ВЕСТНИК РОССИЙСКОГО УНИВЕРСИТЕТА ДРУЖБЫ НАРОДОВ.

СЕРИЯ: ТЕОРИЯ ЯЗЫКА. СЕМИОТИКА. СЕМАНТИКА

2026 Том 17 № 2

Специальная тема номера:

АКТУАЛЬНЫЕ ВЫЗОВЫ ИСКУССТВЕННОГО ИНТЕЛЛЕКТА В ПРИКЛАДНОЙ ЛИНГВИСТИКЕ

Приглашенные редакторы ***А.В. Джунковский и К.П. Жикулина***

DOI: 10.22363/2313-2299-2026-17-2

http://journals.rudn.ru/semiotics-semantics

Научный журнал

Издается с 2010 г.

Издание зарегистрировано Федеральной службой по надзору в сфере связи, информационных технологий и массовых коммуникаций (Роскомнадзор)

Свидетельство о регистрации ПИ № ФС 77-61215 от 30.03.2015 г.

Учредитель: Федеральное государственное автономное образовательное учреждение высшего образования «Российский университет дружбы народов имени Патриса Лумумбы»

ГЛАВНЫЙ РЕДАКТОР
Денисенко Владимир Никифорович, доктор филологических наук, профессор, Российский университет дружбы народов, Москва, Российская Федерация
E-mail: denisenko-vn@rudn.ru

ЗАМЕСТИТЕЛЬ ГЛАВНОГО РЕДАКТОРА
Новоспасская Наталья Викторовна, кандидат филологических наук, доцент, Российский университет дружбы народов, Москва, Российская Федерация
E-mail: novospasskaya-nv@rudn.ru

ОТВЕТСТВЕННЫЙ СЕКРЕТАРЬ
Лазарева Олеся Викторовна, кандидат филологических наук, доцент, Российский университет дружбы народов, Москва, Российская Федерация
E-mail: lazareva-ov@rudn.ru

ЧЛЕНЫ РЕДАКЦИОННОЙ КОЛЛЕГИИ

Беднарова-Гибова Клаудиа, доктор филологических наук, доцент, Институт британских и американских исследований Университета г. Прешов, Прешов, Словакия

Болдырев Николай Николаевич, доктор филологических наук, профессор, Тамбовский государственный университет имени Г.Р. Державина, Тамбов, Российская Федерация

Владимирова Татьяна Евгеньевна, доктор филологических наук, профессор, Российский университет дружбы народов, Москва, Российская Федерация

Джусупов Маханбет, доктор филологических наук, профессор, заслуженный профессор, Узбекский государственный университет мировых языков, Ташкент, Республика Узбекистан

Евграфова Юлия Александровна, доктор филологических наук, доцент, Национальный исследовательский университет «МЭИ», Москва, Российская Федерация

Иванов Евгений Евгеньевич, доктор филологических наук, доцент, МГУ им. А.А. Кулешова, Могилев, Республика Беларусь

Карасик Владимир Ильич, доктор филологических наук, профессор, ГИРЯ им. А.С. Пушкина, Москва, Российская Федерация

Маслова Валентина Авраамовна, доктор филологических наук, профессор, Витебский государственный университет им. П.М. Машерова, Витебск, Республика Беларусь

Монфорте Дюпре Роберто, доктор филологических наук, доцент, Университет Страны Басков, Витория-Гастейс, Испания

Новикова Марина Львовна, доктор филологических наук, доцент, Российский университет дружбы народов, Москва, Российская Федерация

Поплавская Татьяна Викторовна, доктор филологических наук, профессор, Минский государственный лингвистический университет, Минск, Республика Беларусь

Синячкин Владимир Павлович, доктор филологических наук, доцент, Российский университет дружбы народов, Москва, Российская Федерация

Темиргазина Зифа Какбаевна, доктор филологических наук, профессор, Павлодарский государственный педагогический институт, Павлодар, Республика Казахстан

Цзюй Юнь Шэн, доктор филологических наук, доцент, Хэйлунцзянский университет, Харбин, Китайская Народная Республика

Вестник Российского университета дружбы народов. Серия: ТЕОРИЯ ЯЗЫКА. СЕМИОТИКА. СЕМАНТИКА

ISSN 2313-2299 (Print); ISSN 2411-1236 (Online)
4 выпуска в год (ежеквартально).
http://journals.rudn.ru/semiotics-semantics
Языки: русский, английский.
Индексируется Белом списке РЦНИ, РИНЦ Научной электронной библиотеки, Scopus, RSCI, DOAJ, Google Scholar.
Включен в каталог периодических изданий (Ulrich's Periodicals Directory: http://www.ulrichsweb.com).

Цель и тематика

Журнал *«Вестник Российского университета дружбы народов. Серия: Теория языка. Семиотика. Семантика»* издается с 2010 г. и является периодическим рецензируемым научным изданием, входит в Перечень рецензируемых журналов ВАК РФ по специальности 5.9.8. Теоретическая, прикладная и сравнительно-сопоставительная лингвистика (филологические науки). Журнал ориентирован на широкую международную аудиторию.

Журнал углубляет и разрабатывает вопросы общей теории языка; теории речевой деятельности; семиотических характеристик знаковых систем, единиц языка разных уровней и текста; семиотики и поэтики художественных текстов; функциональной семантики лексических и грамматических единиц. Журнал предлагает вниманию комплексное и сопоставительное исследование категорий и единиц языка. Общие цели и задачи издания включают интегральную характеристику символической и коммуникативной парадигм филологического и гуманитарного знания. В прикладном аспекте разрабатываются методология и комплексные методики теоретических исследований языка и социума, системной лингвистики; осуществляется языковое моделирование. В журнале также есть разделы «Рецензии», «Научные обзоры», «Научная хроника». Издание адресовано филологам и другим представителям гуманитарного знания. Его авторы — доктора и кандидаты филологических наук, аспиранты, докторанты России, ближнего и дальнего зарубежья.

Правила оформления статей и другая информация о журнале размещена на сайте: http://journals.rudn.ru/semiotics-semantics/about

Каждая статья рецензируется анонимно двумя экспертами. Редакционная коллегия принимает решение о публикации с учетом мнения рецензентов.

Электронный адрес: semioticj@rudn.ru

Литературный редактор *К.В. Зенкин*
Редактор англоязычных текстов *О.Н. Колышева*
Компьютерная верстка: *Н.А. Ясько*

Адрес редакции:
Россия, 115419, Москва, ул. Орджоникидзе, д. 3
Тел.: +7 (495) 955-07-16; e-mail: publishing@rudn.ru

Почтовый адрес редакции:
Россия, 117198, Москва, ул. Миклухо-Маклая, д. 10, корп. 2
Тел.: +7 (495) 434-20-12; e-mail: semioticj@rudn.ru

Подписано в печать 15.06.2026. Выход в свет 30.06.2026. Формат 70×108/16.
Бумага офсетная. Печать офсетная. Гарнитура «Times New Roman».
Усл. печ. л. 40,78. Тираж 500 экз. Заказ № 560. Цена свободная.
Федеральное государственное автономное образовательное учреждение высшего образования
«Российский университет дружбы народов имени Патриса Лумумбы» (РУДН)
Россия, 117198, Москва, ул. Миклухо-Маклая, д. 6

Отпечатано в типографии ИПК РУДН
Россия, 115419, Москва, ул. Орджоникидзе, д. 3,
Тел. +7 (495) 955-08-61; e-mail: publishing@rudn.ru

© Российский университет дружбы народов, 2026

RUDN JOURNAL OF LANGUAGE STUDIES, SEMIOTICS AND SEMANTICS

2026 VOLUME 17 No. 2

Special Theme of the Issue:
ARTIFICIAL INTELLIGENCE IN APPLIED LINGUISTICS: CURRENT CHALLENGES AND PROSPECTS

Guest Editors **Andrey V. Dzhunkovskiy & Kristina P. Zhikulina**

DOI: 10.22363/2313-2299-2026-17-2

http://journals.rudn.ru/semiotics-semantics

Founded in 2010

Founder: Peoples' Friendship University of Russia named after Patrice Lumumba

EDITOR-IN-CHIEF
Vladimir N. Denisenko
RUDN University, Moscow, Russian Federation
E-mail: denisenko-vn@rudn.ru

DEPUTY EDITOR
Natalia V. Novospasskaya
RUDN University, Moscow, Russian Federation
E-mail: novospasskaya-nv@rudn.ru

EXECUTIVE SECRETARY
Olesya V. Lazareva
RUDN University, Moscow, Russian Federation
E-mail: lazareva-ov@rudn.ru

EDITORIAL BOARD

Klaudia Bednárová-Gibová, University of Prešov, Prešov, Slovak Republic
Nikolay N. Boldyrev, Tambov State University named after G.R. Derzhavin, Tambov, Russian Federation
Tatyana E. Vladimirova, RUDN University, Moscow, Russian Federation
Makhanbet Dzhusupov, Uzbek State World Languages University, Tashkent, Republic of Uzbekistan
Yulia A. Evgrafova, Moscow Power Engineering Institute, Moscow, Russian Federation
Evgeny E. Ivanov, Mogilev State A. Kuleshov University, Mogilev, Republic of Belarus
Vladimir I. Karasik, Pushkin State Russian Language Institute, Moscow, Russian Federation
Valentina A. Maslova, Vitebsk State University n.a. P.M. Masherov, Vitebsk, Republic of Belarus
Roberto V. Monforte Dupret, University of Basque Country, Vitoria-Gasteiz, Spain
Marina L. Novikova, RUDN University, Moscow, Russian Federation
Tatiana V. Poplavskaya, Minsk State Linguistic University, Minsk, Republic of Belarus
Vladimir P. Sinyachkin, RUDN University, Moscow, Russian Federation
Zifa K. Temirgazina, Pavlodar State Pedagogical University, Pavlodar, Republic of Kazakhstan
Yunsheng Ju, Heilongjiang University, Harbin, People's Republic of China

RUDN JOURNAL OF LANGUAGE STUDIES, SEMIOTICS AND SEMANTICS
Published by the Peoples' Friendship University of Russia named after Patrice Lumumba (RUDN University), Moscow, Russian Federation

ISSN 2313-2299 (Print); ISSN 2411-1236 (Online)
Publication frequency: quarterly.
http://journals.rudn.ru/semiotics-semantics
Languages: Russian, English.
Indexing: White List RCSI, Russian Index of Science Citation, Scopus, RSCI, RSCI, DOAJ, Ulrich's Periodicals Directory.

Aims and Scope

Journal of Language Studies, Semiotics and Semantics deepens and develops aspects of general theory of language; theories of speech activity; semiotic characteristics of sign systems, language units of different levels and text; semiotics and poetics of literary texts; functional semantics of lexical and grammatical units. The journal draws a reader's attention to a comprehensive and comparative study of the language categories and units. The general goals and objectives of the journal include an integral characteristic of symbolic and communicative paradigms of philological and humanitarian knowledge. The articles cover methodology and complex methods of theoretical studies of language and society, develop system linguistics and carry out language modeling. The journal introduces sections Reviews and Events. The journal is addressed to philologists and other representatives of the humanities. Our author are Russian and international scientists and postgraduate students.

Submission requirements and stylesheet guidelines are available online: http://journals.rudn.ru/semiotics-semantics/about

Each article is being reviewed anonymously (peer-reviewing) by two experts. The editorial board makes up a final decision on publication referring to the opinion of the reviewers.

Access to full-text files of issues and articles on the journal website is open to all users.

E-mail: semioticj@rudn.ru

Review Editor ***K.V. Zenkin***
English Text Editor ***O.N. Kolysheva***
Computer Design ***N.A. Yasko***

Address of the Editorial Board:
3 Ordzhonikidze st., Moscow, 115419, Russian Federation
Ph. +7 (495) 955-07-16; e-mail: publishing@rudn.ru

Postal Address of the Editorial Board:
10 Miklukho-Maklaya st., bldg. 2, Moscow, 117198, Russian Federation
Ph. +7 (495) 434-20-12; e-mail: semioticj@rudn.ru

Printing run 500 copies. Open price
Peoples' Friendship University of Russia named after Patrice Lumumba
6 Miklukho-Maklaya st., Moscow, 117198, Russian Federation

Printed at RUDN Publishing House:
3 Ordzhonikidze st., Moscow, 115419, Russian Federation
Ph. +7 (495) 952-08-61; e-mail: publishing@rudn.ru

© RUDN University, 2026

СОДЕРЖАНИЕ

АКТУАЛЬНЫЕ ВЫЗОВЫ ИСКУССТВЕННОГО ИНТЕЛЛЕКТА В ПРИКЛАДНОЙ ЛИНГВИСТИКЕ

От редакции …….. 391

Dzhunkovskiy A.V. (*Moscow, Russian Federation*). Forensic Linguistic Research of Actor-Based AI-Generated Speech in Audio-Video Games Post-Release (Юрислингвистическое исследование сгенерированной ИИ речи на базе голосов актеров озвучки в аудио-видеоиграх) …… 392

Ахренова Н.А. (*Москва, Российская Федерация*). Алгоритмический поворот в интернет-дискурсе …….. 401

Комалова Л.Р. (*Москва, Российская Федерация*). В поисках алгоритма определения текста, сгенерированного нейросетевой моделью, с опорой на его языковые особенности ……... 427

Максименко О.И., Беляков М.В. (*Москва, Российская Федерация*). Промт-инжиниринг как лингвистическая проблема …….. 442

Харламов А.А. (*Москва, Российская Федерация*). TextAnalyst и GPT: модель текста (модель текстового описания сущего) и семантика текста …… 459

Евграфова Ю.А. (*Москва, Российская Федерация*). Генеративные нейросети как инструмент транспозиции печатного текста в экранную форму на примере работ нейроиллюстратора DJ БлокNote …… 474

Новикова М.Л., Новиков Ф.Н. (Москва, *Российская Федерация*). Произведения больших языковых моделей как симулякр смысла. Семантическая энтропия и особенности новейших алгоритмов обнаружения использования искусственного интеллекта …… 496

Kiose M.I., Khulkhachieva Zh.S., Barmin A.V. (*Moscow, Russian Federation*). Exploring Cross-Cultural Knowledge Transfer: Ai-Integrated Semantic Framework (Изучение кросс-культурного трансфера знания: семантическое решение с встроенным ии-компонентом) ……... 517

Ваничкина А.С., Изюмская-Капитонова В.В. (*Москва, Российская Федерация*). Моделирование демографически обусловленной речи в больших языковых моделях: построение синтетического респондента для лингвистических исследований ……... 534

Жикулина К.П. (*Москва, Российская Федерация*). Токен как алгоритмическая «тень» морфемы …… 557

Popova M.V. (*Moscow, Russian Federation*). Transformation-Based Generation of English Language Exercises: A Deterministic and CEFR-Guided Approach (Генерация упражнений по английскому языку на основе текстовой трансформации: детерминированная модель с учётом уровней CEFR) …… 569

Zabolotskikh A.V., Sokolova E.E., Tavberidze D.V. (*Moscow, Russian Federation*). Machine vs Human: Large Language Model Errors as a Means of Creating Stylistically Appropriate Text (Человек или машина: ошибки больших языковых моделей как средство создания стилистически уместного текста) …… 585

ДИСКУРСИВНЫЕ ИССЛЕДОВАНИЯ

Malyuga E.N., Malkina M.P. (*Moscow, Russian Federation*). Mitigation Strategies in South African English Corporate Communication (Стратегии митигации в корпоративной коммуникации на материале южноафриканского варианта английского языка) …… 605

Калинин О.И. (*Москва, Российская Федерация*). Взаимосвязь метафоры и оценки в репрезентации образа России в англоязычном медиадискурсе Великобритании и Индии …….. 625

Ларионова М.В., Мосина С.К. (*Москва, Российская Федерация*). Национально-культурные коды в мексиканском политическом дискурсе: инаугурация сквозь призму лингвосемиотики ……... 648

Хромов С.С., Распопов А.А. (*Москва, Российская Федерация*). Эволюция поликодового текста в российском медиадискурсе: семантика, семиотика, институциональность …… 668

Topping K. (*Dundee, Scotland, UK*), Egorova L.A., Nikashina N.V., Nagornova E.V. (*Moscow, Russian Federation*). Professional and Popular Science: Mapping Linguistic Overlap and Divergence (Язык академической науки и научно-популярной литературы: сходства и различия) …… 690

ФУНКЦИОНАЛЬНАЯ ГРАММАТИКА И СЕМАНТИКА

Куликова Э.Г., Барабаш В.В. (*Москва, Российская Федерация*). Динамика морфологических норм русского языка XXI века: элиминация системных лакун vs грамматические архаизмы …… 712

Червоный А.М. (*Таганрог, Российская Федерация*), Арутюнова Ж.М., Нелюбова Н.Ю. (*Москва, Российская Федерация*). Функциональный и коллокационный потенциал существительного coup в современном французском языке ……. 728

Денисенко В.Н., Сафаралиева Л.А., Недопекина Е.М. (*Москва, Российская Федерация*), Чжан Кэ (*Чанчунь, Китайская Народная Республика*). Концепт DIEU ‘БОГ’ как ценностная доминанта французской лингвокультуры ……. 743

Мыскин С.В., Вашунина И.В. (*Москва, Российская Федерация*), Кабакова М.П., Муса К.М. (*Almaty, the Republic of Kazakhstan*). Модель этнокультурной семантической универсалии в эдусемиотике на примере казахского языка …….. 757

Сербин В.А. (*Москва, Российская Федерация*). Эволюция политической медиариторики во Вьетнаме 1945–2025 гг.: лингвистический анализ ……... 778

ЛИНГВИСТИКА ХУДОЖЕСТВЕННОГО ТЕКСТА

Ganina V.V., Kartashkova F.I. (*Ivanovo, Russian Federation*). Non-Verbal Behavior as a Marker of Altered States of Consciousness in a Literary Piece (Невербальное поведение как признак измененных состояний сознания в литературном произведении) …….. 796

Анисимов В.Е., Анисимова Е.Д., Савастенко Р.А. (*Москва, Российская Федерация*). Трансформации прецедентных феноменов в дискурсе современной сетевой литературы ……. 813

Ismailova Sh.A., Bayanbaeva Zh.A., Daribaev S.D. (*Almaty, the Republic of Kazakhstan*). The Birth of a Myth in the Literary Works by S. Sanbaev: the Archetype of Cosmogony (Рождение мифа в творчестве С. Санбаева: архетипика космогонии) …… 836

CONTENTS

ARTIFICIAL INTELLIGENCE IN APPLIED LINGUISTICS: CURRENT CHALLENGES AND PROSPECTS

Editorial Note 391

Dzhunkovskiy A.V. (*Moscow, Russian Federation*). Forensic Linguistic Research of Actor-Based AI-Generated Speech in Audio-Video Games Post-Release 392

Akhrenova N.A. (*Moscow, Russian Federation*). The Algorithmic Turn in Internet Discourse 401

Komalova L.R. (*Moscow, Russian Federation*). Towards an Algorithm for Identifying Text Generated by Neural Network Models through Linguistic Cues 427

Maksimenko O.I., Belyakov M.V. (*Moscow, Russian Federation*). Promt-Engineering as a Linguistic Problem 442

Kharlamov A.A. (*Moscow, Russian Federation*). Text Analyst and GPT: Text Model (Text Description Model) and Text Semantics 459

Yevgrafova Yu.A. (*Moscow, Russian Federation*). Generative Neural Networks as a Tool for Transposing Printed Text into Screen Format: Case Study of Neuro-Illustrator DJ БлокNote 474

Novikova M.L., Novikov Ph.N. (*Moscow, Russian Federation*). Work of Large Language Models as Simulations of Meaning. Semantic Entropy and Features of the Latest AI Detection Algorithms 496

Kiose M.I., Khulkhachieva Zh.S., Barmin A.V. (*Moscow, Russian Federation*). Exploring Cross-Cultural Knowledge Transfer: Ai-Integrated Semantic Framework 517

Vanichkina A.S., Izyumskaya-Kapitonova V.V. (*Moscow, Russian Federation*). Modeling Demographically Conditioned Language in Large Language Models: Towards a Synthetic Respondent for Linguistic Research 534

Zhikulina K.P. (*Moscow, Russian Federation*). Token as an Algorithmic 'Shadow' of a Morpheme 557

Popova M.V. (*Moscow, Russian Federation*). Transformation-Based Generation of English Language Exercises: A Deterministic and CEFR-Guided Approach 569

Zabolotskikh A.V., Sokolova E.E., Tavberidze D.V. (*Moscow, Russian Federation*). Machine vs Human: Large Language Model Errors as a Means of Creating Stylistically Appropriate Text 585

DISCOURSE STUDIES

Malyuga E.N., Malkina M.P. (*Moscow, Russian Federation*). Mitigation Strategies in South African English Corporate Communication 605

Kalinin O.I. (*Moscow, Russian Federation*). The Interrelation of Metaphor and Evaluation in the Representation of Russia's Image in English-Language Media Discourse in Great Britain and India ……. 625

Larionova M.V., Mosina S.K. (*Moscow, Russian Federation*). National-Cultural Codes in Mexican Political Discourse: Inauguration through the Prism of Linguosemiotics …… 648

Khromov S.S., Raspopov A.A. (*Moscow, Russian Federation*). Evolution of the Polycode Text in Russian Media Discourse: Semantics, Semiotics, and Institutional Dimensions …… 668

