© 2026 International Conference on Multidisciplinary Sciences and Innovations . All Rights Reserved.

ПРИРОДА ФЕНОМЕНА ЭВФЕМИЗМА: ЭВОЛЮЦИЯ ОТ ТРАДИЦИОННОЙ ЛИНГВИСТИКИ К КОМПЬЮТЕРНОЙ И ВЫЧИСЛИТЕЛЬНОЙ ЛИНГВИСТИКЕ

EVFEMIZM HODISASI TABIATI: AN'ANAVIY TILSHUNOSLIKDAN KOMPYUTER VA HISOBLASH LINGVISTIKASIGA EVOLYUTSIYASI

THE NATURE OF THE EUPHEMISM PHENOMENON: EVOLUTION FROM TRADITIONAL LINGUISTICS TO COMPUTATIONAL LINGUISTICS

Азимова Гузаль Бахтияровна
Учитель школы № 58 Музрабатского района Сурхандарьинской области

Аннотация

В данной статье рассматривается эволюция феномена эвфемизма от традиционных лингвокультурологических и риторических концепций до парадигмы современной компьютерной и вычислительной лингвистики. Анализируются механизмы табуирования, прагматической митигации и когнитивного маскирования в естественном языке, а также специфика их формализации в системах автоматической обработки естественного языка (NLP). Особое внимание уделяется функционированию эвфемизмов в дискурсе искусственного интеллекта, модерации контента, детекции токсичности и обработке пользовательских промптов в больших языковых моделях (LLM).

Ключевые слова: эвфемизм, табу, компьютерная лингвистика, вычислительная лингвистика, обработка естественного языка (NLP), большие языковые модели (LLM), контент-модерация, прагматика.

Annotatsiya

Mazkur maqolada evfemizm hodisasining ontologik tabiati, uning an'anaviy lisoniy-madaniy va ritorik ta'limotlardan tortib zamonaviy kompyuter va hisoblash lingvistikasi paradigmasigacha bo'lgan nazariy-metodologik evolyutsiyasi tadqiq etiladi. Tabu, muloqot odobi (lingvistik mitigatsiya) hamda kognitiv ma'no pardalash vositasi sanalgan evfemizmlarning tabiiy tildan raqamli muhitga ko'chish mexanizmlari, shuningdek, tabiiy tilni avtomatik qayta ishlash (NLP) modellarida formal tahlil qilinishi yoritiladi. Tadqiqotda katta til modellarida (LLM) axloqiy filtrlarni boshqarish, prompt muhandisligida evfemik birliklarning roli hamda ularni avtomatik aniqlash algoritmlarining istiqbollari ilmiy jihatdan asoslanadi.

Kalit so'zlar: evfemizm, tabu, kompyuter tilshunosligi, hisoblash lingvistikasi, NLP, katta til modellari (LLM), xavfsizlik filtrlari, pragmatika, kognitiv semantika.

Annotation

This article examines the evolution of the euphemism phenomenon from traditional linguacultural and rhetorical paradigms to modern computational linguistics and natural language processing. The mechanisms of tabooing, pragmatic mitigation, and cognitive masking in natural language are explored alongside their formalization within artificial intelligence systems. Special emphasis is placed on the role of euphemisms in prompt engineering, automated content moderation, toxicity detection, and safety alignment in large language models (LLMs).

© 2026 International Conference on Multidisciplinary Sciences and Innovations . All Rights Reserved.

Keywords: euphemism, taboo, computational linguistics, natural language processing (NLP), large language models (LLM), content moderation, safety filters, pragmatics, prompt engineering.

The multi-century evolution of human linguistic thought demonstrates that language functions not merely as a nominative tool directly reflecting objective reality, but also as an intricate cultural-pragmatic mechanism regulating social interactions. Within this system, the phenomenon of euphemism constitutes one of the central categories intimately linked with sociolinguistic, cognitive, and psychological factors. In the archaic worldview, the process of euphemization emerged from primitive taboos, mystical fear, and the instinctual necessity of seeking protection against sacred powers. Believing in the ontological, material potency of words, early humans were driven to express crude, offensive, or hazardous concepts through veiled, covert linguistic means. An animistic belief prevailed that uttering the direct designation of a tabooed entity would inevitably provoke unforeseen catastrophes or the wrath of malevolent spirits. Consequently, members of archaic societies established the tradition of substituting benign designations for concepts related to death, disease, predatory beasts, and divine forces[1]. With the historical civilizing of human society, this primitive imperative transformed into cultural-ethical codes of etiquette, culminating in communicative and linguistic mitigation. As cultural stratification within society intensified, adhering to norms of communicative politeness, safeguarding interlocutors from offense, and complying with established social conventions acquired paramount importance. Thus, within traditional linguistics, euphemisms were long interpreted through a narrow lens as mere figurative embellishments belonging to stylistics, rhetoric, and lexicology. Early textbooks and scholarly treatises primarily treated euphemisms as a variety of synonymy or an expressive artistic device designed to heighten rhetorical eloquence[2].

