of "Wealth" in Russian and Uzbek phraseology and paremias. European International Journal of Multidisciplinary Research and Management Studies, 5(2), 15–17.
- Zaripova, M. K. (2024). Russian language: Textbook for the educational program in Law (Part I, pp. 103–141). Termez.
- Zaripova, M. K. (2024). Russian language: Textbook for the educational program in Law (Part II, pp. 203–310). Termez.
- Zaripova, M. K. (2025). Russian language: Textbook for the educational program in History (Part I, pp. 119–164). Termez.
- Zaripova, M. K. (2025). Russian phraseology and paremiology: Textbook for the educational program in Linguistics (Russian Language). Termez.