Topping K. (*Dundee, Scotland, UK*), Egorova L.A., Nikashina N.V., Nagornova E.V. (*Moscow, Russian Federation*). Professional and Popular Science: Mapping Linguistic Overlap and Divergence …….. 690

FUNCTIONAL GRAMMAR AND SEMANTICS

Kulikova E.G., Barabash, V.V. (*Moscow, Russian Federation*). Dynamics of Morphological Norms of the Russian Language in the XXI Century: Elimination of Systemic Lacunae vs Grammatical Archaisms ……. 712

Chervony A.M. (*Taganrog, Russian Federation*), Arutyunova Zh.M., Neliubova N.Yu. (*Moscow, Russian Federation*). Functional and Collocational Potential of the Noun Coup in Modern French …… 728

Denisenko V.N., Safaralieva, L.A., Nedopekina E.M. (*Moscow, Russian Federation*), Zhang Ke (*Changchun, The People's Republic of China*). The concept of DIEU 'GOD' as the Dominant Value of French Linguoculture …… 743

Myskin S.V., Vashunina I.V. (*Moscow, Russian Federation*), Kabakova M.P., Musa K.M. (*Almaty, the Republic of Kazakhstan*). The Model of Ethnocultural Semantic Universality in Edusemiotics on the Example of the Kazakh Language ……. 757

Serbin V.A. (*Moscow, Russian Federation*). Evolution of Political Media Rhetoric in Vietnam 1945–2025: A Linguistic Analysis …….. 778

LITERARY TEXT LINGUISTICS

Ganina V.V., Kartashkova F.I. (*Ivanovo, Russian Federation*). Non-Verbal Behavior as a Marker of Altered States of Consciousness in a Literary Piece …….. 796

Anisimov V.E., Anisimova E.D., Savastenko R.A. (*Moscow, Russian Federation*). Transformation Strategies of Precedent Phenomena in the Discourse of Modern Web-Literature …… 813

Ismailova Sh.A., Bayanbaeva Zh.A., Daribaev S.D. (*Almaty, the Republic of Kazakhstan*). The Birth of a Myth in the Literary Works by S. Sanbaev: the Archetype of Cosmogony ……. 836

RUDN Journal of Language Studies, Semiotics and Semantics
2313-2299 (Print), ISSN 2411-1236 (Online)
Вестник Российского университета дружбы народов.
Серия: ТЕОРИЯ ЯЗЫКА. СЕМИОТИКА. СЕМАНТИКА
2026 Vol. 17 No. 2 391
http://journals.rudn.ru/semiotics-semantics

АКТУАЛЬНЫЕ ВЫЗОВЫ ИСКУССТВЕННОГО ИНТЕЛЛЕКТА В ПРИКЛАДНОЙ ЛИНГВИСТИКЕ

ARTIFICIAL INTELLIGENCE IN APPLIED LINGUISTICS: CURRENT CHALLENGES AND PROSPECTS

DOI: 10.22363/2313-2299-2026-17-2-391
EDN: LFSKXO

Редакционная статья / Editorial

Специальный выпуск «Актуальные вызовы искусственного интеллекта в прикладной лингвистике» содержит ряд междисциплинарных лингвистических исследований, ставших возможными из-за новейших веяний современного мира: в начале второй четверти XXI в. возникают вопросы о том, как повлияет искусственный интеллект на когнитивные способности человека, какие профессии он создаст и уничтожит, что ожидает ИИ-индустрию в ближайшем будущем.

Мы с коллегами рады представить уважаемым читателям исследования, которые объединяет проблематика взаимодействия лингвистики с технологиями искусственного интеллекта: авторы обращаются к анализу промпта как нового типа управляющего текста, требующего переосмысления категорий семантики и прагматики; предлагают детерминированную модель генерации учебных упражнений на основе текстовой трансформации с калибровкой по уровням CEFR; рассматривают LLM как инструмент обучения академическому письму; выявляют системные искажения морфемной структуры при BPE-токенизации русскоязычных текстов и вводят типологию таких искажений; сопоставляют интерпретируемую семантическую сеть с непрозрачными механизмами генеративных трансформеров; разрабатывают юрислингвистические методы атрибуции синтезированной речи в аудио-видеоиграх и стеганографические механизмы защиты голосовых данных; предлагают трехуровневую систему верификации текстов, сгенерированных нейросетевыми моделями; создают семантический фреймворк с ИИ-компонентом для выявления прямого и непрямого трансфера знания из языка-донора в язык-реципиент; а также оценивают потенциал синтетических респондентов для моделирования гендерно маркированной вариативности на материале дневниковых текстов. Все работы отличаются методологической строгостью, опорой на корпусные данные и вносят вклад в формирование теоретических оснований компьютерной и цифровой лингвистики.

Редакционная коллегия

RUDN Journal of Language Studies, Semiotics and Semantics
2313-2299 (Print), ISSN 2411-1236 (Online)
Вестник Российского университета дружбы народов.
Серия: ТЕОРИЯ ЯЗЫКА. СЕМИОТИКА. СЕМАНТИКА

2026 Vol. 17 No. 2 392–400
http://journals.rudn.ru/semiotics-semantics

DOI: 10.22363/2313-2299-2026-17-2-392-400
EDN: LGKZBF

Research Article / Научная статья

Forensic Linguistic Research of Actor-Based AI-Generated Speech in Audio-Video Games Post-Release

Andrey V. Dzhunkovskiy

Moscow State Linguistic University, *Moscow, Russian Federation*
✉ Vetinari01@gmail.com

Abstract. Since 2023, the audio-video game industry has witnessed the active implementation of generative artificial intelligence for voice-acting game characters. In some cases, AI completely replaces the role of voice actors; in others, actors and developers enter into an agreement whereby the developer reserves the right to use neural networks and the actor's voice samples to synthesize new lines as needed. With proper algorithm training and a sufficient training dataset, the results of speech synthesis based on an actor's voice can be indistinguishable from natural speech except through deep phonoscopic analysis. Another trend in the modern audio-video game industry is noteworthy, which, in combination with the previous one, creates the foundation of the problem we aim to research: "games-as-a-service" (GaaS), a phenomenon in which a game is not a static, finished product, but evolves over time, receiving updates with new content. Under these conditions, a situation arises where it is possible to synthesize speech indistinguishable at first glance from human speech by a specific individual who has given legal consent for this, yet lacks control over which phrases will be publicly voiced. The research methodology relies on works in the field of specialized forensic linguistics of text and speech. The aim of the research is to formulate an algorithm for the linguistic analysis of such utterances, which has been accomplished.

Keywords: AI-speech synthesis, forensic linguistics, speaker attribution, steganography

Conflict of interests: The author states that there were no conflicts of interest.

Article history: Received: 17.01.2026; Accepted: 27.02.2026.

For citation: Dzhunkovskiy, A.V. (2026). Forensic Linguistic Research of Actor-Based AI-Generated Speech in Audio-Video Games Post-Release. *RUDN Journal of Language Studies, Semiotics and Semantics, 17*(2), 392–400. https://doi.org/10.22363/2313-2299-2026-17-2-392-400 EDN LGKZBF

© Dzhunkovskiy A.V., 2026

This work is licensed under a Creative Commons Attribution-NonCommercial 4.0 International License
https://creativecommons.org/licenses/by-nc/4.0/legalcode

Юрислингвистическое исследование сгенерированной ИИ речи на базе голосов актеров озвучки в аудио-видеоиграх

А.В. Джунковский

Московский государственный лингвистический университет, *Москва, Российская Федерация*
✉Vetinari01@gmail.com

Аннотация. Актуальность обусловлена активным внедрением разработчиками развлекательного программного обеспечения генеративного искусственного интеллекта для озвучки персонажей в индустрии аудио-видеоигр. В некоторых случаях ИИ полностью заменяет роль актеров озвучки, в других случаях актеры и разработчики заключают соглашение, согласно которому разработчик оставляет за собой право использовать нейронные сети и образцы голоса актера для синтезирования новых реплик по мере необходимости. При должном обучении алгоритма и достаточной обучающей выборке результаты синтеза речи на основе голоса актера могут быть неотличимы от естественного речепроизводства без глубокого фоноскопического анализа. Дополнительно необходимо отметить другой тренд современной индустрии аудио-видеоигры, в совокупности с предыдущим, формирующим проблему нашего исследования: «игры-как-сервис» (англ. GaaS; Games as a Service), явление, при котором игра не статична, а развивается с течением времени, получая обновления с новым контентом. В этих условиях возникает ситуация, при которой возможен синтез неотличимой на первый взгляд человеческой речи конкретным человеком, который дал на это юридическое согласие, однако не имеет контроля над тем, какие фразы будут озвучены его голосом в публичном пространстве. Методология исследования опирается на юрислингвистических работах в сфере специализированного изучения текста и речи. Цель исследования заключается в том, чтобы сформировать алгоритм лингвистического анализа таких высказываний, что было осуществлено.

Ключевые слова: ИИ-синтез речи, юрислингвистика, атрибуция говорящего, стеганография

Конфликт интересов: автор заявил об отсутствии конфликтов интересов.

История статьи: дата поступления: 17.01.2026; дата приема в печать: 27.02.2026.

Для цитирования: *Dzhunkovskiy A.V.* Forensic Linguistic Research of Actor-Based AI-Generated Speech in Audio-Video Games Post-Release // Вестник Российского университета дружбы народов. Серия: Теория языка. Семиотика. Семантика. 2026. Т. 17. № 2. С. 392–400. https://doi.org/10.22363/2313-2299-2026-17-2-392-400 EDN LGKZBF

Introduction

The audio-video game industry (colloquially simply referred to as “The game industry”) is not a new phenomenon, but it is becoming increasingly impossible to ignore its global impact on lifestyle, economy, politics and society. Despite this, modern Russian methods of Forensic Linguistics are not yet reworked to be able to analyze the products of this industry: audio-video games.

For the purposes of our research, we shall specify that from a linguistic viewpoint we consider an audio-video game to be an interactive, multimodal, polycode, creolized intertextual text with variable content. The last attribute requires additional explanation: audio-video game content is seldom linear — certain parts may not be encountered, may be skipped by choice or due to

circumstances of a particular playthrough, hence shaping a unique perception of the product for different users. This is further complicated by the ability of developers to add new content via updates.

The totality of what we described creates a unique challenge when it comes to analyzing linguistic contents of audio-video games. Studies indicate that, as of 2025, the average age of a gamer is 41 years old, with a gender balance of 51% to 49% slightly favoring males[1], and a total of approximately 3.57 billion consumers of such works worldwide. Concurrently, certain estimates placed the value of the industry at $189 billion USD in 2025[2]. In the same year, the music industry's value amounted to $29.6 billion USD[3], while the film industry's reached $33.5 billion USD[4]. Thus, the audio-video game industry is approximately three times more valuable than the film and music industries combined in 2025. This state of affairs is not novel: the assertion that the audio-video game industry is valued higher than the film and music industries combined held true each year from 2007 onward[5]. We have not discovered indicators that state that this trend will change.

Narratives, themes, and dialogue in audio-video games frequently address the topics of philosophy, politics, religion, social issues, ideology. Moreover, these works achieve instantaneous global reach through a sophisticated system of digital distribution. Under these circumstances, audio-video games can only be classified as powerful tools with the capacity for linguistic, ideological, political, psychological, and other forms of influence, as well as prospective repositories for propaganda and extremism.

We have to touch upon two important audio-video game industry trends. The first one has to do with the advent of AI-generated voice acting in audio-video games. It is possible to use fully AI-generated voices to reduce game development costs, as well as decrease the time it takes to produce a new game. Another scenario has become not uncommon since at least 2023: game developers hire real voice actors and negotiate in their contract the continued use of their voice with AI-generated voice lines for future content. This future content is required for the "games-as-a-service" (GaaS) distribution model, which entails continual addition of new content to already existing games via updates. Hence, a modern audio-video game is often not static, but a changing body of text, parts of which may be added, reworked or even removed over time. Analysis of audio-video games is

[1] Power of Play: 2025 Global Video Games Report. In: *Entertainment Software Association*. URL: https://www.theesa.com/resources/the-global-power-of-play-report (accessed: 17.02.2026)

[2] Global games market to hit $189 billion in 2025 as growth shifts to console. In: *Newzoo*. URL: https://newzoo.com/resources/blog/global-games-market-to-hit-189-billion-in-2025 (accessed: 17.02.2026).

[3] How Big Is the Music Industry in 2025? In: *Badenstock*. URL: https://artists.badenstock.com/how-big-is-the-music-industry-in-2025/ (accessed: 17.02.2026).

[4] Global Box Office Estimated At $33.5B; Anime Explodes In Up-And-Down Year: Studio Report Cards. In: Deadline. URL: https://deadline.com/2026/01/global-box-office-2025-report-hollywood-studio-rankings-1236660512 (accessed: 17.02.2026).

[5] Newzoo Global Games Market Report 2019. In: *Newzoo*. URL: https://newzoo.com/resources/trend-reports/newzoo-global-games-market-report-2019-light-version (accessed: 17.02.2026).

only possible via version control where researchers reach conclusions regarding a particular static temporal slice of an evolving body of text under the same name.

Theoretical Framework

Forensic linguistic research of speech activity artifacts represents a branch of applied linguistics within which a wide range of issues concerning semantic and stylistic text analysis is addressed for the purposes of regulating legal relations in society and establishing linguistic facts. In a broader sense, forensic linguistic research also encompasses similar analytical activities conducted by a linguist that are not regulated by Russian Federal Law № 73-FZ "On State Forensic Expert Activity in the Russian Federation."[6] The objective of studying speech activity products in this context is not the pursuit of new knowledge but rather the evaluation of linguistic facts pertaining to a speech product.

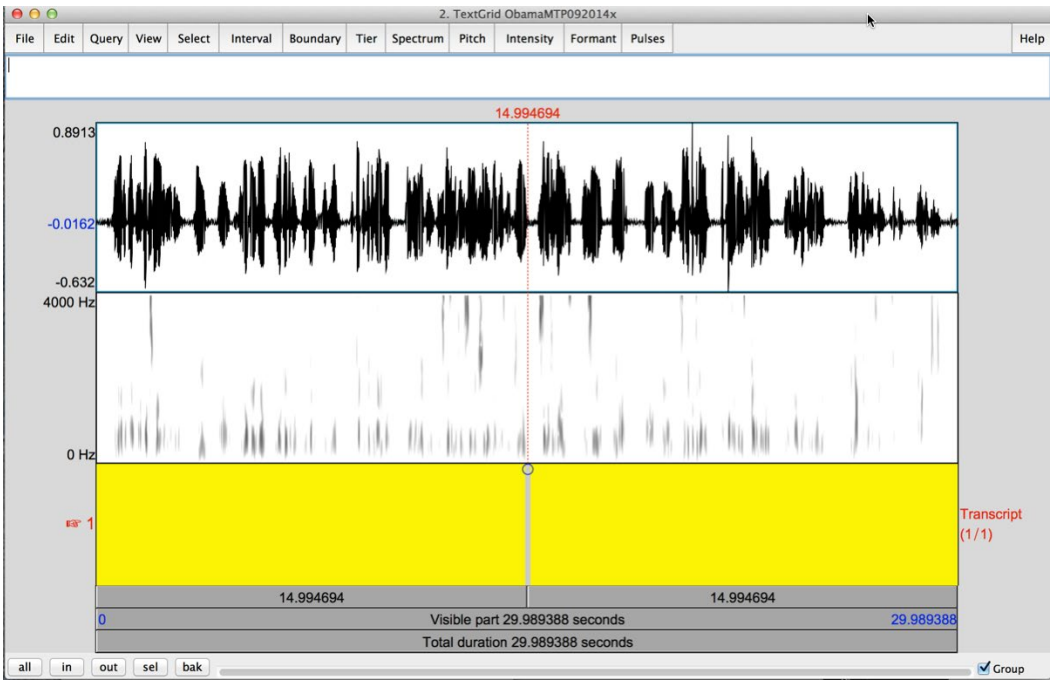
At present, such research in Russia is frequently conducted on trademarks, media texts, correspondence, social media messages, literary works, music, cinematography, as well as audio and video recordings of speech activity [1]. The latter four types of research typically involve the engagement of an acoustics expert. Meanwhile, the substantive analysis of the linguistic object is still performed by a linguist. Thus, a gap in the list of speech activity products amenable to analysis is revealed: interactive audiovisual works, commonly referred to as computer or video games (or, by some researchers, us included, audio-videogames) [2–4].

Let us propose a thought experiment. In the process of creating an audio-video game, a voice actor reaches an agreement with the developers where they not only voice one or multiple characters, but also grant the rights to use their voice samples to train a generative AI speech synthesis model. Their direct engagement with the development process ceases, but as the game evolves in a "games as a service" framework, the developers create new voice lines nigh-indistinguishable from real speech production without complicated phonoscopic analysis. These new AI-synthesized voice lines include content that is illegal to be voiced openly in certain jurisdictions. In an even more complex scenario, the trained AI speech synthesis model may leak, allowing anyone the ability to synthesize any utterance desired using the voice of a real person.

In contemporary forensic linguistics, experts use specialized software to determine if recorded speech has been tampered with, or simply to determine if recorded speech is indeed uttered by the same individual (using comparison to a controlled sample guaranteed to contain speech by the same individual). While more sophisticated software is used professionally, we may easily demonstrate what such a process looks like using open software (Figure).

[6] Russian Federal Law № 73-FZ "On State Forensic Expert Activity in the Russian Federation." In: *Garant*. URL: https://base.garant.ru/12123142/ (accessed: 17.02.2026). (In Russ.).

Above we may see a spectrogram of human speech. For the purposes of this research, it is not necessary to explain in depth the machinations on display. Spectral analysis allows forensic linguists to uncover a host of data that can be compared to control samples, thus allowing one to produce data-based conclusions that are later used in court.



PRAAT audio analysis software
Source: compiled by Andrey V. Dzhunkovskiy.

The tool pictured above is PRAAT[7], developed by Paul Boersma and David Weenink at the University of Amsterdam. It stands as a cornerstone in the field of audio signal analysis, offering a sophisticated, open-source platform for phonetic and speech research. Primarily designed for the visualization, manipulation, and quantitative assessment of acoustic signals, PRAAT enables researchers to extract precise measurements from waveforms, spectrograms, and derived parameters such as formant frequencies, pitch contours, and intensity envelopes. Its scripting language further extends its utility, automating complex analyses across large corpora of audio data.

At its core, PRAAT's audio signal analysis workflow begins with signal loading and preprocessing. Users import digitized audio files typically sampled at rates exceeding 16 kHz to capture human speech harmonics and apply filtering techniques, such as bandpass or pre-emphasis, to mitigate noise and enhance spectral clarity. The software's autocorrelation-based pitch detection algorithm,

[7] PRAAT: doing phonetics by computer. In: *Praat*. URL: https://www.fon.hum.uva.nl/praat/ (accessed: 17.02.2026).

which computes periodicity via time-domain correlations, yields fundamental frequency (F0) tracks with sub-millisecond resolution, crucial for prosodic studies. For instance, in analyzing intonation patterns, researchers select voiced segments and invoke the "To Pitch" command, specifying autocorrelation parameters like time step (e.g., 0.001 s) and frequency range (75–600 Hz for adult speech), generating pitch objects amenable to statistical modeling.

Spectral analysis constitutes another pillar of PRAAT's capabilities. Fast Fourier Transform (FFT) spectrograms, configurable with window lengths from 0.005 to 0.1 s and dynamic ranges up to 70 dB, reveal formant structures, resonant peaks in the vocal tract transfer function. Burg's linear predictive coding (LPC) method, implemented in the "To Formant (burg)" procedure, estimates formant frequencies (F1–F5) by minimizing prediction error over frames, with order typically set to 2.5 times the number of formants desired. This facilitates vowel quality mapping in acoustic spaces like the F1–F2 plane, where normalization schemes (e.g., Lobanov or Watt-Fabricius) account for speaker-intrinsic variation.

Intensity and duration metrics further enrich PRAAT's analytical repertoire. Root-mean-square (RMS) energy computations, smoothed via Gaussian kernels, delineate stress and phrasing boundaries, while tiered TextGrid annotations synchronize phonetic transcriptions with temporal landmarks. Advanced users leverage scripts for batch processing, such as extracting F0 means, standard deviations, and formant dispersions from aligned corpora, enabling multivariate analyses via exported CSV matrices interfaced with R or Python.

Empirical validation underscores PRAAT's reliability; inter-rater agreement studies report intraclass correlation coefficients exceeding 0.95 for formant tracking across diverse accents. Limitations persist, notably in handling non-stationary signals or severe distortions, where extensions like deep learning plugins (e.g., via PRAAT's plugin architecture) are emerging. Nonetheless, PRAAT remains indispensable for rigorous audio signal analysis, bridging raw acoustics to perceptual linguistics in experimental paradigms.

When AI-synthesized speech is concerned, this becomes considerably more complicated and contentious, especially in cases when the AI is trained on a body of work created by the same individual, like in the thought experiment presented above.

It is evident that the voice actor is no longer the author of these new utterances, but they may face a reality where they would need to prove in a court of law or public opinion that the utterance is a product of AI, and not themselves. Parallels may be easily drawn with the AI deepfake problem.