In the second half of the twentieth century, the consolidation of the anthropocentric paradigm fundamentally altered scholarly perspectives on euphemism. Linguists began to examine the language system in direct conjunction with human cognitive activity, social agency, and communicative environments. Consequently, researchers recognized euphemisms not as superficial verbal ornaments substituting for vulgarities, but as strategic communicative mechanisms instrumental in realizing the speaker's pragmatic intentions. It became increasingly evident that every euphemistic formulation embodies a deliberate communicative intent, a distinct social stance, and an objective to exert psychological influence over the recipient. Politeness Theory, formulated by R.Lakoff, P.Brown, and S.Levinson, conceptualized euphemisms as essential discursive instruments for safeguarding social status and communicative face[3]. Under this framework, speakers systematically avoid direct pejorative lexis to avert threatening either the positive or negative face of their interlocutors. As a result, linguistic mitigation emerges as a foundational mechanism for maintaining equilibrium in interpersonal interaction. Proponents of cognitive linguistics, in turn, elucidated euphemization as an operation of conceptual remodeling of reality. Human cognition resists direct encounters with adverse associations, systematically replacing them with mitigated metaphorical mappings. Analyzed through the lens of George Lakoff and Mark Johnson's Conceptual Metaphor Theory, a euphemism represents the cognitive perception and restructuring of a target domain via an alternative source domain. Thus, across traditional and cognitive paradigms, euphemisms were rigorously investigated as complex linguistic phenomena grounded in semantic shifts, metonymy, and metaphorical transference[4].

However, the dawn of the information age and the rapid integration of digital technologies necessitated situating the phenomenon of euphemism within an entirely new scientific coordinate system. The emergence of virtual communication environments fundamentally reshaped the established conventions and structural parameters of human verbal behavior. As the dominant share of societal interaction migrated to the internet, instant messengers, social networks, and global digital

© 2026 International Conference on Multidisciplinary Sciences and Innovations . All Rights Reserved.

platforms, the imperative for automated processing of natural language data escalated. The exponential proliferation of large-scale text corpora exposed the limitations of traditional descriptive methodologies in processing modern linguistic data. At this juncture, computational linguistics undertook the empirical study of euphemisms, modeling their distributional frequency, syntagmatic relations, and digital dynamics using statistical and mathematical frameworks. Corpus linguistic methodologies including concordance extraction, frequency profiling, and collocation indexing mapped the empirical distribution of euphemistic units across real-world texts[5]. Massive electronic textual repositories provided an empirical foundation for tracing real-time lexical shifts within living language. In tasks such as automated text indexing, information retrieval, and computational lexicography, euphemisms emerged as formidable polysemic challenges for machine interpretation. Because conventional search algorithms rely predominantly on rigid lexical matching, they frequently fail to classify semantically veiled expressions accurately. This failure stems from the fact that euphemistic units diverge from their primary denotative definitions, activating secondary connotations exclusively within specialized contexts. The displacement of denotative meaning by contextual connotation generates severe semantic ambiguity that challenges computational architectures[6].

With advancements in computational linguistics and Natural Language Processing (NLP), this dilemma has acquired critical relevance. In engineering computational models of natural language, algorithmically capturing nuanced layers of semantic meaning became a primary objective. Contemporary systems are tasked not merely with indexing static euphemisms, but with dynamically detecting and accurately decoding them within real-time text streams. This analytical process requires algorithms to factor in the syntactic configurations and semantic valencies of every constituent lexical item. Machine learning architectures and deep neural networks deployed for sentiment analysis confront severe operational hurdles due to euphemistic expressions. Automated sentiment classification algorithms typically estimate the polarity of texts based on pre-assigned positive or negative lexical weights. Consequently, when a softened euphemistic construction is employed in place of overt negative lexis, neural networks frequently misclassify the content as positive or neutral. Such misclassifications induce algorithmic distortion in evaluating public opinion, social feedback, and overall textual sentiment. This systematic bias results in critical computational errors, ultimately misrepresenting the author's authentic communicative intentionality. The distortion of pragmatic intentionality undermines the overall academic validity and operational reliability of outputs generated by intelligent systems[7].