Discussion and Results

The issue presented above is currently becoming more important and prominent. While use of generative AI in audio-video games is often criticized by users and creatives whose jobs are ostensibly at risk, the audio-video games industry is heavily invested in promoting the use of generative AI. Prominent

examples include the free 2013 audio-video game "The Finals"[8] [5] and the 2025 audio-video game "ARC Raiders"[9]. These are not small, relatively unknown entities. Data aggregators estimate ownership of the former at 18 million and the latter at 9.38 million users worldwide[10].

A question naturally arises: with such outreach and impact of audio-video games, what steps may be taken to protect voice actors from potential misuse of generative AI voice synthesis models trained with their voices? A potential solution may be found in the field of linguistic steganography.

Linguistic steganography is a branch of steganography dealing with how verbal, non-verbal, and paraverbal variables may be used to embed secret information into text and voice recordings. Steganography in and of itself deals with the issue of covertly relaying secret messages without third parties becoming cognizant of confidential information being "in transit". Unlike cryptography, the goal of steganography is obfuscation of the very fact that additional information is being relayed between parties [6–9].

One prominent use of linguistic steganography is watermarking which helps protect copyright and determine sources of potential leaks. Watermarking entails embedding hidden markers that allow one to claim ownership of a digital entity.

There is research that details the algorithm of how additional information may be imbedded into audio. This additional information is added via altering the audio signal by altering wavelengths at peaks. This is not perceptible to the human ear and is only detectable on a spectrogram via forensic linguistic analysis [10].

We posit the following: the selfsame steganography algorithm for relaying additional info by altering the audio signal may be used to watermark AI-synthesized speech that was trained with the use of real human voice actors. Such a procedure would provide a powerful, subtle and cost-effective protection for voice actors, and for developers if the embedding of this additional info is built into the AI model itself in case it becomes available to a malicious third party.

Final Remarks

While linguistic forensic analysis of audio-video games is in its infancy, it is both possible and necessary to predict and preemptively solve potential problems that may arise from not having the sufficient methods and tools that allow experts to conduct such analysis when it comes to problematic content in other forms of media: deepfakes in news broadcasts, fabricated celebrity endorsements, courtroom audio tampering. Without robust methodologies, audio-video games risk becoming vectors for disinformation: AI-generated propaganda disseminated via in-game voiceover, indistinguishable from voice actor-created or even developer-sanctioned

[8] The Finals Criticised by Actors and Devs for Using AI-Generated Voice Overs. In: *IGN*. URL: https://www.ign.com/articles/the-finals-criticised-by-actors-and-devs-for-using-ai-generated-voice-overs (accessed: 17.02.2026).

[9] Arc Raiders and the Ethical Use of Generative AI. In*: AI and Games*. URL: https://www.aiandgames.com/p/arc-raiders-and-the-ethical-use-of (accessed: 17.02.2026).

[10] Steam Database. URL: https://steamdb.info/ (accessed: 17.02.2026).

content, eroding player trust and inciting real-world unrest. Preemptive development of tools—such as real-time spectrographic watermarking can avert these pitfalls by embedding verifiable markers during production, enabling post-hoc authentication even in live-service environments where content changes over time.

The problem is additionally compounded by the rise of generative AI that creates unique and novel challenges. Generative AI exacerbates these challenges exponentially. Models like those powering dialogues or procedural narratives produce hyper-realistic, contextually adaptive speech that evades conventional phonetics: prosody mimics human variability and synthesis artifacts are minimized through training. This leads us to "uncanny valley" forensic linguistics: subtle indicators like unnatural micro-pauses or spectral inconsistencies that elude human perception without additional expert analysis. In "games as a service" ecosystems, speaker attribution fractures. Proactive countermeasures may entail integrating forensic linguistics into audio-video game development pipelines from inception. Developers could mandate "audio passports" containing metadata schemas logging synthesis parameters, training data provenance, and edit histories.

References

1. Kukushkina, O.V., Safopova, U.A., &. Sekerazh, T.P. (2011). *Theoretical and Methodical Basis for Forensic Psychological and Linguistic Text Research for Extremism*. Moscow: Russian Ministry of Justice publ. (In Russ.).
2. Kamaruddin, N., et al. (2018). Review of Text Watermarking: Theory, Methods, and Applications. *IEEE Access*, 6, 8011–8028. https://doi.org/10.1109/ACCESS.2018.2796585
3. Kuleshov, S.V., Zaytseva, A.A., Aksenov, A.Yu. (2024). Analysis of Statistical Characteristics of Artificially Generated Texts. *Journal of Instrument Engineering*, 67(11), 958–968. (In Russ.). https://doi.org/10.17586/0021-3454-2024-67-11-958-968 EDN: HIYQAL
4. Ogorelkov, I.V., & Makhaeva, E.V. (2025). Artificial Neural Network Text Generation as a New Type of Masking the Authorship. *Virtual Communication and Social Networks*, *4*(2), 163–171. (In Russ.). https://doi.org/10.21603/2782-4799-2025-4-2-163-171 EDN: HIYQAL
5. Prokhorov, A.I., & Asadchaya, K.V. (2023). Tools for Determining the Text Generated Using a Neural Network. In: *Nauchny Vektor*, proceeding (Vol. 9, pp. 250–253). Rostov-on-Don. (In Russ.). EDN: MGTGIE
6. Potapova, R.K., & Dzhunkovskiy, A.V. (2021). Trigger-Containers in Linguistic Steganography. *Russian Linguistic Bulletin*, *26*(2), 19–21. (In Russ.). https://doi.org/10.18454/RULB.2021.26.2.10 EDN: DAKDOO
7. Potapova, R.K. (2015). *Speech: Communication, Information, Cybernetics*. Moscow: URSS. (In Russ.).
8. Potapova, R. (2002). Linguistic Knowledge and New Technologies. *Acoustic Physics*, *48*(4), 486–493. (In Russ.). EDN: NXXVUV
9. Potapova, R., & Dzhunkovskiy, A. (2020). Preliminary Investigation of Potential Steganographic Container Localization. In: *Lecture Notes in Artificial Intelligence* (12335, pp. 389–398). Springer. https://doi.org/10.1007/978-3-030-60276-5_38 EDN: RLZNGS
10. Salagaj, M.O. (2011). *Prosodic Means of Information Protection* (*Experimental Phonetic Inquiry in the Field of Steganography*) [PhD Thesis]. Moscow: MSLU publ. (In Russ). EDN: QFOJLJ

Список литературы

1. *Кукушкина О.В., Сафопова Ю.А., Секераж Т.П.* Теоретические и методические основы судебной психолого-лингвистической экспертизы текстов по делам, связанным с противодействием экстремизму. М. : Министерство Юстиции РФ, 2011.
2. *Kamaruddin N. et al.* Review of Text Watermarking: Theory, Methods, and Applications. // IEEE Access. 2018. Vol. 6. P. 8011–8028. https://doi.org/10.1109/ACCESS.2018.2796585
3. *Кулешов С.В., Зайцева А.А., Аксенов А.Ю.* Анализ статистических характеристик искусственно сгенерированных текстов // Известия вузов. Приборостроение. 2024. Т. 67. № 11. С. 958–968. https://doi.org/10.17586/0021-3454-2024-67-11-958-968 EDN: HIYQAL
4. *Огорелков И.В., Махаева Е.В.* Искусственная нейросетевая генерация текста как новый вид маскировки письменной речи автора // Виртуальная коммуникация и социальные сети. 2025. Т. 4. № 2. С. 163–171. https://doi.org/10.21603/2782-4799-2025-4-2-163-171 EDN: HIYQAL
5. *Прохоров А.И., Асадчая К.В.* Инструментальные средства определения текста, сгенерированного при помощи нейросети // Научный вектор: сб. научных трудов / ред. Е.Н. Макаренко. Т. 9. Ростов-на-Дону : Ростовский государственный экономический университет «РИНХ», 2023. С. 250–253. EDN: MGTGIE
6. *Потапова Р.К., Джунковский А.В.* Понятие «Триггер-контейнер» в лингвистической стеганографии // Russian Linguistic Bulletin. 2021. № 26 (2). С. 19–21. https://doi.org/10.18454/RULB.2021.26.2.10 EDN: DAKDOO
7. *Потапова Р. К.* Речь: коммуникация, информация, кибернетика. М. : Изд-во УРСС, 2010.
8. *Потапова Р.К.* Лингвистические знания и новые технологии // Акустический журнал. 2002. Т. 48. № 4. С. 552–559. EDN: NXXVUV
9. *Potapova R., Dzhunkovskiy A.* Preliminary Investigation of Potential Steganographic Container Localization // Lecture Notes in Computer Science. № 12335. Springer, 2020. P. 389–398. EDN: RLZNGS
10. *Салагай М.О.* Просодические средства защиты смысловой информации (экспериментально-фонетическое исследование в области стеганографии) : дис. … канд. филол. наук. М. : МГЛУ, 2011. EDN: QFOJLJ

Information about the author:
Andrey V. Dzhunkovskiy, PhD in Philology, Applied and Experimental Linguistics Department Chair, Lead Researcher at Experimental Phonetics and Forensic Linguistics Laboratory, Moscow State Linguistic University (38 build 1 Ostozhenka str., Moscow, Russian Federation, 119034); *Research interests*: corpus linguistics, experimental linguistics, pragmalinguistics, forensic linguistics; *e-mail*: Vetinari01@gmail.com
ORCID: 0000-0002-3761-8010, Scopus ID: 57210805837; ResearcherID: J-6081-2017, SPIN-code: 8541-6226, AuthorID: 930082.

Сведения об авторе:
Джунковский Андрей Владимирович, кандидат филологических наук, заведующий кафедрой прикладной и экспериментальной лингвистики, ведущий научный сотрудник экспериментально-фонетической лаборатории криминалистики по речеведению, Московский государственный лингвистический университет (119034, Российская Федерация, г. Москва, ул. Остоженка, д. 38 с. 1); *сфера научных интересов*: корпусная лингвистика, экспериментальная лингвистика, лингвопрагматика, юрислингвистика; *e-mail*: Vetinari01@gmail.com
ORCID: 0000-0002-3761-8010, Scopus ID: 57210805837; ResearcherID: J-6081-2017, SPIN-код: 8541-6226, AuthorID: 930082.

RUDN Journal of Language Studies, Semiotics and Semantics
2313-2299 (Print), ISSN 2411-1236 (Online)
2026 Vol. 17 No. 2 401–426
http://journals.rudn.ru/semiotics-semantics
Вестник Российского университета дружбы народов.
Серия: ТЕОРИЯ ЯЗЫКА. СЕМИОТИКА. СЕМАНТИКА

DOI: 10.22363/2313-2299-2026-17-2-401-426
EDN: LHQFQK

Научная статья / Research Article

Алгоритмический поворот в интернет-дискурсе

Н.А. Ахренова

Российский университет дружбы народов, *Москва, Российская Федерация*
nakhrenova@mail.ru

Аннотация. Интернет-дискурс, традиционно изучавшийся в парадигме компьютерно-опосредованной коммуникации как гибридное образование, сочетающее черты устной и письменной речи, вступает в качественно новую фазу своего развития. Публичный запуск больших языковых моделей (GPT-4 и аналоги) в 2022–2023 гг. знаменует собой алгоритмический поворот: генеративный искусственный интеллект перестает быть нейтральным каналом передачи сообщений и превращается в активного соучастника коммуникации, способного самостоятельно продуцировать контент, вступать в диалог и влиять на языковые практики миллионов пользователей. Лингвистическая теория, однако до сих пор оперирует понятийным аппаратом, ориентированным на антропоцентрическую модель дискурса, что не позволяет адекватно описывать новую коммуникативную реальность, в которой границы между автором, инструментом и сообщением размываются. Цель. Выявить и систематизировать фундаментальные трансформации интернет-дискурса в эпоху генеративного искусственного интеллекта, предложив теоретическую рамку для их анализа в контексте цифровой антрополингвистики. Материалы и методы. Исследование выполнено в междисциплинарном дизайне, сочетающем методы корпусной лингвистики и критического дискурс-анализа. Материалом послужили три оригинальных контрастных корпуса общим объемом 1,5 млн словоупотреблений на русском и английском языках: (1) корпус человеческого общения (H2H), репрезентирующий «доалгоритмическую» эпоху (до 2021 г.); (2) корпус промптов (H2A), фиксирующий новую жанровую реальность общения человека с алгоритмом; (3) корпус текстов, сгенерированных искусственным интеллектом (A2H), отражающий специфику алгоритмического дискурса. Результаты. Количественный анализ впервые на репрезентативном материале подтвердил гипотезу о гомогенизации дискурса: корпус A2H демонстрирует снижение лексического разнообразия на 37,6 % и синтаксического разнообразия на 45,3 % по сравнению с корпусом человеческого общения. Выявлена жанровая специфика промпта как нового типа дискурсивной практики — директивно-эксплораторного жанра, характеризующегося императивностью, параметризацией запроса и метаязыковой рефлексией пользователя. Обнаружен феномен алгоритмической нормы — устойчивого воспроизводства клишированных конструкций книжного регистра («важно отметить», «таким образом»), формирующих усредненный, стилистически гомогенизированный регистр. Впервые зафиксированы и описаны гибридные дискурсивные акты — коммуникативные единицы, возникающие в итеративном взаимодействии человека и алгоритма, авторство которых не может быть однозначно атрибутировано ни одному из участников. Заключение. В ответ на вызовы новой

© Ахренова Н.А., 2026
This work is licensed under a Creative Commons Attribution-NonCommercial 4.0 International License
https://creativecommons.org/licenses/by-nc/4.0/legalcode

эпохи интернет-дискурса вводятся теоретические концепты, позволяющие описывать трансформированную коммуникативную реальность: параметр «степень алгоритмической опосредованности» (DAM) и понятие «гибридного дискурсивного акта». Сопоставительный анализ русскоязычной и англоязычной частей корпусов демонстрирует универсальность выявленных трансформаций, что позволяет говорить не о локальных языковых изменениях, а об общецивилизационном сдвиге — алгоритмическом повороте, требующем пересмотра фундаментальных категорий лингвистики (авторство, текст, дискурс, языковая личность).
Ключевые слова: искусственный интеллект, алгоритмическая опосредованность, гибридная агентность, цифровая коммуникация, цифровая антрополингвистика

Конфликт интересов: автор заявил об отсутствии конфликтов интересов.

История статьи: дата поступления: 20.01.2026; дата приема в печать: 10.02.2026.

Для цитирования: *Ахренова Н.А.* Алгоритмический поворот в интернет-дискурсе // Вестник Российского университета дружбы народов. Серия: Теория языка. Семиотика. Семантика. 2026. Т. 17. № 2. С. 401–426. https://doi.org/10.22363/2313-2299-2026-17-2-401-426 EDN LHQFQK

The Algorithmic Turn in Internet Discourse

Natalia A. Akhrenova

RUDN University, *Moscow, Russian Federation*
✉nakhrenova@mail.ru

Abstract. Background. Internet discourse, traditionally studied within the paradigm of computer-mediated communication as a hybrid formation combining features of oral and written speech, is entering a qualitatively new phase of its development. The public launch of large language models (GPT-4 and equivalents) in 2022–2023 marks an algorithmic turn: generative artificial intelligence ceases to be a neutral channel for message transmission and transforms into an active participant in communication, capable of independently producing content, engaging in dialogue, and influencing the linguistic practices of millions of users. However, linguistic theory still operates within a conceptual apparatus oriented toward an anthropocentric model of discourse, which does not allow us to adequately describe the new communicative reality in which the boundaries between author, tool, and message are becoming blurred. Objective. To identify and systematize the fundamental transformations of internet discourse in the era of generative artificial intelligence, proposing a theoretical framework for their analysis within the context of digital anthropolinguistics. Materials and Methods. The study employs an interdisciplinary design combining methods of corpus linguistics and critical discourse analysis. The material comprises three original contrasting corpora with a total volume of 1.5 million word usages in Russian and English: (1) a corpus of human-to-human communication (H2H), representing the “pre-algorithmic” era (up to 2021); (2) a corpus of prompts (H2A), capturing the new genre reality of human-algorithm communication; and (3) a corpus of AI-generated texts (A2H), reflecting the specificity of algorithmic discourse. Results. Quantitative analysis, for the first time on representative material, confirmed the hypothesis of discourse homogenization: the A2H corpus demonstrates a 37.6% reduction in lexical diversity and a 45.3% reduction in syntactic diversity compared to the human communication corpus. The genre specificity of the prompt as a new type of discursive practice has been identified — a directive-exploratory genre characterized by imperativeness, query parameterization, and the user’s metalinguistic reflection. The phenomenon of the algorithmic norm has been discovered — the stable reproduction of clichéd bookish register constructions (“it is important to note,” “thus”), forming an averaged, stylistically homogenized register. Hybrid discursive acts have been recorded and described for the first time — communicative units arising from the iterative interaction between

human and algorithm, the authorship of which cannot be unambiguously attributed to either participant. Conclusion. In response to the challenges of the new era of internet discourse, theoretical concepts are introduced that allow for the description of the transformed communicative reality: the parameter “Degree of Algorithmic Mediation” (DAM) and the concept of the “hybrid discursive act.” A comparative analysis of the Russian and English parts of the corpora demonstrates the universality of the identified transformations, suggesting not local linguistic changes but a civilizational shift — an algorithmic turn that requires a reconsideration of the fundamental categories of linguistics (authorship, text, discourse, linguistic persona).

Keywords: artificial intelligence, algorithmic mediation, hybrid agency, digital communication, digital anthropolinguistics

Conflict of interests: The author stated that there were no conflicts of interest.

Article history: Received: 20.01.2026; Accepted: 10.02.2026.

For citation: Akhrenova, N.A. (2026). The Algorithmic Turn in Internet Discourse. *RUDN Journal of Language Studies, Semiotics and Semantics, 17*(2), 401–426. (In Russ.). https://doi.org/10.22363/2313-2299-2026-17-2-401-426 EDN LHQFQK

Введение

Компьютерно-опосредованная коммуникация (КОК) на протяжении последних трех десятилетий остается одной из магистральных тем лингвистических исследований. От ранних работ по анализу чатов и форумов до современных исследований социальных сетей дискурс-анализ успешно описывал специфику виртуального общения, опираясь на такие категории, как гибридность устной и письменной речи, гипертекстуальность, анонимность и игровая природа коммуникации [1; 2]. Классические модели КОК строились вокруг идеи о том, что компьютер является нейтральным каналом (медиумом) передачи сообщения между людьми, а единственным агентом, продуцирующим смыслы, остается человек.

Однако публичный запуск больших языковых моделей (Large Language Models, LLM), таких как GPT-4 и его аналоги, в 2022–2023 гг. ознаменовал собой наступление качественно новой эры. Генеративный искусственный интеллект (ГИИ) перестал быть инструментом пассивной обработки данных и превратился в активного соучастника коммуникации. Пользователи не просто общаются через компьютер, но все чаще вступают в диалог с алгоритмом или используют его для продуцирования контента, который затем циркулирует в человеческой среде. Возникает ситуация, которую уместно обозначить термином «алгоритмическая опосредованность» (algorithmic mediation): алгоритм выступает не просто посредником, а соавтором или даже генератором сообщения, что радикально меняет структуру коммуникативного акта.

Как справедливо отмечает Т.А. Литвинова с соавторами, наступление эры LLM требует переосмысления самого понятия «письмо» как вида деятельности [3], а мы можем добавить, что также переосмысливается понятие «текст». Лингвистическая теория сегодня сталкивается с вызовом, для которого в классическом инструментарии пока нет адекватного ответа: как описывать дискурс, в котором авторство текстов распределено между человеком

и машиной; как квалифицировать речевой акт, где интенция и ее реализация разделены между биологическим и алгоритмическим интеллектом.

Осмысление указанных процессов происходит в рамках формирующейся междисциплинарной области — цифровой антрополингвистики, которая исследует фундаментальные изменения, вносимые цифровой средой в понимание языка, культуры и их неразрывной связи.

Традиционная антрополингвистика фокусировалась на языке в его культурном контексте, изучая, как речь-действие формирует социальные отношения, ритуалы и мировоззрение в рамках конкретного сообщества. Ее методологической базой было полевое исследование — включенное наблюдение за речевыми практиками территориально-локализованных групп. Однако цифровая эпоха создала принципиально новую среду бытования языка — гибридное коммуникативное пространство, стирающее границы между офлайн и онлайн, локальным и глобальным, устным и письменным. Это требует пересмотра базовых категорий дисциплины [Морозова, 2023].

Переход от классической антрополингвистики к цифровой знаменует собой смену парадигмы по нескольким ключевым направлениям:

1. Трансформация сообщества. Классическая антрополингвистика работала с территориально-локализованными сообществами, чьи языковые практики были неразрывно связаны с физическим местом, общими ритуалами и прямым («живым») межличностным взаимодействием («речевой общиной»). Цифровая антрополингвистика изучает сетевые сообщества (digital tribes), собранные вокруг общих интересов, практик или аффектов, а не их географического расположения. Их связывает не общая территория, а общая платформа, хештег или игровая вселенная. Членство в таком сообществе добровольно, перформативно и часто множественно.

2. Трансформация метода. Классический полевой этнографический метод — включенное наблюдение — трансформируется в нетнографию (netnography). Полем становится цифровая среда (форум, чат, социальная сеть, игровой мир). Исследователь должен осваивать роль «цифрового аборигена», что подразумевает понимание специфического жаргона, мемов, норм поведения сообщества, разработку этики работы с цифровым следом и работу с Большими данными (Big Data), используя методы корпусной лингвистики.

3. Трансформация единицы анализа. Единицей анализа становится поликодовый, полимодальный дискурс. Сообщение создается одновременной работой вербального кода (текст), визуального кода (эмодзи, мем, гифка, аватар) и технического кода (гиперссылка, хештег, функция репоста, алгоритмическая подача). Наблюдается поликодирование текста — интеграция вербального и невербального кодов в единое семиотическое целое. Эмодзи становятся визуальным диалектом, интернациональным языком эмоций и действий. Мемы функционируют как культурные единицы, требующие от реципиента специфической культурной компетенции для декодирования.

4. Появление нового актора. В классической модели взаимодействие было горизонтальным: человек — человек в рамках культурного контекста. В цифровую эпоху появляется новый, нечеловеческий актор, опосредующий и формирующий это взаимодействие — алгоритм. Алгоритм выступает как цензор и модератор, определяя приемлемость высказываний и устанавливая новую цифровую норму. Алгоритм как куратор контента (рекомендательные системы) формирует «лингвокультурный пузырь» пользователя, активно влияя на его языковой выбор и культурный кругозор. Наконец, алгоритм как генератор языка через большие языковые модели создает тексты, которые участвуют в коммуникации наравне с человеческими, стирая границы между «естественной» и «искусственной» речью.

Появление алгоритма как активного участника коммуникации ставит перед теорией дискурса проблему агентности (agency). В классической коммуникативной схеме Р. Якобсона «адресант — сообщение — адресат» адресант всегда мыслился как человек — носитель интенции, эмоций и культурного контекста. В случае с генеративным ИИ возникает феномен, который мы предлагаем обозначить как гибридная агентность.

Гибридная агентность характеризуется распределением интенциональности и ответственности за высказывание между человеком и алгоритмом. Человек задает цель, формулирует запрос (промпт), оценивает результат и осуществляет финальную редактуру. Алгоритм обеспечивает лингвистическую реализацию цели, оперируя статистическими паттернами, извлеченными из гигантских корпусов человеческой речи. Результат такого сотрудничества не может быть однозначно приписан ни одному из участников. Это коррелирует с концепцией гибридного авторства, разрабатываемой в западной лингвистике и литературоведении, где атрибуция текста становится сложной многослойной задачей [4; 5].

Ближайшей аналогией в доцифровую эпоху могло бы быть редактирование или соавторство, но эти модели предполагают, что все соавторы являются людьми и действуют в рамках единого культурного кода. Алгоритм же действует на основе иного — статистического, а не культурного — кода, что порождает феномен алгоритмической интертекстуальности: текст генерируется как статистически вероятностная комбинация элементов обучающей выборки, а не как осмысленная интертекстуальная отсылка.

Анализ научных публикаций западных и российских авторов последних лет позволяет выделить несколько устойчивых направлений в осмыслении роли ИИ в коммуникации.

Технологическое направление фокусируется на инженерных аспектах разработки и оптимизации LLM, а также на методах промпт-инжиниринга как способа «настройки» модели под конкретные задачи [6; 7]. Лингвистический анализ здесь вторичен и служит преимущественно прикладным целям повышения эффективности генерации.

Социальное направление исследует области применения генеративного ИИ: в образовании, медицине, журналистике, бизнесе. Работы этого направления фиксируют новые практики, но редко углубляются в лингвистическую природу порождаемых текстов.

Антропологическое направление рассматривает влияние ИИ на человека, его идентичность и когнитивные способности. Именно в этом русле возникают вопросы о трансформации языковой личности, о новых формах грамотности (AI literacy) и об этических аспектах делегирования речи алгоритмам [3].

Несмотря на растущий корпус исследований, сохраняется ряд существенных лакун: отсутствует комплексное лингвистическое описание дискурсивных практик, опосредованных генеративным ИИ, с опорой на репрезентативные корпусные данные; не разработана теоретическая модель, которая включала бы «алгоритмическую опосредованность» как системообразующий параметр дискурса; не введены понятия для описания гибридных (человеко-машинных) речевых актов.

В этой связи целью данной статьи является выявление и систематизация фундаментальных трансформаций интернет-дискурса, вызванных интеграцией генеративных нейросетей, а также будет предложена теоретическая рамка для их анализа в контексте цифровой антрополингвистики.

Для достижения поставленной цели нами были решены следующие задачи:

1) сформированы контрастные корпуса, репрезентирующие три типа коммуникации: человеческую (до эры массового ИИ), директивную (промпты) и алгоритмическую (тексты, сгенерированные ИИ);

2) gроведен количественный (корпусный) анализ лексических и синтаксических особенностей каждого типа дискурса;

3) выявлены качественные (дискурсивные) характеристики промпт-инженерии как нового речевого жанра;

4) описан феномен гибридного авторства и предложены способы его концептуализации;

5) введены параметр степень алгоритмической опосредованности (Degree of Algorithmic Mediation, DAM) и понятие гибридного дискурсивного акта.

Научная новизна исследования заключается в следующем: впервые предлагается комплексное сопоставление трех типов дискурса (человек — человек, человек — алгоритм, алгоритм — человек) на единой корпусной основе; вводится и операционализируется понятие алгоритмической интертекстуальности, описывающее специфику порождения текста LLM; разрабатывается таксономия коммуникативных стратегий в жанре промпта; предлагается расширение модели интернет-дискурса за счет параметра алгоритмической опосредованности.

Теоретическая значимость работы состоит в развитии категориального аппарата цифровой лингвистики и антрополингвистики в рамках цифровой

гуманитаристики. Введение понятий «гибридная агентность», «алгоритмический дискурс» и «гибридный дискурсивный акт» позволяет более точно описывать современную коммуникативную реальность, выходя за рамки антропоцентрической парадигмы.

Практическая значимость связана с возможностью применения результатов в разработке методик лингвистической экспертизы для атрибуции авторства, в преподавании курсов по цифровой грамотности (digital literacy), а также в оптимизации взаимодействия человека с генеративными нейросетями (промпт-инжиниринг).

Данное исследование призвано внести вклад в понимание того, как цифровая среда и алгоритмы меняют природу человеческого общения, культурной передачи и языковой идентичности. Цифровая антрополингвистика становится не просто новым направлением, а необходимостью, поскольку цифровая среда превратилась в основную среду бытования языка для миллиардов людей.

Материалы и методы исследования

В данном разделе представлено детальное описание эмпирической базы исследования, принципов формирования корпусов, а также методов количественного и качественного анализа, примененных для верификации выдвинутых теоретических положений. Методологический дизайн работы строится на сочетании традиций корпусной лингвистики, критического дискурс-анализа и нетнографического подхода, что позволяет обеспечить как статистическую надежность выводов, так и глубину интерпретации выявленных феноменов.

Принципы формирования корпусов. Для решения поставленных исследовательских задач был разработан дизайн контрастивного корпусного исследования, предполагающий сопоставление трех типов дискурса, различающихся по параметру участия человека и алгоритма в порождении текста. Основным критерием отбора материалов являлась возможность однозначной атрибуции источника текста (человек / алгоритм) и хронологическая привязка к периоду до и после массового распространения генеративных нейросетей.

При формировании корпуса учитывался принципиальный хронологический рубеж — конец 2022 г., когда произошел публичный запуск ChatGPT и начало массового внедрения больших языковых моделей в повседневную коммуникативную практику. Корпус человеческого общения (H2H) формировался из текстов, созданных до 2021 г. включительно, чтобы минимизировать возможное влияние ИИ-генераций на человеческую речь. Корпуса промптов (H2A) и ИИ-генераций (A2H) собирались из материалов 2023–2025 гг.

Материалы собраны на двух языках — русском и английском — в пропорции примерно 60% к 40% соответственно. Двуязычный состав корпусов позволяет верифицировать универсальность выявленных закономерностей и

отделить языковые особенности от общих дискурсивных тенденций, связанных с алгоритмической опосредованностью.

При сборе материалов соблюдались принципы исследовательской этики: использовались только публично доступные тексты, все персональные данные были анонимизированы, идентификаторы пользователей удалены. Для корпуса промптов использовались материалы из открытых репозиториев, где пользователи добровольно публикуют свои диалоги с ИИ в исследовательских целях.

Структура и объем корпусов. Итоговый объем сформированных материалов составил 1 545 070 словоупотреблений (токенов). Распределение по корпусам представлено в табл. 1.

Таблица 1. **Структура и объем корпусов исследования**
Table 1. **Structure and Volume of the Research Corpora**

Корпус / Corpora	**Обозначение / Designation**	**Описание / Description**	**Объем (токены) / Volume (tokens)**	**Количество текстов / Number of texts**
Корпус человеческой коммуникации / Human Communication Corpora	H2H (Human-to-Human)	Тексты, созданные людьми для общения с людьми до 2021 г. (форумы, блоги, личная переписка, комментарии в соцсетях)	512 847	3 427
Корпус промптов / Prompt Corpora	H2A (Human-to-AI)	Диалоги пользователей с ChatGPT, запросы к генеративным нейросетям	463 291	5 891
Корпус ИИ-генераций / AI Generation Corpora	A2H (AI-to-Human)	Тексты, сгенерированные ИИ и маркированные как таковые (ответы ChatGPT, статьи, созданные нейросетями)	568 932	2 143
ИТОГО / Total			**1 545 070**	**11 461**

Источник: исследование Н.А. Ахреновой.
Source: compiled by Natalia A. Akhrenova.

Корпус H2H (Human-to-Human). Корпус человеческого общения составлен из текстов, созданных в период 2015–2021 гг. Источниками послужили: открытые форумы и дискуссионные платформы (Reddit, Pikabu, специализированные форумы по различной тематике); посты и комментарии в блогах (LiveJournal, WordPress); публично доступные фрагменты личной переписки (из корпусов разговорной речи); тексты электронных писем (с согласия участников и в анонимизированной форме).

Критерии отбора: текст должен быть создан человеком, предназначен для восприятия другим человеком, не содержать признаков автоматической генерации. В корпус не включались тексты, прошедшие интенсивную автоматическую обработку (например, автоматически сгенерированные новости на сайтах информагентств).

Корпус H2A (Human-to-AI). Корпус промптов составлен из запросов пользователей к генеративным нейросетям, преимущественно к ChatGPT (версии 3.5 и 4). Источниками послужили: открытые репозитории диалогов ShareGPT и PromptBase (с согласия пользователей); публикации промптов в тематических сообществах (Reddit r / ChatGPT, Telegram-каналы по промпт-инжинирингу); собственный сбор материалов через краудсорсинг (участники добровольно предоставляли логи своих диалогов с ИИ в анонимизированной форме). В корпус включались как однократные запросы, так и многошаговые диалоговые цепочки, что позволяет анализировать итеративные стратегии взаимодействия.

Корпус A2H (AI-to-Human). Корпус текстов, сгенерированных искусственным интеллектом, составлен из ответов нейросетей на запросы пользователей, а также из текстов, опубликованных как созданные ИИ. Источниками послужили: ответы ChatGPT, сопряженные с запросами из корпуса H2A (для анализа диалогового взаимодействия); статьи и посты, явно маркированные авторами как «сгенерировано ИИ№ (в блогах, социальных сетях); тексты из специализированных датасетов ИИ-генераций, используемых для обучения детекторов ИИ-текста. В корпус включались только тексты, в отношении которых имеется достоверная информация об их алгоритмическом происхождении (метаданные, явные указания авторов, техническая документация).

Методы анализа. Исследование реализовано в рамках междисциплинарного подхода (mixed-methods design), сочетающего количественные методы корпусной лингвистики и качественные методы дискурсивного анализа.

Корпусно-статистические методы. Для обработки корпусных данных использовалось программное обеспечение AntConc (версия 4.2.0) и специализированные скрипты на языке Python (библиотеки NLTK, pandas, scipy). Применялись следующие методы:

Частотный анализ. Подсчет абсолютных и относительных частот словоформ и лемм в каждом корпусе с последующим сопоставлением. Частотность нормализована на 1 миллион словоупотреблений для обеспечения сопоставимости корпусов разного объема.

Анализ ключевых слов. Выявление лексических единиц, статистически значимо отличающихся по частоте употребления в одном корпусе по сравнению с другим. В качестве меры ключевости использовался показатель Log-likelihood (LL) с порогом статистической значимости $p < 0.01$ ($LL > 6.63$). Сравнение проводилось попарно: H2A против H2H, A2H против H2H, A2H против H2A.

Анализ коллокаций. Выявление статистически значимых сочетаний слов для ключевых терминов. Использовался метод Mutual Information (MI-score) с порогом MI > 3 для отбора значимых коллокатов. Диапазон контекста — ±4 слова от узлового слова.

Лексическое разнообразие. Расчет коэффициента тип-токен (Type-Token Ratio, TTR) для каждого корпуса. Для нивелирования влияния объема корпуса использовался стандартизированный показатель STTR (среднее TTR для сегментов по 1000 слов).

Синтаксический анализ. Автоматическое определение типов предложений (простые, сложносочиненные, сложноподчиненные) с использованием синтаксических парсеров (spaCy для английского языка, Stanza для русского). Подсчет средней длины предложения и распределения синтаксических конструкций.

Дискурсивный анализ. Для интерпретации количественных данных и выявления глубинных дискурсивных характеристик применялись методы качественного анализа: критический дискурс-анализ (КДА) в традициях Н. Фэркло и Р. Водак анализ фокусировался на выявлении: способов репрезентации участников коммуникации (человека, алгоритма); конструирования социальных отношений и идентичностей в тексте; идеологических импликаций выбора тех или иных языковых средств.

Жанровый анализ. Классификация промптов по коммуникативным стратегиям с выделением устойчивых типов и моделей. Анализ структурно-композиционных особенностей промпта как нового речевого жанра.

Анализ диалоговых структур. Исследование итеративных взаимодействий в цепочках «человек — ИИ — человек» с выделением паттернов гибридного авторства и редактуры.

Нетнографический подход. В соответствии с парадигмой цифровой антрополингвистики часть анализа реализована в рамках нетнографического подхода, предполагающего: погружение в цифровые сообщества пользователей ИИ (тематические форумы, чаты, группы) для понимания контекста и норм коммуникации; анализ метадискурса пользователей о взаимодействии с ИИ (обсуждения эффективных промптов, стратегий «общения» с алгоритмом); учет культурного контекста цифровых платформ как среды бытования изучаемых дискурсивных практик.

Процедура анализа. Исследование проводилось в несколько этапов.

Этап 1. Формирование корпусов (январь–март 2025 г.). Сбор, очистка, анонимизация и первичная обработка текстовых материалов. Токенизация, лемматизация, снятие омонимии.

Этап 2. Количественный анализ (апрель–июнь 2025 г.). Проведение частотного анализа, выявление ключевых слов, анализ коллокаций, подсчет лексического разнообразия и синтаксических характеристик. Статистическая обработка результатов.

Этап 3. Качественный анализ (июль–сентябрь 2025 г.). Интерпретация количественных данных, дискурсивный анализ репрезентативных примеров, типологизация стратегий промптинга, выделение гибридных дискурсивных актов.

Этап 4. Синтез результатов и теоретическое обобщение (октябрь–ноябрь 2025 г.). Построение теоретической модели, введение параметра алгоритмической опосредованности, формулировка выводов.

Ограничения исследования. При интерпретации результатов необходимо учитывать следующие ограничения.

1. Репрезентативность корпусов. Несмотря на значительный объем, корпуса не могут считаться полностью репрезентативными для всех типов интернет-коммуникации. Корпус промптов смещен в сторону англоязычных и русскоязычных пользователей, имеющих доступ к технологиям и готовых делиться логами диалогов.

2. Динамичность объекта. Генеративные нейросети быстро развиваются, и тексты, сгенерированные GPT 3.5, могут отличаться от текстов более новых версий. Исследование фиксирует состояние на 2023–2025 гг.

3. Проблема атрибуции. В корпус H2H могли попасть тексты, уже испытавшие влияние ИИ (например, прошедшие автоматическую проверку грамматики). Однако хронологические рамки (до 2021 г.) минимизируют этот риск, поскольку массовые ИИ-генерации отсутствовали.

4. Языковая специфика. Выявленные закономерности могут частично различаться для русского и английского языков, что требует дополнительного сравнительного анализа, выходящего за рамки данной статьи.

Воспроизводимость исследования. Для обеспечения возможности верификации результатов сторонними исследователями.

1. Методика формирования корпусов документирована и может быть воспроизведена.

2. Использованные программные инструменты являются общедоступными.

3. Статистические методы соответствуют стандартам корпусной лингвистики.

4. Примеры в аналитической части приводятся с сохранением орфографии и пунктуации оригиналов (с необходимой анонимизацией).

Результаты и анализ

Проведенный комплексный анализ трех контрастных корпусов (H2H, H2A, A2H) позволил выявить статистически значимые различия на лексическом, синтаксическом и дискурсивном уровнях. Учитывая двуязычный состав корпусов (русский и английский языки в соотношении 60:40), все количественные показатели рассчитывались отдельно для каждого языка с последующим сопоставлением для выявления универсальных закономерностей алго-

ритмического дискурса и специфичных особенностей языка. В данном разделе мы последовательно представляем результаты количественного корпусного анализа, подкрепленные качественной интерпретацией в рамках критического дискурс-анализа.

Лексико-статистический анализ корпусов. Первым этапом анализа стало вычисление базовых статистических параметров, характеризующих лексическое разнообразие и синтаксическую сложность текстов в каждом корпусе. Для обеспечения сопоставимости данных показатели рассчитывались отдельно для русскоязычной и англоязычной частей корпусов (табл. 2).

Таблица 2. **Общие статистические характеристики корпусов (с распределением по языкам)**
Table 2. **General Statistical Characteristics of Corpora (distributed by language)**

Параметр / Parameter	Язык / Language	Корпус H2H / Corpora H2H	Корпус H2A / Corpora H2A	Корпус A2H / Corpora A2H
Общий объем (токенов) / Total volume (tokens)	Русский	312 847	278 291	341 932
	Английский	200 000	185 000	227 000
	Всего	**512 847**	**463 291**	**568 932**
Количество типов (уникальных слов) / Number of types (unique words)	Русский	42 847	24 891	28 763
	Английский	25 585	16 396	18 800
Лексическое разнообразие (TTR) / Lexical diversity	Русский	0.137	0.089	0.084
	Английский	0.128	0.088	0.082
	Среднее	**0.133**	**0.089**	**0.083**
Средняя длина предложения (слов) / Average sentence length (words)	Русский	12.8	8.9	17.3
	Английский	16.2	10.8	20.4
	Среднее	**14.3**	**9.8**	**18.7**
Индекс удобочитаемости* / Readability index*	Русский	48.5	68.2	38.4
	Английский (Flesch-Kincaid)	58.2	72.5	46.8

Источник: исследование Н.А. Ахреновой.
Source: compiled by Natalia A. Akhrenova.

Примечание: для русскоязычных текстов использовался адаптированный индекс удобочитаемости (формула Оборневой), для англоязычных — стандартный индекс Flesch-Kincaid. Прямое сопоставление абсолютных значений между языками некорректно, однако динамика изменений внутри каждого языка показательна.

Интерпретация результатов. Лексическое разнообразие (Type-Token Ratio, TTR). Наиболее высокий показатель TTR демонстрирует корпус H2H в обоих языках (русский: 0.137, английский: 0.128), что отражает естествен-

ное разнообразие человеческой речи, богатство идиолектов и стилистическую вариативность. Несколько более высокий показатель в русскоязычной части может объясняться более развитой системой словоизменения и словообразования в русском языке, а также большей частотностью дискурсивных слов, увеличивающих лексическое разнообразие.

Корпус A2H демонстрирует наиболее низкий TTR в обоих языках (русский: 0.084, английский: 0.082), что в среднем на 37,6 % ниже, чем в человеческом общении. Это является прямым количественным подтверждением тезиса о стилистической гомогенизации алгоритмического дискурса. Примечательно, что снижение лексического разнообразия практически одинаково для обоих языков (русский: снижение на 38,7 %, английский: снижение на 35,9%), что указывает на универсальный характер данного феномена, определяемый архитектурой языковых моделей, а не спецификой конкретного языка.

Корпус промптов (H2A) также показывает сниженное лексическое разнообразие (русский: 0.089, английский: 0.088), что объясняется функциональной ограниченностью жанра: пользователи используют ограниченный набор клишированных конструкций для запросов (*напиши/write, объясни/explain, сделай/make, переведи/translate*).

Средняя длина предложения. В обоих языках наблюдается идентичная закономерность: промпты (H2A) характеризуются минимальной длиной предложения (русский: 8,9 слов, английский: 10,8 слов). Это соответствует коммуникативной стратегии «экономии усилий» в директивных жанрах. Бóльшая длина английских промптов может объясняться аналитическим строем английского языка, требующим большего количества служебных слов для выражения тех же смыслов.

Тексты A2H демонстрируют максимальную длину предложения в обоих языках (русский: 17,3 слова, английский: 20,4 слова), значительно превышающую средние показатели человеческой речи. Анализ показывает, что ИИ тяготеет к развернутым, синтаксически сложным конструкциям с придаточными предложениями, причастными и деепричастными оборотами (в русском) и герундиальными конструкциями (в английском), что создает эффект «академичности» даже в бытовых контекстах.