Furthermore, the widespread deployment of Large Language Models (LLMs) and generative artificial intelligence architectures has elevated the functional status of euphemisms to an unprecedented plane. Generative systems such as GPT, Claude, and Gemini have achieved performance benchmarks approaching human-level fluency in natural language understanding and generation. Within modern artificial intelligence architectures, developers integrate stringent safety guidelines alongside content filtering mechanisms designed to suppress undesirable generations. These automated filters systematically intercept and neutralize prompts associated with aggression, violence, harassment, and illicit instructions. Users, operating within the domain of prompt engineering, increasingly leverage euphemisms as sophisticated discursive mechanisms to circumvent these ethical guardrails and articulate sensitive objectives[8]. Framing restricted concepts through oblique euphemistic analogues and metaphors has evolved into a widespread adversarial technique for bypassing model alignment. While neural networks reliably reject overtly prohibited directives, they struggle to identify adversarial queries cloaked in complex figurative and euphemistic phrasing. An algorithm calibrated to surface-level lexical statistics frequently fails to detect latent risks buried within subtext, inadvertently generating prohibited content. This vulnerability

© 2026 International Conference on Multidisciplinary Sciences and Innovations . All Rights Reserved.

necessitates advanced cyberlinguistic analysis and the training of next-generation neural architectures to safeguard information security. Cyberlinguistics seeks to integrate cognitive-pragmatic rules into computational algorithms to detect adversarial manipulations of natural language.

Simultaneously, the automated moderation of toxic discourse, harassment, and hate speech across social platforms constitutes one of the most challenging frontiers in computational linguistics. Confronted with rigorous automated censorship, malicious actors routinely substitute explicit slurs with covert, discriminatory euphemisms. Harmful content is thus propagated through veiled sarcasm and superficially benign formulations, effectively evading automated moderation filters. To counteract this evasion, computational linguistics deploys contextual vector embeddings and transformer-based architectures. Deep learning representations such as Word2Vec, BERT, and RoBERTa enable the multi-dimensional geometric mapping of lexical tokens. These models are optimized to decode latent euphemistic senses by evaluating both left and right syntagmatic contexts surrounding a given token. Through contextualized vectorization, algorithms can expose the caustic, discriminatory subtext underlying expressions that appear lexically neutral on the surface[9].

At present, computational linguistics focuses on formalizing euphemisms into structured logical frameworks, constructing dedicated domain ontologies, and embedding them into Knowledge Graphs. Knowledge Graphs facilitate the systematic representation of latent associative links across semantic concepts through multi-tiered relational nodes. By its very nature, the euphemism represents a dynamic, fluid, and culturally sensitive linguistic entity. Rather than remaining static, it continuously undergoes functional renewal across historical epochs and shifting socio-cultural climates[10]. A lexical unit serving as a polite euphemism today often undergoes pejoration over time, ultimately deteriorating into an explicit dysphemism. Integrating such dynamic semantic drift into automated systems poses substantial lexicographical and algorithmic challenges for computational linguists.

In conclusion, the phenomenon of euphemism has traversed a profound evolutionary trajectory, originating from primitive taboos and classical rhetorical arts to emerge as a cornerstone of modern information technology and artificial intelligence. Within contemporary computational and corpus linguistics, euphemisms transcend simple lexical classifications; they serve as critical benchmarks for deep natural language understanding, semantic security, and ethical filter alignment[11]. The mathematical and computational formalization of these units significantly enhances a machine's capacity to decode human intentionality accurately. Looking forward, developing automated analytical models of euphemisms across diverse typological systems—most notably for the agglutinative Uzbek language will propel computational linguistics and practical artificial intelligence applications to new methodological heights. Integrating the rich morphological and semantic resources of the Uzbek language into digital architectures will ensure that national natural language processing systems align with international standards in interpreting intricate linguistic phenomena.

REFERENCES:

1. Allan K., Burridge K. Euphemism and Dysphemism: Language Used as Shield and Weapon. – Oxford: Oxford University Press, 1991. – 263 p.
2. Frazer J.G. The Golden Bough: A Study in Magic and Religion. – London: Macmillan, 1922. – 756 p.
3. Ismatullayev N. O‘zbek tilida evfemizmlar. – Toshkent: Fan, 1964. – 128 b.
4. Lakoff R. The Logic of Politeness: or, Minding Your P’s and Q’s // Papers from the Ninth Regional Meeting of the Chicago Linguistic Society, 1973. – pp. 292–305.
5. Galperin I.R. Stylistics. – Moscow: Higher School, 1977. – 334 p.

© 2026 International Conference on Multidisciplinary Sciences and Innovations . All Rights Reserved.

6. Searle J.R. Expression and Meaning: Studies in the Theory of Speech Acts. – Cambridge: Cambridge University Press, 1979. – 187 p.
7. Brown P., Levinson S.C. Politeness: Some Universals in Language Usage. – Cambridge: Cambridge University Press, 1987. – 345 p.
8. Lakoff G., Johnson M. Metaphors We Live By. – Chicago: University of Chicago Press, 1980. – 242 p.
9. Omonov P. O‘zbek tili leksikasining semantik taraqqiyoti. – Toshkent: Mumtoz so‘z, 2017. – 180 b.
10. Crystal D. Language and the Internet. – Cambridge: Cambridge University Press, 2006. – 304 p.
11. Manning C.D., Schütze H. Foundations of Statistical Natural Language Processing. – Cambridge: MIT Press, 1999. – 680 p.