Индекс удобочитаемости. В обоих языках тексты A2H оказались наиболее сложными для восприятия (русский: 38.4, английский: 46.8 по соответствующим шкалам, где более низкие значения указывают на большую сложность). Парадоксальным образом, будучи призванным упрощать коммуникацию, генеративный ИИ в своем «базовом» режиме продуцирует тексты, требующие более высокой читательской компетенции, чем средняя человеческая речь. Промпты же максимально просты и доступны (русский: 68.2, английский: 72.5).

Частотный анализ и анализ ключевых слов. Для выявления содержательных доминант каждого корпуса был проведен анализ частотности и ключевых

слов с использованием меры Log-likelihood (порог значимости $p < 0.01$, $LL > 6.63$). Ввиду значительных различий в грамматическом строе русского и английского языков ключевые слова анализировались отдельно для каждого языка с последующим сопоставлением функционально-семантических классов (табл. 3, 4).

Таблица 3. **Топ-10 ключевых слов (с положительной ключевостью) для каждого корпуса (русскоязычная часть)**
Table 3. **Top 10 Keywords (with Positive Keyword Density) for Each Corpora (Russian-language part)**

Ранг / Rank	Корпус H2H (рус.) / H2H corpora (Russian)	Корпус H2A (рус.) / H2A corpora (Russian)	Корпус A2H (рус.) / A2H corpora (Russian)
1	*ну* (дискурсивная частица)	*напиши*	*важно отметить*
2	*вот*	*сделай*	*следует учитывать*
3	*типа*	*помоги*	*таким образом*
4	*короче*	*объясни*	*в контексте*
5	*привет*	*переведи*	*исследование показывает*
6	*ребята*	*кратко*	*данные свидетельствуют*
7	*кстати*	*подробно*	*с одной стороны*
8	*слушай*	*формат*	*репрезентативный*
9	*нормально*	*список*	*релевантный*
10	*жесть*	*пример*	*дискурсивный*

Источник: исследование Н.А. Ахреновой.
Source: compiled by Natalia A. Akhrenova.

Таблица 4. **Топ-10 ключевых слов (с положительной ключевостью) для каждого корпуса (англоязычная часть)**
Table 4. **Top 10 Keywords (with Positive Keyword Activity) for Each Corpora (English part)**

Ранг/ Rank	Корпус H2H (англ.) / H2H corpora (English)	Корпус H2A (англ.) / H2A corpora (English)	Корпус A2H (англ.) / A2H corpora (English)
1	*like* (discourse marker)	*write*	*it is important to note*
2	*just*	*create*	*it should be noted*
3	*guys*	*help*	*thus*
4	*honestly*	*explain*	*in the context of*
5	*literally*	*translate*	*research shows*
6	*anyway*	*briefly*	*data suggest*
7	*basically*	*detailed*	*on the one hand*
8	*hey*	*format*	*representative*
9	*awesome*	*list*	*relevant*
10	*literally*	*example*	*discursive*

Источник: исследование Н.А. Ахреновой.
Source: compiled by Natalia A. Akhrenova.

Сопоставительный анализ ключевых слов. Корпус H2H в обоих языках насыщен дискурсивными маркерами устной речи (*ну*/*like*, *вот*/*just*, *типа*/*basically*, *короче*/*anyway*), формами обращения (*ребята*/*guys*,

слушай/*hey*) и эмоционально-оценочной лексикой (*жесть*/*awesome*, *нормально*/*literally* в функции интенсификатора). Это подтверждает гибридную природу интернет-коммуникации, совмещающей черты письменной и устной речи, а также высокую степень экспрессивности и межличностной ориентированности.

Межъязыковые различия минимальны и касаются преимущественно грамматического оформления: в русском языке высока частотность частиц (ну, вот), выполняющих функции хезитации и акцентирования; в английском — высока частотность дискурсивного маркера *like* и интенсификатора *literally*, отражающих специфику разговорной речи.

Корпус H2A в обоих языках демонстрирует доминирование глаголов в императиве (*напиши/write, сделай/create, помоги/help, объясни/explain, переведи/translate*), что однозначно характеризует жанр промпта как директивный. Высокая частотность слов *кратко/briefly, подробно/detailed, формат/format, список/list, пример/example* указывает на параметрический характер запроса: пользователь не просто задает тему, а уточняет жанровые и структурные параметры ожидаемого ответа.

Сопоставление языков показывает, что англоязычные промпты несколько чаще используют вежливые формы (косвенные вопросы с *could you, would you*), однако в корпусном анализе ключевых слов они не попадают в топ-10, уступая прямым императивам. Русскоязычные промпты тяготеют к использованию инфинитивных конструкций (*написать, сделать*) наряду с императивом.

Корпус A2H в обоих языках ключевые слова представлены клишированными вводными конструкциями (*важно отметить/it is important to note, следует учитывать/it should be noted*), маркерами логической связности (*таким образом/thus, с одной стороны/on the one hand*) и абстрактной лексикой книжного регистра (репрезентативный/representative, релевантный/relevant, дискурсивный/discursive). Это формирует эффект научности или официальности, который становится маркером алгоритмической нормы.

Примечательно, что русскоязычные ИИ-тексты демонстрируют бóльшую склонность к безличным и пассивным конструкциям (*следует учитывать, важно отметить*), в то время как англоязычные чаще используют конструкции с формальным подлежащим *it*. Это соответствует грамматическим нормам языков, но в ИИ-генерациях данные конструкции гипертрофированно частотны.

Анализ коллокаций: семантические поля. Для углубленного анализа было проведено исследование коллокаций (статистически значимых соседей с MI > 3) для ключевых понятий в обоих языках.

Семантическое поле ЧЕЛОВЕК / HUMAN

В корпусе H2H лексемы, обозначающие человека, сочетаются с глаголами, обозначающими эмоциональное состояние и повседневные действия:

Русский: *я + чувствую, думаю, пошел, сделал*

Английский*: I + feel, think, went, did*

В корпусе A2H наблюдается принципиально иная картина. Лексема *человек/human* взаимодействует с абстрактными существительными и глаголами состояния:

Русский: *человек + фактор, роль, деятельность, является, представляет собой*

Английский: *human + factor, role, activity, is, represents*

Человек в алгоритмическом дискурсе предстает не как конкретный индивид с эмоциями и действиями, а как абстрактная категория, объект классификации и теоретизирования. Это свидетельствует о деантропоморфизации дискурса: даже когда речь идет о человеке, алгоритм описывает его в терминах абстрактных сущностей.

Семантическое поле КОММУНИКАЦИЯ / COMMUNICATION

В корпусе H2A глаголы говорения/письма демонстрируют специфическое распределение:

Русский: *сказать* практически отсутствует (частотность 0.3 на млн), его замещают *написать* (45.2 на млн), *сгенерировать* (12.8 на млн), *выдать* (8.3 на млн)

Английский: *say* имеет низкую частотность (1.2 на млн), доминируют *write* (52.3 на млн), *generate* (15.7 на млн), *produce* (10.2 на млн)

Это указывает на восприятие ИИ не как собеседника, а как инструмента продуцирования текста. Коммуникация с алгоритмом концептуализируется не как диалог (обмен репликами), а как техническая операция по производству контента.

В корпусе A2H глагол *общаться*/*communicate* встречается в сочетании с лексемами:

Русский: *эффективно, продуктивно, в рамках, в контексте*

Английский: *effectively, productively, within, in the context of*

Это отражает инструментальный и нормативный подход к коммуникации, характерный для алгоритмических систем: общение рассматривается как деятельность, подлежащая оптимизации и оценке по критериям эффективности.

Дискурсивный анализ промптов (H2A)

Качественный дискурсивный анализ корпуса промптов позволил выделить устойчивые коммуникативные стратегии пользователей, которые демонстрируют сходные паттерны в обоих языках.

Типология стратегий промптинга:

1. Стратегия делегирования (48 % выборки, равномерно распределена между языками). Пользователь передает ИИ полную ответственность за создание текста, формулируя максимально общий запрос.

(1) *Напиши сочинение на тему «Осень».* (рус.)
(2) *Write an essay on «Autumn».* (англ.)

Характеризуется минимальной параметризацией, высокой степенью неопределенности. Пользователь ожидает, что ИИ самостоятельно решает все жанрово-стилистические вопросы.

2. Стратегия параметризации (35 % выборки). Пользователь задает четкие параметры будущего текста: объем, стиль, целевую аудиторию, тональность.

(3) *Напиши пост для Instagram[*] о новом кафе. Длина — до 1000 знаков с пробелами. Тон — дружелюбный, молодежный. Используй 3–5 эмодзи. Добавь призыв к действию в конце.* (рус.)
(4) *Write an Instagram[*] post about a new cafe. Max 1000 characters with spaces. Tone: friendly, youthful. Use 3–5 emojis. Add a call to action at the end.* (англ.)

Эта стратегия наиболее интересна с лингвистической точки зрения, так как она требует от пользователя развития метаязыковой рефлексии: он должен эксплицировать те жанровые и стилистические нормы, которые в человеческой коммуникации усваиваются имплицитно. Примечательно, что англоязычные пользователи несколько чаще специфицируют тональность и аудиторию, в то время как русскоязычные — объем и формат.

3. Стратегия итеративного уточнения (17% выборки). Многошаговый диалог, где пользователь последовательно корректирует результат.

(5) *Пользователь*: *Напиши письмо с отказом от вакансии. (рус.)*
(6) *ИИ*: *[генерирует текст]*
(7) *Пользователь*: *Слишком формально. Сделай мягче, добавь благодарность за уделенное время и упомяни, что мне предложили другую позицию. (рус.)*
(8) *Пользователь*: *Write a job rejection letter. (англ.)*
(9) *ИИ*: *[генерирует текст]*
(10) *Пользователь*: *Too formal. Make it softer, add thanks for their time and mention I got another offer.* (англ.)

Эта стратегия моделирует ситуацию редактуры и соавторства, приближаясь к тому, что мы называем гибридным дискурсивным актом. В англоязычной части корпуса итеративные уточнения встречаются несколько чаще (19 % против 16 % в русской), что может быть связано с более ранним и массовым распространением ChatGPT в англоязычной среде.

Анализ гомогенизации стиля в корпусе A2H. Для верификации тезиса о стилистической гомогенизации был проведен сравнительный анализ вариативности синтаксических конструкций в русскоязычной и англоязычной частях корпуса (табл. 5).

[*] Принадлежит Meta, признана экстремисткой и запрещенной на территории РФ.

Таблица 5. **Вариативность синтаксических конструкций (с распределением по языкам)**
Table 5. **Variability of Syntactic Constructions (with distribution by languages)**

Параметр / Parameter	Язык / Language	Корпус H2H / H2H corpora	Корпус A2H / A2H corpora
Количество уникальных синтаксических паттернов (на 10 000 предл.) / Number of unique syntactic patterns (per 10,000 sentences)	Русский	356	192
	Английский	328	182
	Среднее	**342**	**187**
Доля простых предложений, % / The proportion of simple sentences, %	Русский	42	24
	Английский	34	20
	Среднее	**38**	**22**
Доля сложносочиненных предложений, % / The proportion of complex sentences, %	Русский	25	29
	Английский	29	33
	Среднее	**27**	**31**
Доля сложноподчиненных предложений, % / The proportion of complex sentences, %	Русский	33	47
	Английский	37	47
	Среднее	**35**	**47**

Источник: исследование Н.А. Ахреновой.
Source: compiled by Natalia A. Akhrenova.

Анализ синтаксической вариативности. Корпус A2H демонстрирует значительно более низкое разнообразие синтаксических паттернов в обоих языках (снижение на 46,1 % в русском и на 44,5 % в английском). Это подтверждает универсальный характер синтаксической гомогенизации как свойства алгоритмического дискурса (табл. 6).

Таблица 6. **Наиболее частотные синтаксические паттерны в корпусе A2H**
Table 6. **The Most Frequent Syntactic Patterns in the A2H Corpora**

Русский язык / Russian	Английский язык / English
[Вводная конструкция], что [придаточное изъяснительное]	[Introductory phrase] that [noun clause]
Если [условие], то [следствие]	If [condition], then [result]
С одной стороны, [тезис 1], с другой стороны, [тезис 2]	On the one hand, [thesis 1], on the other hand, [thesis 2]
Важно отметить, что [утверждение]	It is important to note that [statement]
Таким образом, [вывод]	Thus, [conclusion]

Источник: исследование Н.А. Ахреновой.
Source: compiled by Natalia A. Akhrenova.

При этом наблюдается устойчивое смещение в сторону сложноподчиненных конструкций с придаточными определительными и изъяснительными. В русскоязычной части A2H доля сложноподчиненных предложений возрастает с 33 до 47 % (прирост 42 %), в англоязычной — с 37 до 47 % (прирост 27 %). Меньший прирост в английском объясняется исходно более высокой долей сложных предложений в человеческой английской речи, связанной с грамматическим строем языка.

Качественный анализ показывает, что ответы ИИ на различные стимулы (от бытовых до научных) демонстрируют удивительное стилистическое единообразие. Разница между «разговорным» и «официальным» режимами часто сводится к замене нескольких лексем при сохранении общей синтаксической матрицы. Например, в обоих языках ИИ устойчиво воспроизводит паттерн «*вводная конструкция* + *тезис* + *аргументация* + *вывод*» независимо от запрашиваемого жанра.

Сопоставительный анализ гибридных дискурсивных актов в русском и английском языках. На пересечении корпусов H2A и A2H (полные диалоговые цепочки) нами были выделены случаи, квалифицируемые как гибридные дискурсивные акты. Сопоставительный анализ русскоязычных и англоязычных примеров позволяет выявить как универсальные характеристики феномена, так и культурно-языковые особенности.

Критерии выделения гибридного дискурсивного акта: 1) наличие минимум двух итераций взаимодействия «человек-ИИ»; 2) наличие следов редактуры человеческого текста ИИ или следов адаптации ИИ-текста человеком; 3) неоднозначность атрибуции финального высказывания; 4) интеграция результата в человеческую коммуникацию.

Пример 1 (русскоязычный, адаптировано из корпуса):

(11) *Пользователь*: *Напиши тост на юбилей начальнику. Он строгий, любит точность, работает в IT.*
(12) *ИИ*: «*Дорогой Иван Иванович! В этот знаменательный день хочется пожелать вам четкости кода и ясности мысли, стабильности в работе и радости в жизни. Пусть ваш код всегда компилируется без ошибок, а баги обходят стороной!*»
(13) *Пользователь*: *Слишком пафосно. Перепиши короче, добавь шутку про дедлайны, но чтобы не обидно.*
(14) *ИИ*: *«Иван Иванович! Желаю, чтобы в жизни, как и в коде, всё было без багов, дедлайны горели только от счастья, а аптайм был 99.9%. С юбилеем!»*
(15) *Пользователь (в мессенджере коллегам)*: *Завтра тостую, готовьтесь. Накидал тут с ChatGPT, но вроде норм.*

Пример 2 (англоязычный, адаптировано из корпуса):

(16) *User*: *Write a best man speech for my brother's wedding. Funny but not offensive, 3 minutes.*
(17) *AI*: *"Good evening everyone! For those who don't know me, I'm [Name], the lucky guy who got to be the best man. I've known the groom for 25 years, which means I have 25 years of embarrassing stories — but don't worry, I'll only share the ones that won't get me uninvited from Christmas dinner..."*
(18) *User*: *Too long. Cut to 2 minutes, add a story about when we were kids.*
(19) *AI*: *[сокращенная версия с добавлением истории]*
(20) *User (posting on Reddit)*: *Here's my speech. Used ChatGPT for the structure, but added personal stories. Thoughts?*

Таблица 7. **Сопоставительный анализ гибридных дискурсивных актов**
Table 7. **Comparative Analysis of Hybrid Discursive Acts**

Параметр / Parameter	**Русскоязычные ГДА / Russian-speaking Discursive Acts**	**Англоязычные ГДА / English speaking Discursive Acts**
Частотность в корпусе / Frequency in the case	12.3 % диалоговых цепочек	15.8 % диалоговых цепочек
Среднее количество итераций / Average number of iterations	2.8	3.2
Наиболее частые объекты редактирования / The most common editing objects	Тональность (48 %), длина (42 %)	Тональность (52 %), персонализация (44 %)
Маркеры гибридности в пост-коммуникации / Markers of hybridity in post-communication	*накидал с нейросети, поправил немного, ChatGPT помог*	*used AI for the draft, ChatGPT helped, AI-generated with edits*

Источник: исследование Н.А. Ахреновой.
Source: compiled by Natalia A. Akhrenova.

Интерпретация. В обоих языках гибридные дискурсивные акты демонстрируют сходную структуру: человеческая инициация (промпт) → машинная генерация → человеческая редактура (часто многошаговая) → презентация результата как (квази-)собственного в человеческой коммуникации. Различия касаются преимущественно дискурсивных маркеров, сопровождающих презентацию: англоязычные пользователи более склонны открыто признавать использование ИИ (*used AI*, *AI-generated*), в то время как русскоязычные чаще используют эвфемизмы (*накидал*, *поправил*), минимизирующие роль алгоритма.

Финальный продукт (тост, речь, письмо) является результатом гибридной интенциональности: цель, тема и финальная редактура принадлежат человеку, базовая структура и значительная часть лексического наполнения сгенерированы ИИ. Атрибуция авторства такого высказывания в рамках классической теории невозможна, что подтверждает необходимость введения понятия гибридного дискурсивного акта как новой единицы анализа цифровой коммуникации.

Обобщение результатов анализа. Проведенный многоуровневый сопоставительный анализ русскоязычной и англоязычной частей корпусов позволяет сделать следующие обобщающие выводы.

1. Количественно подтверждена гипотеза о гомогенизации: корпус A2H демонстрирует снижение лексического разнообразия в среднем на 37,6 % (русский: 38,7 %, английский: 35,9 %) и синтаксического разнообразия на 45,3 % (русский: 46,1 %, английский: 44,5 %) по сравнению с корпусом H2H. Близость показателей в обоих языках свидетельствует об универсальном характере данного феномена, определяемом архитектурой языковых моделей.

2. Выявлена специфика жанра промпта: корпус H2A в обоих языках характеризуется высокочастотными императивными конструкциями и метаязыковыми маркерами параметризации, что позволяет квалифицировать промпт как новый речевой жанр — директивно-эксплораторный. Межъязыковые различия минимальны и касаются преимущественно грамматического оформления императива.

3. Обнаружен феномен алгоритмической нормы: тексты A2H в обоих языках устойчиво воспроизводят клишированные конструкции книжного регистра и абстрактную лексику, формируя новый тип дискурсивной практики. Русскоязычные ИИ-тексты демонстрируют бóльшую склонность к безличным конструкциям, англоязычные — к конструкциям с формальным подлежащим.

4. Зафиксированы гибридные дискурсивные акты: в диалогах человек-ИИ выделены коммуникативные единицы, авторство которых не может быть однозначно приписано ни человеку, ни машине. В англоязычной части корпуса частота ГДА несколько выше (15,8 % против 12,3 %), что может объясняться более длительной историей взаимодействия с технологией.

5. Установлена универсальность ключевых трансформаций: несмотря на типологические различия русского и английского языков, основные тенденции — гомогенизация, параметризация промптов, алгоритмическая норма, гибридное авторство — проявляются в обоих языках сходным образом, что позволяет говорить об общецивилизационном характере описываемых процессов.

Обсуждение результатов

Полученные в ходе эмпирического анализа результаты требуют содержательной интерпретации в контексте современных теоретических дискуссий о природе цифровой коммуникации.

Наше исследование подтверждает тезис о гомогенизации алгоритмического дискурса, что согласуется с наблюдениями Т.А. Литвиновой [3] о «сглаживании» идиостиля в текстах, созданных при участии ИИ. Однако наши данные заставляют усомниться в оптимистичных прогнозах о том, что ИИ может выступать инструментом стилистического обогащения речи. Напротив, «базовый» режим генерации тяготеет к усредненности. В терминах, предложенных С. Херринг [3], можно говорить о том, что фактор технологической опосредованности (technological mediation) в случае с LLM перерастает в фактор «алгоритмической детерминации», жестко задающий стилистические параметры высказывания.

Выделенная стратегия параметризации промптов коррелирует с выводами исследователей промпт-инжиниринга [6; 9], которые подчеркивают необходимость экспликации «неявного знания» для эффективного взаимодействия с ИИ. С лингвистической точки зрения это означает, что жанр промпта требует от пользователя уникальной метакоммуникативной

компетенции — способности вербализовывать то, что в человеческом общении передается через контекст и интонацию. Это принципиально отличает его от традиционных директивных речевых жанров.

Проблема агентности и концепция гибридного автора. Введение нами понятия «гибридный дискурсивный акт» развивает идеи западных коллег о «гибридном авторстве» [4; 5]. Наш вклад заключается в операционализации этого понятия через конкретные критерии (наличие итераций, следы редактуры, неоднозначность атрибуции) и выявлении его языковых маркеров. Зафиксированное в нашем корпусе различие в степени открытости признания использования ИИ (прямое признание в английском и эвфемизмы в русском) указывает не только на языковые, но и на культурные различия в восприятии статуса алгоритма как соавтора.

Цифровая антрополингвистика и трансформация языковой личности. Наконец, наши данные подтверждают тезис о необходимости смены парадигмы, развитый в рамках цифровой антрополингвистики. Появление нового актора (алгоритма) и новой единицы анализа (гибридного дискурсивного акта) означает, что мы больше не можем изучать языковую личность изолированно. Формируется «расширенная языковая личность», чья речевая деятельность неразрывно связана с алгоритмическими инструментами, что требует пересмотра методологических подходов к исследованию дискурса в XXI в. Традиционное понимание интернет-дискурса как гибридного образования, сочетающего черты устной и письменной речи в условиях компьютерно-опосредованной коммуникации [11], сохраняя свою эвристическую ценность, более не отражает всей полноты и сложности изучаемого феномена.

Заключение

Проведенное исследование, выполненное на материале трех контрастных корпусов общим объемом 1,5 млн словоупотреблений на русском и английском языках, позволяет сделать ряд теоретически и эмпирически значимых выводов, подтверждающих тезис об алгоритмическом повороте в интернет-дискурсе.

Так, количественный анализ доказал, что тексты, генерируемые большими языковыми моделями (корпус A2H), демонстрируют статистически значимое снижение лексического (на 37,6 %) и синтаксического (на 45,3 %) разнообразия по сравнению с естественной человеческой коммуникацией (корпус H2H). Близость показателей для русского и английского языков свидетельствует об универсальном характере данного феномена, обусловленного архитектурой LLM, которые продуцируют текст на основе усредненных статистических паттернов.

Корпус запросов к ИИ (H2A) позволил выделить и описать директивно-эксплораторный жанр промпта. Его ключевыми характеристиками являются высокая частотность императивных конструкций, параметризация запроса

(объем, стиль, тональность) и метаязыковая рефлексия пользователя, вынужденного эксплицировать имплицитные нормы коммуникации. Выделенная типология стратегий (делегирование, параметризация, итеративное уточнение) отражает эволюцию взаимодействия человека с алгоритмом.

На пересечении корпусов H2A и A2H зафиксированы коммуникативные единицы, авторство которых не может быть однозначно атрибутировано. Феномен гибридной агентности, при котором интенциональность распределена между человеком (цель, тема, редактура) и алгоритмом (лексико-синтаксическая реализация), потребовал введения нового понятия — гибридный дискурсивный акт. Этот концепт позволяет описывать реальность, в которой граница между «естественным» и «искусственным» в тексте стирается.

Анализ ключевых слов и коллокаций в корпусе A2H выявил устойчивое воспроизводство клишированных конструкций книжного регистра («важно отметить», «таким образом», «исследование показывает») и абстрактной лексики. Это формирует особую алгоритмическую норму — усредненный, стилистически гомогенизированный регистр, который начинает оказывать обратное влияние на пользователей, формируя новые стандарты «правильного» письма.

Сопоставительный анализ русскоязычной и англоязычной частей корпусов показал, что, несмотря на типологические различия языков, ключевые тенденции (гомогенизация, параметризация промптов, формирование алгоритмической нормы) проявляются сходным образом. Это позволяет говорить о глобальном, а не узколингвистическом характере описываемых процессов, что вносит вклад в развитие цифровой антрополингвистики.

Проведенное комплексное исследование, опирающееся на репрезентативный корпусный материал и межъязыковой сопоставительный анализ, позволяет не только обобщить выявленные трансформационные процессы, но и предложить концептуальное переосмысление самого феномена интернет-дискурса в новую алгоритмическую эпоху. На основании полученных результатов мы предлагаем расширенную дефиницию, учитывающую произошедший алгоритмический поворот

Интернет-дискурс в современном понимании представляет собой многокомпонентное, динамически развивающееся коммуникативное пространство, в котором взаимодействуют человеческие и алгоритмические агенты, продуцирующие поликодовые, полимодальные тексты в условиях гибридной агентности и варьирующейся степени алгоритмической опосредованности, что приводит к формированию новых жанровых форм (промпт), трансформации традиционных категорий авторства и читательства, а также к стилистической гомогенизации дискурсивных практик под влиянием статистически усредненных языковых моделей.

Данное определение фиксирует принципиальные изменения, выявленные в ходе исследования:

1. Расширение субъектного состава дискурса. К традиционной дихотомии «адресант — адресат» добавляется третий, нечеловеческий актор —

алгоритм, который может выступать в различных ипостасях: генератора, модератора, соавтора или инструмента. Это требует пересмотра категории агентности в дискурсивном анализе.

2. Трансформация текстуальности. Единицей анализа более не может выступать стабильный, линейный текст. Современный интернет-дискурс реализуется в форме поликодовых, полимодальных образований, где вербальный код неразрывно связан с визуальным (эмодзи, мемы, гифки) и техническим (хештеги, гиперссылки, алгоритмическая курация).

3. Появление новых жанров, формирующих и направляющих коммуникацию. Возникновение промпта как самостоятельного речевого жанра (директивно-эксплораторного) свидетельствует о формировании принципиально новых коммуникативных практик, требующих от пользователя развития метаязыковой компетенции — способности эксплицировать имплицитные нормы общения.

4. Выявленное снижение лексического и синтаксического разнообразия в алгоритмическом дискурсе (на 37,6 и 45,3 % соответственно) указывает на объективные, статистически верифицируемые изменения языкового ландшафта интернет-коммуникации.

Резюмируя сказанное, следует подчеркнуть, что алгоритмический поворот в интернет-дискурсе знаменует собой не локальное изменение отдельных языковых практик, а фундаментальный сдвиг в самой природе человеческой коммуникации. Цифровая среда, перестав быть нейтральным каналом передачи сообщений, превратилась в активного со-творца смыслов. Формирующаяся алгоритмическая норма, основанная на статистическом усреднении, начинает оказывать обратное влияние на речевую деятельность человека, создавая прецедент, при котором «искусственное» (порожденное моделью) постепенно становится образцом для «естественного».

В этих условиях цифровая антрополингвистика, выдвигающая на первый план изучение гибридной агентности, алгоритмической опосредованности и трансформирующейся языковой личности, становится не просто новым научным направлением, а необходимым методологическим основанием для понимания того, как технологии переформатируют самое древнее и фундаментальное свойство человека — способность к языку и созданию культуры. Дальнейшие исследования в этой области должны быть направлены на разработку методов лингвистической экспертизы в условиях гибридного авторства, изучение долгосрочных последствий алгоритмической гомогенизации для языкового разнообразия, а также на этическую рефлексию последствий делегирования коммуникативных функций искусственному интеллекту.

Список литературы

1. *Crystal D.* Language and the Internet. Cambridge : Cambridge University Press, 2006.
2. *Herring S.C.* A faceted classification scheme for computer-mediated discourse // Language@Internet. 2007. Vol. 4. Article 1.

3. *Litvinova T.A., Mikros G.K., Dekhnich O.V.* Writing in the era of large language models: a bibliometric analysis of research field // Research Result. Theoretical and Applied Linguistics. 2024. Т. 10. № 4. С. 15–16. https://doi.org/10.18413/2313-8912-2024-10-4-0-1 EDN: RLVGOC
4. *Bozkurt A.* Tell Me Your Prompts and I Will Make Them True: The Alchemy of Prompt Engineering and Generative AI // Open Praxis. 2024. Vol. 16. № 2. P. 111–118.
5. *Henrickson L., Meroño-Peñuela A.* Prompting meaning: a hermeneutic approach to optimising prompt engineering with ChatGPT // AI & SOCIETY. https://doi.org/10.1007/s00146-023-01752-8 EDN: ROVQJU
6. *Velásquez-Henao J.D., Franco-Cardona C.J., Cadavid-Higuita L.* Prompt Engineering: a methodology for optimizing interactions with AI-Language Models in the field of engineering // DYNA. 2023. № 90 (230). P. 9–17.
7. *Ма Ч.* Промпт-инжиниринг в лингводидактике: стратегии взаимодействия с чат-ботами в обучении китайских студентов русскому языку как иностранному // Филологические науки. Вопросы теории и практики. 2025. Т. 18. № 12. С. 5159–5166. https://doi.org/10.30853/phil20250698 EDN: OVOCIH
8. *Kochetova L.A.* Linguocultural Specifics of Artificial Intelligence Representation in the English Language Media Discourse: Corpus-Based Approach // Science Journal of Volgograd State University. Linguistics. 2023. Vol. 22. № 5. С. 6–18. https://doi.org/10.15688/jvolsu2.2023.5.1 EDN: OZDUSR
9. *Остапенко С.В., Халина Н.В.* Лингвистическая оптимизация промта в аспекте проблематики промт-инжиниринга // Вестник Тюменского государственного университета. Гуманитарные исследования. Humanitates. 2025. Т. 11. № 1 (41). С. 36–50. https://doi.org/10.21684/2411-197X-2025-11-1-36-50 EDN: XCFQTA
10. *Батурина И.В.* Метатренды и мифологемы тематики искусственного интеллекта западного научного дискурса // Известия Саратовского университета. Новая серия. Серия: Философия. Психология. Педагогика. 2024. Т. 24. № 1. С. 14–18. https://doi.org/10.18500/1819-7671-2024-24-1-14-18 EDN: RJCVCM
11. *Ахренова Н.А.* Интернет-лингвистика: доминантный подход. М. : Русайнс, 2024.

References

1. Crystal, D. (2006). *Language and the Internet*. Cambridge: Cambridge University Press.
2. Herring, S.C. (2007). A Faceted Classification Scheme for Computer-Mediated Discourse. *Language@Internet*, 4, 1.
3. Litvinova, T.A., Mikros, G.K., & Dekhnich, O.V. (2024). Writing in the Era of Large Language Models: a Research Result. *Theoretical and Applied Linguistics*, *10*(4), 15–16. (In Russ.). https://doi.org/10.18413/2313-8912-2024-10-4-0-1 EDN: RLVGOC
4. Bozkurt, A. (2024). Tell Me Your Prompts and I Will Make Them True: The Alchemy of Prompt Engineering and Generative AI. *Open Praxis*, *16*(2), 111–118.
5. Henrickson, L., & Meroño-Peñuela, A. (2023). Prompting Meaning: a Hermeneutic Approach to Optimising Prompt Engineering with ChatGPT. *AI & SOCIETY*. https://doi.org/10.1007/s00146-023-01752-8 EDN: ROVQJU
6. Velásquez-Henao, J.D., Franco-Cardona, C.J., & Cadavid-Higuita, L. (2023). Prompt Engineering: a Methodology for Optimizing Interactions with AI-Language Models in the Field of Engineering. *DYNA*, 90(230), 9–17.
7. Ma, Ch. (2025). Prompt Engineering in Linguodidactics: Strategies for Interacting with Chatbots in Teaching Russian as a Foreign Language to Chinese Students. *Philology. Theory & Practice*, *18*(12), 5159–5166. (In Russ.). https://doi.org/10.30853/phil20250698 EDN: OVOCIH
8. Kochetova, L.A. (2023). Linguocultural Specifics of Artificial Intelligence Representation in the English Language Media Discourse: Corpus-Based Approach.

Science Journal of Volgograd State University. Linguistics, *22*(5), 6–18. https://doi.org/10.15688/jvolsu2.2023.5.1 EDN: OZDUSR
9. Ostapenko, S.V., & Khalina, N.V. (2025). Linguistic Optimization of Prompts Regarding the Challenges in Prompt Engineering. *Tyumen State University Herald. Humanities Research. Humanitates*, *11*, 1 (41), 36–50. (In Russ.). https://doi.org/10.21684/2411-197X-2025-11-1-36-50
10. Baturina, I.V. (2024). Megatrends and Mythologems of Artificial Intelligence Topics in Western Scientific Discourse. *Izvestiya of Saratov University. New Series. Series: Philosophy. Psychology. Pedagogy*, *24*(1), 14–18. (In Russ.). https://doi.org/10.18500/1819-7671-2024-24-1-14-18 EDN: RJCVCM
11. Akhrenova, N.A. (2024). *Internet Linguistics: a Dominant Approach*. Moscow: Rusains. (In Russ.).

Сведения об авторе:
Ахренова Наталья Александровна, доктор филологических наук, доцент, руководитель отделения постдипломного образования, профессор кафедры теории и практики иностранных языков, ведущий научный сотрудник института иностранных языков, Российский университет дружбы народов (117198, Российская Федерация, г. Москва, ул. Миклухо-Маклая, д. 6); *сфера научных интересов*: дискурс-анализ, цифровая гуманитаристика, медиалингвистика, интернет-лингвистика, межкультурная коммуникация; *e-mail*: nakhrenova@mail.ru
ORCID: 0000-0002-1265-5595, SPIN-код: 9818-1046, AuthorID: 265006.

Information about the author:
Natalia A. Akhrenova, Dr. Sc. (Philology), Associate Professor, Head of the Postgraduate Education Department of the Institute of Foreign Languages, RUDN University (6 Miklukho-Maklaya St., Moscow, Russian Federation, 117198); *Research interests*: discourse analysis, digital humanities, media linguistics, Internet linguistics, intercultural communication; *e-mail*: nakhrenova@mail.ru
ORCID: 0000-0002-1265-5595, SPIN-code: 9818-1046, AuthorID: 265006.

RUDN Journal of Language Studies, Semiotics and Semantics
2313-2299 (Print), ISSN 2411-1236 (Online)
Вестник Российского университета дружбы народов.
Серия: ТЕОРИЯ ЯЗЫКА. СЕМИОТИКА. СЕМАНТИКА

2026 Vol. 17 No. 2 427–441
http://journals.rudn.ru/semiotics-semantics

DOI: 10.22363/2313-2299-2026-17-2-427-441
EDN: LIUVUU

Научная статья / Research Article

В поисках алгоритма определения текста, сгенерированного нейросетевой моделью, с опорой на его языковые особенности

Л.Р. Комалова

Институт научной информации по общественным наукам Российской академии наук, *Москва, Российская Федерация*
Московский государственный лингвистический университет, *Москва, Российская Федерация*
✉komalova@inion.ru

Аннотация. С 2023 г. наблюдается устойчивый рост научных публикаций, в которых присутствует текст, сгенерированный нейронными сетями на основе больших языковых моделей. Развитие методов машинного обучения и глубокого обучения применительно к задачам обработки естественного языка приближает результат генерации текста к текстам, написанным человеком. На данном этапе развития лингвистика не располагает разработанным инструментарием, позволяющим разграничить текст, написанный человеком, и текст, сгенерированный нейросетью. Цель проводимого исследования заключается в выработке алгоритма формальной проверки рукописи, поступившей в редакцию научного журнала / научному руководителю, с опорой на доступные научному редактору / научному руководителю инструменты (в основном открытого доступа) детекции сгенерированного нейросетями текста на русском языке. Методология исследования опирается на результаты изысканий в области компьютерной и математической лингвистики по описанию «стиля письма» различных нейросетей при генерации текстов научных жанров на русском языке. Материалом исследования послужили научные публикации за 2021–2025 гг., доступные через базы данных научной информации, авторами которых являются исследователи, аффилированные преимущественно с российскими научными и образовательными организациями, а также с российскими технологическими компаниями. В ходе исследования было выявлено, что практика применения Gen-AI при написании научных текстов намного опережает рефлексию ее результатов и прогнозирование последствий в долгосрочной перспективе. При этом калейдоскоп публикаций, описывающих отличительные черты результата генерации различными нейросетями, позволяет выработать подход к обнаружению подобных текстов, опираясь на ошибки в генерации на различных языковых уровнях.

Ключевые слова: генеративный искусственный интеллект, обнаружение сгенерированного текста

© Комалова Л.Р., 2026
This work is licensed under a Creative Commons Attribution-NonCommercial 4.0 International License
https://creativecommons.org/licenses/by-nc/4.0/legalcode

Финансирование: Исследование выполнено в рамках государственного задания ФГБУН «Институт научной информации по общественным наукам РАН» по теме «Новейшие тенденции в развитии гуманитарной мысли России и Запада: философские течения, развитие культурологии и духовная жизнь общества в цифровую эпоху».

Конфликт интересов: автор заявил об отсутствии конфликтов интересов.

История статьи: дата поступления: 23.12.2025; дата приема в печать: 15.02.2026.

Для цитирования: *Комалова Л.Р.* В поисках алгоритма определения текста, сгенерированного нейросетевой моделью, с опорой на его языковые особенности // Вестник Российского университета дружбы народов. Серия: Теория языка. Семиотика. Семантика. 2026. Т. 17. № 2. С. 427–441. https://doi.org/10.22363/2313-2299-2026-17-2-427-441 EDN LIUVUU

Towards an Algorithm for Identifying Text Generated by Neural Network Models through Linguistic Cues

Liliya R. Komalova

Institute of Scientific Information for Social Sciences of the Russian Academy of Sciences, *Moscow, Russian Federation*
Moscow State Linguistic University, *Moscow, Russian Federation*
✉komalova@inion.ru

Abstract. Since 2023, there has been a steady increase in research papers containing text generated by neural networks based on large language models. The development of machine and deep learning methods applied to natural language processing tasks is bringing the results of text generation closer to those of human authors. At this stage of development, linguistics lacks practical tools to distinguish between human-written text and text generated by a neural network. The goal of this study is to develop an algorithm for manuscripts formal verification by editors of scientific journals / academic supervisors, relying on (primarily open access) tools available to the scientific editor / supervisor for detecting text generated by neural networks in Russian language. The study methodology is based on results of computational and mathematical linguistics researches to describe the "writing style" of various neural networks when generating scientific texts in Russian language. We analyzed research papers issued from 2021 to 2025, accessible through scientific information databases, authored by researchers primarily affiliated with Russian scientific and educational organizations, as well as Russian technology companies. The study revealed that the use of Gen-AI in academic writing significantly outpaces the reflection of its results and the prediction of its long-term implications. Furthermore, the plethora of publications describing the distinctive features of the output generated by various neural networks allows for the development of an approach to detecting such texts based on generation errors at various linguistic levels.

Keywords: generative artificial intelligence, generated text detection, linguistic features of generated text

Financing: The research is carried out within the framework of the state assignment to the Federal State Budgetary Institution of Science «Institute of Scientific Information for Social Sciences of the Russian Academy of Sciences», project «The latest trends in the development of humanitarian thought in Russia and the West: philosophical movements, the development of Cultural Studies and the spiritual life of society in the digital era».

Conflict of interests: The author stats that there were no conflicts of interest.

Article history: Received: 23.12.2025; Accepted: 15.02.2025.

For citation: Komalova, L.R. (2026). Towards an Algorithm for Identifying Text Generated by Neural Network Models through Linguistic Cues. *RUDN Journal of Language Studies, Semiotics and Semantics, 17*(2), 427–441. (In Russ.). https://doi.org/10.22363/2313-2299-2026-17-2-427-441 EDN LIUVUU

Генеративный искусственный интеллект — новая текстовая реальность

В статье 5 Национальной стратегии развития искусственного интеллекта на период до 2030 г., утвержденной Указом Президента Российской Федерации от 10 октября 2019 г. № 4901, искусственный интеллект определяется как «комплекс технологических решений, позволяющий имитировать когнитивные функции человека (включая самообучение и поиск решений без заранее заданного алгоритма) и получать при выполнении конкретных задач результаты, сопоставимые с результатами интеллектуальной деятельности человека или превосходящие их»[1]. В стратегии говорится об ускоренном внедрении технологических решений, разработанных на основе искусственного интеллекта, в различные отрасли экономики и сферы общественных отношений. В документе указывается как на позитивные изменения в ключевых сферах жизнедеятельности, которые сопровождаются использованием технологий искусственного интеллекта (ИИ), так и на негативные последствия, вызванные социальными и технологическими изменениями, которые сопутствуют развитию технологий искусственного интеллекта.

Многие эксперты и практики отмечают, что применение ИИ-инструментов имеет потенциал для преобразования различных сфер жизни, включая образование и научные исследования [1; 2]. Всплеск активности на уровне рядовых пользователей произошел после введения в массовое пользование предназначенных для генерации текста в письменной форме чат-ботов на основе предобученных нейронных сетевых языковых моделей, основанных на архитектуре типа «трансформеры» (generative pre-trained transformer, GPT), на механизме самовнимания (self-attention) [3] и на парадигме самообучения (self-supervised learning) на огромном объеме текстовых данных. В отличие от традиционных задач классификации и регрессии генеративный ИИ (Gen-AI) обладает способностью автономно генерировать новые данные, включая изображения, музыку и текст [4; цит. по: 5].

Генерация текста реализуется с помощью двух базовых подходов [6. С. 283]: (1) использование предопределенного шаблона для генерации текста на основе различных входных данных (например, используя предопределенный шаблон, искусственный интеллект может сгенерировать определенное описание продукта на основе типа продукта, его характеристик и преимуществ); (2) использование метода неконтролируемого (глубоком)

[1] См.: Указ Президента РФ от 10 октября 2019 г. № 490 «О развитии искусственного интеллекта в Российской Федерации» URL: https://base.garant.ru/72838946/?ysclid=ml9m8kv11q126675642 (дата обращения: 05.12.2025).

обучения, который изучает нюансы языковых структур и функционирует с условием наличия больших объемов входных данных — по сравнению с подходом на основе шаблонов результат генерации текста на основе метода глубокого обучения приближается к тексту, написанному человеком.

«Генеративные нейросети настроены на создание наиболее предсказуемых, статистически чаще встречающихся последовательностей слов, вызывающих наименьшее недоумение у читателя и с большей вероятностью воспринимаемых как осмысленные» [7. С. 91]. В основе работы моделей лежат статистические пермутации (перестановки элементов), «пропущенные через миллионные обучающие выборки на примерах» [8. С. 470]. Развитие методов машинного глубокого обучения применительно к задачам обработки естественного языка «позволяет получать искусственно созданные тексты, степень сходства которых с написанными человеком с годами становится лишь выше» [9. С. 308]. Однако статистическое «понимание» языка, которым обладают генеративные нейросетевые модели, отлично от семантического понимания языка человеком [10. С. 235].

Сгенерированный текст как лингвистическая проблема

Начиная с 2020 г. мир захлестнула волна публикаций, искусственно сгенерированных с помощью технологии GPT под видом статей, написанных человеком [11]. К настоящему времени «сгенерированные нейросетями тексты уже стали частью нашей коммуникативной среды и культуры» [12. С. 790]. В частности, в образовательной среде возникла коллизия между широким функционалом Gen-AI (системами аннотирования, реферирования, перефразирования для подготовки и написания рефератов, обзорных и исследовательских статей, выпускных квалификационных работ) и развитием творческих исследовательских способностей у обучающихся вузов, а также возможностями преподавателей оценить интеллектуальные компетенции обучающихся в сложившейся ситуации [13. С. 58—59]. Опасность появления большого количества синтетических данных заключается в том, что они становятся (пусть и формально) частью научной онтологии, материалом для обучения следующего поколения генеративных нейросетевых языковых моделей, в то время как «программы-детекторы работают пока недостаточно надежно» [8. С. 476].

Можно выделить несколько типов недобросовестного использования генеративных нейросетей [9; 14]:

— из-за огромного объема данных для обучения модель генерирует ответ, который не всегда соответствует действительности, при этом пользователи не всегда проверяют информацию, полученную в результате генерации;

— генерация научного или учебного текста, гарантированно проходящего проверку на наличие плагиата;

— перефразирование текста для того, чтобы затруднить системам обнаружения заимствований выявление того, что текст был ранее опубликован.

Задача «проверяющего» усложняется тем, что текст чаще всего не генерируется моделями целиком: как правило, это умелая комбинация машинно сгенерированных и написанных человеком текстовых фрагментов [15; 16].

Следует зафиксировать, что на данном этапе развития лингвистика не располагает разработанным инструментарием (устойчивым комплексом признаков), позволяющим разграничить текст, написанный человеком, и текст, сгенерированный нейросетью [17. С. 165]. При этом «тексты одной жанровой принадлежности, но сгенерированные разными языковыми моделями, существенно различаются» как по языковым параметрам, так и по ошибкам, которые делают эти генеративные нейросети [18; цит. по: 12. С. 789]. Результат генерации зависит также и от задачи, которую пользователь формулирует в промпте (сбор, реферирование, аннотирование текстового материала, его перефразирование и т.д.), и от субъективных показателей автора промпта [13. С. 62].

Цель проводимого нами исследования заключается в выработке алгоритма формальной проверки рукописи, поступившей в редакцию научного журнала / научному руководителю, с опорой на доступные научному редактору / научному руководителю инструменты (в основном открытого доступа) детекции сгенерированного нейросетями текста на русском языке.

Методология исследования опирается на результаты изысканий в области компьютерной лингвистики по описанию «стиля письма» различных нейросетей при генерации текстов научных жанров на русском языке.

Материалом для исследования послужили научные публикации за 2021–2025 г., доступные через базы данных научной информации, авторами которых являются исследователи, аффилированные преимущественно с российскими научными и образовательными организациями, а также с российскими технологическими компаниями.

Автоматическая детекция сгенерированного текста

Как отмечается в ряде работ в области математической лингвистики, существующие методы не решают проблемы детекции текстов большого объема [9], так как длина фрагментов, на которые разбивается анализируемый текст, оказывает влияние на результат детекции [19; 20]. Для разных языков оптимальная длина варьируется, но в каждом существует плато, позволяющее зафиксировать качество классификации [9]. Для русского языка это порядка 4000 символов.

Высокий уровень полноты детекции и большая обобщающая способность модели достигаются при обучении детектирующей модели на мультидоменных текстах: чем больше разнородных по тематикам текстов будет на стадии обучения, тем выше будут метрики качества классификации при использовании детектора в реальных задачах поиска искусственно сгенерированных фрагментов текста [21; цит. по: 9. С. 312].

Автоматическая детекция сгенерированного текста строится на различных статистических метриках [6; 7; 22–24]: частоте совместной встречаемости n-грамм, вычислении коэффициентов перекрытия слов, длины предложений, распределения слов в предложении, значений «бурности» (burstiness) и «недоумения» (perplexity), вычислении различных метрик сложности текста. Если после генерации текста нейросетью проводится постредактирование, его идентификация становится еще более сложной. Известны следующие подходы к детекции такого типа искусственно сгенерированных текстов [25. С. 963]:

— обнаружение неконтекстуальных токенов (non-contextual tokens): например, ChatGPT-модель часто вставляет специальные токены или маркеры в свои ответы для указания начала или конца ответа;

— анализ статистических свойств сгенерированного текста: распределение слов, грамматические или стилистические особенности (проверка использования редких или специфичных слов и фраз, обнаружение неправильной грамматики или «плохого» стиля);

— выявление метаданных: для обнаружения применения, например, ChatGPT-модели используется дополнительная информация, такая как сведения о запросах, метаданные сеанса и временные отметки;

— анализ контекста вопросов и ответов: если в ответах модели отсутствуют устойчивые или явные связи с предыдущими вопросами либо контекст быстро теряется, это может быть признаком автоматической генерации;

— сравнение следующих токенов при совпадающих предыдущих (идентичном левом контексте) в проверяемом тексте и в тексте, генерируемом GPT-подобной моделью;

— проверка фрагментов текста путем их поиска в интернет-архивах.

Системы, детектирующие искусственно сгенерированный текст, широко представлены в сети Интернет[2]. Результат детекции подается либо в виде ответа типа «текст сгенерирован / текст написан человеком», либо в виде указания доли предположительно искусственно сгенерированного текста.

Экспертное обнаружение сгенерированного текста

В результате анализа научной литературы было принято решение объединить выявленные маркеры сгенерированности текста в три группы. Макроуровень охватывает всю рукопись и также касается взаимодействия структурных частей рукописи между собой, мезоуровень относится к отдельным абзацам и взаимодействию между ними, микроуровень характеризует особенности текста на уровне предложения и взаимодействия между предложениями в рамках одного абзаца.

[2] См., например: GigaCheck. URL: https://developers.sber.ru/portal/products/gigacheck (accessed: 05.12.2025); GPTZero. URL: https://gptzero.me (accessed: 05.12.2025); ZeroGPT. URL: https://www.zerogpt.com (accessed: 05.12.2025); Copyleaks. URL: https://copyleaks.com (accessed: 05.12.2025); Contentatscale. URL: https://ai.tenorshare.com/products/contentatscale-ai-detector (accessed: 05.12.2025) и др.

Макроуровень

Наиболее частотной ошибкой, упоминаемой в научной литературе, которая отличает сгенерированный текст от текста, написанного человеком, является некорректная цитация и составление списка литературы. Внутритекстовые отсылки к списку литературы в сгенерированных текстах либо не соответствуют списку литературы, либо являются выдуманными, т.е. таких источников в научной онтологии не существует, а указанные в списке идентификаторы (например DOI) связаны с иными публикациями [24; 26–28]. Частота данной особенности сгенерированных текстов вынуждает экспертов (научных редакторов и руководителей) проверять каждую цитату и каждую позицию в списке литературы.

Второй и более серьезной частотной погрешностью сгенерированных текстов является формулировка недостоверной информации и искажение фактов [5; 12; 25; 29], что требует высокой компетенции проверяющего в предметной области и тематике проверяемой рукописи. Фактчекинг — серьезный вызов, т.к. в случае включения недостоверной информации в научный оборот, впоследствии она может лечь в основу методологии исследования неопытного исследователя, что повлечет за собой дальнейшее искажение научных данных.

Исследователи отмечают, что сгенерированные нейросетью тексты носят поверхностный характер (на основе поверхностного анализа), отличаются неясностью формулировок, в таких текстах взаимосвязь между его частями не всегда прослеживается, исследуемый вопрос в полной мере не раскрывается, такие тексты содержат общую информацию и имеют более простую структуру [5; 7; 29–32]. При генерации текста обзорной части рукописи выявляются многочисленные ошибки при перефразировании, а также явный плагиат [33]. Такие тексты тяготеют к публицистическому, а иногда и обиходно-разговорному функциональному стилю письма. При этом «сгенерированные тексты отличает более четкая организация, соответствующая правилам IMRaD» [24. С. 1002].

Отмечается «неспособность языковых моделей выстраивать долгосрочные семантические зависимости в тексте, что иногда приводит к „вырождению“ текста» [34; цит. по: 17. С. 166]. В сгенерированных нейросетью текстах отмечается единообразный стиль изложения, в то время как текст, написанный человеком, «может иметь изменения стиля и тона по всему тексту» [25. С. 962]. В силу программных особенностей больших языковых моделей в порождаемых нейросетями текстах высока лексическая плотность[3] при низком лексическом разнообразии[4] [5].

[3] Количественная характеристика обширности словарного запаса — соотношение содержательных слов (самостоятельных частей речи: имен существительных, имен прилагательных, глаголов и наречий) в тексте к общему количеству слов в тексте.

[4] Количественная характеристика текста, отражающая степень разнообразия словаря при построении текста заданной длины — соотношение отдельных лексических единиц (лемм) и количества их употребления в тексте (словоформ).

Сгенерированный нейросетью текст отличает «отсутствие идеи, объективно или субъективно характеризующей наличие нового знания» [7. С. 92]. Такой текст не содержит научной новизны, не имеет теоретической и практической значимости [29]. За сгенерированным текстом не «проглядывается» личность автора [24]. Эти особенности может обнаружить только человек — эксперт в предметной области, затронутой в тексте.

Мезоуровень

К особенностям, позволяющим обнаружить сгенерированные фрагменты текста, можно отнести [29. С. 155; 35. С. 50–51]:

— «закольцованность» темы (т.е. повтор темы без ввода рематического компонента);

— наличие множества рем, ни одна из которых в дальнейшем не преобразуется в тему и не продолжает свое текстовое развитие;

— недостаточный учет факторов автора и адресата;

— маркированность текста показателями модуса предположения, некатегоричности, сигнализирующими о наличии других, альтернативных вариантов ответа;

— повторение одной и той же мысли в разных предложениях;

— отсутствие связи сгенерированного фрагмента с остальным текстом.

Микроуровень

На уровне предложений сгенерированные тексты обладают меньшей степенью связности, чем тексты, написанные человеком. В сгенерированных текстах обнаруживается «достраивание нарратива причинно-следственными операторами *поэтому*, *так что*, *вот и*, *действительно* и т.п., что устраняет „нарушения“ в оригинальном тексте-источнике» [12. С. 789]. Генеративные «нейросети создают не только предсказуемый порядок слов, но и строят предложения с несложной структурой со средней длиной порядка 10–20 слов» [7. С. 91], в то время как текст, написанный человеком, более насыщен причастными и деепричастными оборотами.

В сгенерированных текстах встречаются нарушения синтаксической сочетаемости; избыточность клише, штампов, канцеляризмов и слов с более высоким уровнем формальности; неестественная сочетаемость и частота специфической терминологии, которая используется для создания более авторитетного или научного стиля текста; неуместные уточнения; формальные повторы словосочетаний и отдельных слов; аббревиатуры терминов на английском языке, которые приводятся в дополнении к терминам на русском; ошибки или искажения в лексической сочетаемости; нарушение сочетаемости слов по линии предсказуемости / непредсказуемости, обязательности / факультативности синтаксической связи; нарушение правил использования степени сравнения прилагательного, лексической валентности слов, грамматической

сочетаемости при использовании разговорных слов и жаргонизмов, правил употребления предлогов перед согласными; наличие значительного (неоправданнго или избыточного) количества вводных слов и конструкций [5; 17; 24; 25; 29; 30; 35]. В тексте, сгенерированном нейросетью, отсутствует выражение собственного мнения, субъективные высказывания (авторская позиция) [7].

В сгенерированных текстах слова, включенные в заголовок, встречаются по тексту значительно чаще, чем в текстах, написанных человеком [36]. В тексте, сгенерированном нейросетью, очень редко встречаются опечатки и орфографические ошибки, не встречаются разные типы тире (-, — и —) или кавычек (“” и «») [25; 37–39].

Заключение

На основе проведенного анализа научной литературы можно заключить, что практика применения Gen-AI при написании научных текстов намного опережает рефлексию ее результатов и прогнозирование последствий в долгосрочной перспективе. Вероятно, в ближайшие годы этот разрыв сократится[5], так как в силу недостаточности обучающего материала генеративные нейронные сети придут к морфологически более стройному и стабильному набору характеристик (например, стереотипных ошибок на разных уровнях языка) результата генерации, который позволит разработать автоматические системы — детекторы генерации текста. До тех пор научный редактор / руководитель / рецензент вынуждены подкреплять интуитивное обнаружение сгенерированного текста разношерстными данными из научной литературы и из личных проб работы с системами-детекторами. Наряду с этим обобщение результатов экспертной оценки позволяет сформулировать следующие выводы.

1. Необходимо сочетать автоматические и ручные способы подтверждения наличия в рукописи сгенерированного текста. Следует использовать параллельно несколько систем автоматической проверки, так как каждая из них «натренирована» на материале конкретных Gen-AI-моделей и на специфическом языковом материале (например, на массиве текстов, написанных в публицистическом стиле). При ручной проверке можно опереться на предложенный в этой статье подход поуровневого анализа ошибок и погрешностей, обнаруживаемых в тексте.

2. Проверку (особенно автоматическую) следует проводить автономно относительно каждого раздела рукописи (для статей это аннотация, введение, обзор литературы / история вопроса, описание методики исследования, описание результатов исследования, обсуждение, выводы / заключение, список литературы). Запрещать использование инструментов Gen-AI, вероятно, непродуктивно, но обозначить границы приемлемого использования необходимо. По нашему мнению, недопустимо передавать на откуп пока еще

[5] См., например: Комерсантъ. URL: https://www.kommersant.ru/doc/6849411 (дата обращения: 05.12.2025).

«сырой» технологии задачу по аналитике практического материала, описанию и обобщению полученных данных. Особенно значимо требование самостоятельности выполнения научной работы и ее описания для начинающих исследователей, которые находятся в процессе освоения профессиональных навыков, погружения в предметную область исследования, становления исследовательской позиции, выработки индивидуального стиля письма.

3. При использовании AI-инструментов на любом этапе подготовки и написания научного текста для авторов должно стать правилом декларирование того, какие именно инструменты, в каком объеме и для генерации какой конкретно части текста применялись.

Список литературы

1. *Лебедев А.Н.* Константа М.Н. Ливанова и психофизиологические закономерности работы мозга // Психологический журнал. 2008. Т. 29. № 1. С. 133–137. EDN: INMIWB
2. *Кузьминов Я.И., Кручинская Е.В., Груздев И.А., Наумов А.А.* Отстающие и опережающие: как студенты используют генеративный искусственный интеллект в образовательных целях // Высшее образование в России. 2025. Т. 34. № 6. С. 9–35. https://doi.org/10.31992/0869-3617-2025-34-6-9-35 EDN: RXDTXQ
3. *Vaswani A., Shazeer N., Parmar N., Uszkoreit J., Jones L., Gomez A.N., Kaiser L., Polosukhin I.* Attention is all you need // Advances in Neural Information Processing Systems 30. Long Beach : Neural Information Processing Systems Foundation. New Orleans : IBM Research, 2017. P. 6000–6010.
4. *Yu H., Guo Y.* Generative artificial intelligence empowers educational reform: current status, issues, and prospects // Frontiers in Education. 2023. Vol. 11. Iss. 7. https://doi.org/10.31992/0869-3617-2025-34-6-9-35 EDN: RXDTXQ
5. *Телицына А.Ю.* Оптимизация научной деятельности через интеграцию ИИ: нейронные сети как инструмент в работе с академической литературой // Мониторинг общественного мнения: экономические и социальные перемены. 2024. № 5. С. 218–236. https://doi.org/10.14515/monitoring.2024.5.2623 EDN: FEHOZO
6. *Голубинский А.Н., Толстых А.А., Толстых М.Ю.* Автоматическая генерация аннотаций научных статей на основе больших языковых моделей // Informatics and Automation. 2025. Т. 24. №. 1. С. 275–301. https://doi.org/10.15622/ia.24.1.10 EDN: ICXSIC
7. *Осипенко Л.Е., Коротков А.В.* Текстовые генеративные нейросети в исследовательской деятельности студентов // Мир науки, культуры, образования. 2024. № 4 (107). С. 90–93. https://doi.org/10.24412/1991-5497-2024-4107-90-93 EDN: COOBKW
8. *Гаврилова Т.А.* «Стохастический попугай» на службе бизнеса: успехи и проблемы генеративного искусственного интеллекта // Российский журнал менеджмента. 2024. Т. 2. № 3. С. 461–482. https://doi.org/10.21638/spbu18.2024.305 EDN: VZIVZB
9. *Грицай Г.М., Грабовой А.В., Кильдяков А.С., Чехович Ю.В.* Поиск искусственно сгенерированных текстовых фрагментов в научных документах // Доклады Российской академии наук. Математика, информатика, процессы управления. 2023. Т. 514. № 2. С. 308–317. https://doi.org/10.31857/S2686954323601677 EDN: GQRWLF
10. *Зашихина И.М.* Научные публикации и большие языковые модели: поймет ли нейросеть российскую науку? // Научный редактор и издатель. 2024. Т. 9. № 52. С. 31–46. https://doi.org/10.24069/SEP-24-11 EDN: KOZFIO
11. *Павлов А.А.* Созданные с использованием искусственного интеллекта «научные тексты»: текущее состояние и перспективы развития // Методологические проблемы цивилистических исследований. 2024. № 6. С. 111–127. https//doi.org/10.33397/2619-0559-2024-6-6-111-127 EDN: LNFESF

12. *Гусаренко С.В., Гусаренко М.К*. Функционально-семантические, макроструктурные и пресуппозиционально-прагматические параметры сгенерированных русскоязычных текстов в лингвистических нейросетях GigaChat, ChatGPT Марти и Яндекс Алиса // Гуманитарные и юридические исследования. 2024. Т. 11. № 4. С. 788–800. https://doi.org/10.37493/24091030.2024.4.23 EDN: JXCRGC
13. *Фомин В.В., Фомина И.К. Тарануха С.Н*. Анализ системы автоматической проверки текстовых документов на наличие заимствований при использовании обучающимися инструментов генеративного искусственного интеллекта // Автоматизация и управление технологическими процессами и производствами. 2024. № 3 (63). С. 57–67.
14. *Николаев В.В., Рахконен М.Е*. Применение различных инструментов и использование чат-бота «ChatGpt» при написании научных работ, проверяемых в программе «Антиплагиат» // Профессиональное юридическое образование и наука. 2023. № 1 (9). С. 78–81. EDN: RJDDBY
15. *Liu Y., Zhang Z., Zhang W., Yue S., Zhao X., Cheng X., Zhang Y., Hu H*. ArguGPT: evaluating, understanding and identifying argumentative essays generated by GPT models // arXiv. 2023. https://doi.org/10.48550/arXiv.2304.07666
16. *Ma Y., Liu J., Yi F., Cheng Q., Huang Y., Lu W., Liu X*. AI vs. human – differentiation analysis of scientific content generation // Computational Linguistics. 2023. Vol. 1. Iss. 1. https://doi.org/arxiv.org/abs/2301.10416
17. *Огорелков И.В., Махаева Е.В*. Искусственная нейросетевая генерация текста как новый вид маскировки письменной речи автора // Виртуальная коммуникация и социальные сети. 2025. Т. 4. № 2. С. 163–171. https://doi.org/10.21603/2782-4799-2025-4-2-163-171 EDN: ENIMQP
18. *Гусаренко С.В*. Дефекты когнитивно-семантических структур как причина высокой энтропии актуального дискурса // Известия Южного федерального университета. Филологические науки. 2009. № 2. С. 68–74. EDN: KVLHWD
19. *Gritsay G., Grabovoy A., Chekhovich Yu*. Automatic detection of machine generated texts: need more tokens // Ivannikov Memorial Workshop (IVMEM). Moscow, 2022. P. 20–26. https://doi.org/10.1109/IVMEM57067.2022.9983964
20. *Ye J., Chen X., Xu N., Zu C., Shao Z., Liu S., Cui Y., Zhou Z., Gong C., Shen Y., Zhou J., Chen S., Gui T., Zhang Q., Huang X*. A comprehensive capability analysis of GPT-3 and GPT-3.5 Series Models // arXiv. 2023. https://doi.org/10.48550/arXiv.2405.10913
21. *Rodriguez J.D., Hay T., Gros D., Shamsi Z., Srinivasan R*. Cross-domain detection of GPT-2-generated technical text // Proceedings of the 2022 Conference of the North American Chapter of the Association for Computational Linguistics: Human Language Technologies. Association for Computational Linguistics. Seattle : Washington. 2022. P. 1213–1233. https://doi.org/10.18653/v1/2022.naacl-main.88
22. *Herbold S., Hautli-Janisz A., Heuer U., Kikteva Z., Trautsch A*. A large-scale comparison of human-written versus ChatGPT-generated essays // Scientific Reports. 2023. Vol. 13. Iss. 1. P. 18617–18627. https://doi.org/10.1038/s41598-023-45644-9 EDN: MFECWA
23. *Прохоров А.И., Асадчая К.В*. Инструментальные средства определения текста, сгенерированного при помощи нейросети // Научный вектор: сборник научных трудов / под ред. Е.Н. Макаренко. Т. 9. Ростов-на-Дону : Ростовский государственный экономический университет «РИНХ», 2023. С. 250–253. EDN: MGTGIE
24. *Хохлова М.В., Корышев М.В*. Сравнительный анализ сгенерированных и оригинальных аннотаций научных статей по филологической тематике // Научно-технический вестник информационных технологий, механики и оптики. 2024. Т. 24. № 6. С. 999–1006. https://doi.org/10.17586/2226-1494-2024-24-6-999-1006 EDN: KDMIJI
25. *Кулешов С.В., Зайцева А.А., Аксенов А.Ю*. Анализ статистических характеристик искусственно сгенерированных текстов // Известия высших учебных заведений. Приборостроение. 2024. Т. 67. № 11. С. 958–968. https://doi.org/10.17586/0021-3454-2024-67-11-958-968 EDN: HIYQAL

26. *Zhao B., Zhang S., Xu C., Sun Y., Deng C*. Deep fake geography? When geospatial data encounter artificial intelligence // Cartography and Geographic Information Science. 2021. Vol. 48. Iss. 4. P. 338–352. https://doi.org/10.1080/15230406.2021.1910075 EDN: MBHDAY
27. *Misra D.P., Chandwar K*. ChatGPT, artificial intelligence and scientific writing: what authors, peer reviewers and editors should know // Journal of the Royal College of Physicians of Edinburgh. 2023. Vol. 53. Iss. 2. P. 90–93. https://doi.org/10.1177/14782715231181023 EDN: DGPSHF
28. *Kacena M.A., Plotkin L.I., Fehrenbacher J.C*. The use of artificial intelligence in writing scientific review articles // Current Osteoporosis Reports. 2024. Vol. 22. P. 115–121. https://doi.org/10.1007/s11914023008520 EDN: PDAEEH
29. *Айдагулова А.Р*. Особенности текстов, сгенерированных искусственным интеллектом // Вестник Башкирского государственного педагогического университета им. М. Акмуллы. 2023. № 4 (72). С. 154–156. EDN: ROEFNK
30. *Козловский А.В., Мельник Я.Э., Волощук В.И*. О подходе для автоматической генерации сюжетно-связного текста // Известия ТулГУ. Технические науки. 2022. Вып. 9. С. 160–167. https://doi.org/10.24412/2071-6168-2022-9-160-168 EDN: DIRFGA
31. *Черкасова М.Н., Тактарова А.В*. Признаки сгенерированного текста в академическом дискурсе: проблема идентификации // Филологические науки. Вопросы теории и практики. 2024. Т. 17. № 7. С. 2226–2232. https://doi.org/10.30853/phil20240307 EDN: BEJDOU
32. *Kolmogorova A.V., Margolina A.V*. Written vs generated text: “naturalness” as a textual and psycholinguistic category // Научный результат. Вопросы теоретической и прикладной лингвистики. 2024. Т. 10. № 2. С. 71–99. https://doi.org/10.18413/2313-8912-2024-10-2-0-4 EDN: CCESHE
33. *Aydın Ö., Karaarslan E*. OpenAI ChatGPT generated literature review: digital twin in healthcare // Emerging Computer Technologies 2. İzmir : İzmir Akademi Dernegi, 2022. P. 22—31. https://doi. org/10.2139/ssrn.4308687
34. *Mikhaylovskiy N., Churilov I*. Autocorrelations decay in texts and applicability limits of language models // Computational Linguistics and Intellectual Technologies: Papers from the Annual International Conference “Dialogue”. 2023. Vol. 22. P. 350–360. https://doi.org/10.48550/arXiv.2305.06615
35. *Осетрова Е.В., Седова А.В*. Характеристики сгенерированного текста: языковой и социально-коммуникативный анализ // Сибирский филологический форум. 2025. Т. 2. № 31. С. 45–55. https://doi.org/10.24412/2587-7844-2025-2-45-55 EDN: WYGEFZ
36. *Тельпов Р.Е., Ларцина С.В*. Типовые различия естественных и сгенерированных нейронной сетью текстов в квантитативном аспекте // Научный диалог. 2023. Т. 12. № 7. С. 47–65. https://doi.org/10.24224/2227-1295-2023-12-7 47-65 EDN: ZZRRKN
37. *Chaka C*. Detecting AI content in responses generated by ChatGPT, YouChat, and Chatsonic: the case of five AI content detection tools // Journal of Applied Learning and Teaching. 2023. Vol. 6. Iss. 2. P. 94–104. https://doi.org/10.37074/jalt.2023.6.2.12 EDN: BJHFHQ
38. *Weber-Wulff D., Anohina-Naumeca A., Bjelobaba S., Foltýnek T., Guerrero-Dib J., Popoola O., Šigut P., Waddington L*. Testing of detection tools for AI-generated text // International Journal for Educational Integration. 2023. Vol. 19. P. 26. https://doi.org/10.1007/s40979-023-00146-z EDN: VCYINN
39. *Yang X., Cheng W., Petzold L., Wang W.Y., Chen H*. DNA-GPT: divergent n-gram analysis for training-free detection of GPT-generated Text // ArXiv. 2023. https://doi.org/10.48550/arXiv.2305.17359

References

1. Lebedev, A.N., & Konstanta, M.N. (2008). Livanov and Psychophysiological Patterns of Brain Function. *Psychological Journal*, *29*(1), 133–137. (In Russ.) EDN: INMIWB

2. Kuzminov, Ya.I., Kruchinskaia, E.V., Gruzdev, I.A., & Naumov, A.A. (2025). Falling Behind and Getting Ahead: Student Use of Generative AI in Education. *Higher Education in Russia*, *34*(6), 9–35. (In Russ.). https://doi.org/10.31992/0869-3617-2025-34-6-9-35 EDN: RXDTXQ
3. Vaswani, A., Shazeer, N., Parmar, N., Uszkoreit, J., Jones, L., Gomez, A.N., Kaiser, Ł., & Polosukhin, I. (2017). Attention is All You Need. In: Guyon, I., et al. (eds.). *Advances in Neural Information Processing Systems 30* (pp. 6000–6010). Long Beach: Neural Information Processing Systems Foundation.
4. Yu, H., & Guo, Y. (2023). Generative Artificial Intelligence Empowers Educational Reform: Current Status, Issues, and Prospects. *Frontiers in Education*, *11*(7). https://doi.org/10.31992/0869-3617-2025-34-6-9-35 EDN: RXDTXQ
5. Telitsyna, A.Yu. (2024). Optimization of Scientific Activities Through AI Integration: Neural Networks as a Tool in Working with Academic Literature. *Monitoring of Public Opinion: Economic and Social Changes*, 5, 218–236. (In Russ.). https://doi.org/10.14515/monitoring.2024.5.2623 EDN: FEHOZO
6. Golubinskiy, A., Tolstykh, A., & Tolstykh, M. (2025). Automatic Generation of Scientific Articles Abstracts Based on Large Language Models. *Informatics and Automation*, *24*(1), 275–301. (In Russ.). https://doi.org/10.15622/ia.24.1.10
7. Osipenko, L.E., & Korotkov, A.V. (2024). Text Generative Neural Networks in Students' Research Activities. *Mir Nauki, Kultury, Obrazovaniya*, 4 (107), 90–93. (In Russ.). https://doi.org/10.24412/1991-5497-2024-4107-90-93 EDN: COOBKW
8. Gavrilova, T.A. (2024). "Stochastic parrot" as the servant of Business: Achievements and Challenges of Generative Artificial Intelligence. *Russian Management Journal*, *2*(3), 461–482. (In Russ.). https://doi.org/10.21638/spbu18.2024.305 EDN: VZIVZB
9. Gritsay, G.M., Grabovoy, A.V., Kildyakov, A.S., & Checkovich, Yu.V. (2023). Artificially Generated Text Fragments Search in Academic Documents. *Doklady Rossijskoj Akademii Nauk. Matematika, Informatika, Processy Upravleniya*, *514*(2), 308–317. (In Russ.). https://doi.org/10.31857/S2686954323601677 EDN: GQRWLF
10. Zashikhina, I.M. (2024). Scientific Publications and Large Language Models: will Neural Networks Understand Russian Science? *Science Editor and Publisher*, *9*(52), 31–46. (In Russ.). https://doi.org/10.24069/SEP-24-11 EDN: KOZFIO
11. Pavlov, A.A. (2024). Artificially Generated "Scientific" Texts: Current Status and Future Perspectives. *Methodological Problems of the Civil Law Researches*, 6, 111–127. (In Russ.). https//doi.org/10.33397/2619-0559-2024-6-6-111-127 EDN: LNFESF
12. Gusarenko, S.V., & Gusarenko, M.K. (2024). Functionalsemantic, Macrostructural and Presuppositional-Pragmatic Parameters of Generated Russian-Language Texts in the Linguistic Neural Networks Gigachat, Chatgpt Marty and Yandex Alice. *Humanities and Law Research*, *11*(4), 788–800. (In Russ.). https://doi.org/10.37493/24091030.2024.4.23 EDN: JXCRGC
13. Fomin, V.V., Fomina, I.K., & Taranukha, S.N. (2024). Analysis of the System of Automatic Verification of Text Documents for the Presence of Borrowings when Students Use Generative Artificial Intelligence Tools. *Informatization of Education and Science*, 3 (63), 57–67. (In Russ.).
14. Nikolaev, V.V., & Rakhkonen, M.E. (2023). The Use of Various Tools and the Use of the Chatgpt Chatbot When Writing Scientific Papers Tested in the Antiplagiate Program. *Professional Legal Education and Science*, 1(9), 78–81 (In Russ.).
15. Liu, Y., Zhang, Z., Zhang, W., Yue, S., Zhao, X., Cheng, X., Zhang, Y., & Hu, H. (2023). ArguGPT: Evaluating, Understanding and Identifying Argumentative Essays Generated by GPT Models. *arXiv*. https://doi.org/10.48550/arXiv.2304.07666
16. Ma, Y., Liu, J., Yi, F., Cheng, Q., Huang, Y., Lu, W., & Liu, X. (2023). AI vs. Human – Differentiation Analysis of Scientific Content Generation. *arXiv*, *1*(1). https://doi.org/arxiv.org/abs/2301.10416

17. Ogorelkov, I.V., & Makhaeva, E.V. (2025). Artificial Neural Network Text Generation as a New Type of Masking the Authorship. *Virtual Communication and Social Networks*, *4*(2), 163–171. (In Russ.). https://doi.org/10.21603/2782-4799-2025-4-2-163-171 EDN: ENIMQP
18. Gusarenko, S.V. (2009). Defects in Cognitive-Semantic Structures as a Cause of High Entropy in Contemporary Discourse. *Proceedings of Southern Federal University. Philology*, 2, 68–74. (In Russ.). EDN: KVLHWD
19. Gritsay, G., Grabovoy, A., & Chekhovich, Yu. (2022). Automatic Detection of Machine Generated Texts: Need More Tokens. In: *Ivannikov Memorial Workshop (IVMEM)* (pp. 20–26). Moscow. https://doi.org/10.1109/IVMEM57067.2022.9983964
20. Ye, J., Chen, X., Xu, N., Zu, C., Shao, Z., Liu, S., Cui, Y., Zhou, Z., Gong, C., Shen, Y., Zhou, J., Chen, S., Gui, T., Zhang, Q., & Huang, X. (2023). A Comprehensive Capability Analysis of GPT-3 and GPT-3.5 Series Models. *arXiv*. https://doi.org/10.48550/arXiv.2405.10913
21. Rodriguez, J.D., Hay, T., Gros, D., Shamsi, Z., & Srinivasan, R. (2022). Cross-Domain Detection of GPT-2-Generated Technical Text. In: *Proceedings of the 2022 Conference of the North American Chapter of the Association for Computational Linguistics: Human Language Technologies* (pp. 1213–1233). Association for Computational Linguistics. https://doi.org/10.18653/v1/2022.naacl-main.88
22. Herbold, S., Hautli-Janisz, A., Heuer, U., Kikteva, Z., & Trautsch, A. (2023). A Large-Scale Comparison of Human-Written Versus ChatGPT-Generated Essays. *Scientific Reports*, *13*(1), 18617–18627. https://doi.org/10.1038/s41598-023-45644-9
23. Prokhorov, A.I., & Asadchaya, K.V. (2023). Tools for Determining the Text Generated Using a Neural Network. *Nauchny Vektor*, *9*, 250–253. (In Russ.). EDN: MGTGIE
24. Khokhlova, M.V., & Koryshev, M.V. (2024). Comparative Analysis of AI-Generated and Original Abstracts of Academic Articles on Philology. *Scientific and Technical Journal of Information Technologies, Mechanics and Optics*, *24*(6), 999–1006. (In Russ.). https://doi.org/10.17586/2226-1494-2024-24-6-999-1006 EDN: KDMIJI
25. Kuleshov, S.V., Zaytseva, A.A., & Aksenov, A.Yu. (2024). Analysis of Statistical Characteristics of Artificially Generated Texts. *Journal of Instrument Engineering*, *67*(11), 958–968. (In Russ.). https://doi.org/10.17586/0021-3454-2024-67-11-958-968 EDN: HIYQAL
26. Zhao, B., Zhang, S., Xu, C., Sun, Y., & Deng, C. (2021). Deep Fake Geography? When Geospatial Data Encounter Artificial Intelligence. *Cartography and Geographic Information Science*, *48*(4), 338–352. https://doi.org/10.1080/15230406.2021.1910075 EDN: MBHDAY
27. Misra, D.P., & Chandwar, K. (2023). ChatGPT, Artificial Intelligence and Scientific Writing: What Authors, Peer Reviewers and Editors Should Know. *Journal of the Royal College of Physicians of Edinburgh*, *53*(2), 90–93. https://doi.org/10.1177/14782715231181023
28. Kacena, M.A., Plotkin, L.I., & Fehrenbacher, J.C. (2024). The Use of Artificial Intelligence in Writing Scientific Review Articles. *Current Osteoporosis Reports*, *22*, 115–121. https://doi.org/10.1007/s11914023008520 EDN: PDAEEH
29. Aidagulova, A.R. (2023). Features of Texts Generated by Artificial Intelligence. *Vestnik Bashkirskogo Gosudarstvennogo Pedagogicheskogo Universiteta imeni M. Akmully*, 4, 154–156. (In Russ.). EDN: ROEFNK
30. Kozlovsky, A.V., Melnik, Ya.E., & Voloshchuk, V.I. (2022). On The Approach for Automatic Generation of Narrative-Linked Text. *Izvestiya TulGU. Tekhnicheskie Nauki*, 9, 160–167. (In Russ.). https://doi.org/10.24412/2071-6168-2022-9-160-168 EDN: DIRFGA
31. Cherkasova, M.N., & Taktarova, A.V. (2024). Attributes Of Generated Text in Academic Discourse: The Problem of Identification. *Philology. Theory & Practice*, *17*(7), 2226–2232. (In Russ.). https://doi.org/10.30853/phil20240307 EDN: BEJDOU
32. Kolmogorova, A.V., & Margolina, A.V. (2024). Written vs Generated Text: “Naturalness” as a Textual and Psycholinguistic Category. *Research Result. Theoretical and Applied Linguistics*, *10*(2), 71–99. https://doi.org/10.18413/2313-8912-2024-10-2-0-4 EDN: CCESHE

33. Aydın, Ö., & Karaarslan, E. (2022). OpenAI ChatGPT Generated Literature Review: Digital Twin in Healthcare. In: *Emerging Computer Technologies 2* (pp. 22–31). İzmir: İzmir Akademi Dernegi. https://doi. org/10.2139/ssrn.4308687
34. Mikhaylovskiy, N., & Churilov, I. (2023). Autocorrelations Decay in Texts and Applicability Limits of Language Models. In: *Computational Linguistics and Intellectual Technologies: Papers from the Annual International Conference "Dialogue"* (Vol. 22, pp. 350–360). https://doi.org/10.48550/arXiv.2305.06615
35. Osetrova, EV., & Sedova, A.V. (2025). Characteristics of the generated Text: Linguistic and Social-Communicative Analysis. *Siberian Philological Forum*, *2*(31), 45–55. (In Russ.). https://doi.org/10.24412/2587-7844-2025-2-45-55 EDN: WYGEFZ
36. Telpov, R.E., & Lartsina, S.V. (2023). Typological Differences of Natural and Neural Network-Generated Texts in a Quantitative Aspect. *Nauchnyi dialog*, *12*(7), 47–65. (In Russ.). https://doi.org/10.24224/2227-1295-2023-12-7 47-65 EDN: ZZRRKN
37. Chaka, C. (2023). Detecting AI Content in Responses Generated by ChatGPT, YouChat, and Chatsonic: The Case of Five AI Content Detection Tools. *Journal of Applied Learning and Teaching*, *6*(2), 94–104. DOI: 10.37074/jalt.2023.6.2.12
38. Weber-Wulff, D., Anohina-Naumeca, A., Bjelobaba, S., Foltýnek, T., Guerrero-Dib, J., Popoola, O., Šigut, P., & Waddington, L. (2023). Testing of Detection Tools for AI-Generated Text. *International Journal for Educational Integration*, *19*, 26. https://doi.org/10.1007/s40979-023-00146-z EDN: VCYINN
39. Yang, X., Cheng, W., Petzold, L., Wang, W.Y., & Chen, H. (2023). DNA-GPT: divergent N-Gram Analysis for Training-Free Detection of GPT-Generated Text. *ArXiv*. https://doi.org/10.48550/arXiv.2305.17359

Сведения об авторе:
Комалова Лилия Ряшитовна, доктор филологических наук, доцент, руководитель и ведущий научный сотрудник Центра эмерджентных практик, Институт научной информации по общественным наукам Российской академии наук (117418, Российская Федерация, г. Москва, Нахимовский проспект, д. 51/21); профессор кафедры прикладной и экспериментальной лингвистики, Московский государственный лингвистический университет (119034, Российская Федерация, г. Москва, ул. Остоженка, д. 38 с. 1); *сфера научных интересов*: корпусная лингвистика, экспериментальная лингвистика, речевая деятельность, кинодискурс; *e-mail*: komalova@inion.ru
ORCID: 0000-0002-0955-5315, Scopus Author ID: 55873286400, ResearcherID: O-2238-2019, SPIN-код: 9330-0170

Information about the author:
Liliya R. Komalova, Dr.Sc. (Philology), Associate Professor, Leading Research Fellow and Head of the Centre for Emerging Practices, Institute of Scientific Information for Social Sciences of the Russian Academy of Sciences (51/21 Nakhimovsky Prospekt, Moscow, Russian Federation, 117418); Professor at the Department of Applied and Experimental Linguistics, Moscow State Linguistic University (38 build 1 Ostozhenka Str., Moscow, Russian Federation, 119034); *Research interests*: corpus linguistics, experimental linguistics, speech activity, cinematic discourse; *e-mail*: komalova@inion.ru
ORCID: 0000-0002-0955-5315, Scopus Author ID: 55873286400, ResearcherID: O-2238-2019, SPIN-code: 9330-0170

RUDN Journal of Language Studies, Semiotics and Semantics
2313-2299 (Print), ISSN 2411-1236 (Online)
Вестник Российского университета дружбы народов.
Серия: ТЕОРИЯ ЯЗЫКА. СЕМИОТИКА. СЕМАНТИКА

2026 Vol. 17 No. 2 442–458
http://journals.rudn.ru/semiotics-semantics

DOI: 10.22363/2313-2299-2026-17-2-442-458
EDN: LJBERK

Научная статья / Research Article

Промт-инжиниринг как лингвистическая проблема

О.И. Максименко[1] ✉, **М.В. Беляков**[2]

[1]Государственный университет просвещения, *Москва, Российская Федерация*
[2]Московский государственный институт международных отношений (университет) МИД России, *Москва, Российская Федерация*
✉maxbel7@yandex.ru

Аннотация. Промпт-инжиниринг — феномен, возникший в результате взаимодействия человека с большими языковыми моделями (LLM). В работе обосновывается тезис, что промпт представляет собой качественно новый тип управляющего текста, конструируемого на естественном языке. Цель исследования — продемонстрировать, что анализ промптов требует привлечения всего арсенала современной лингвистики: семантики, прагматики, дискурс-анализа, когнитивной лингвистики и философии языка. Последовательно рассмотрены семантические особенности промптов с указанием на необходимость обработки размытых категорий и контроля референции в условиях отсутствия у модели человеческого опыта. Прагматический анализ выявляет специфику промпта как директивного речевого акта, где условия успешности должны быть эксплицированы вербально. В рамках дискурсивного подхода демонстрируется, как промпт задает макроструктуру будущего текста, конструирует профессиональные роли и социальные регистры. Особое внимание уделяется когнитивным основаниям промпт-инжиниринга, где формулировка запроса выступает как внешняя метакогнитивная практика, структурирующая человеческое мышление; межъязыковым аспектам, связанным с доминированием англоязычных корпусов, и специфике работы с русским языком. Делается вывод, что промпт-инжиниринг знаменует глубокую трансформацию языковой деятельности и требует формирования новой междисциплинарной области — лингвистики коммуникации «человек — модель».

Ключевые слова: большие языковые модели (LLM), семантика, прагматика, дискурс, управляющий текст, модель

Вклад авторов: равнозначен на все этапах исследования и подготовки статьи к публикации.

Конфликт интересов: авторы заявили об отсутствии конфликтов интересов.

История статьи: дата поступления: 20.01.2026; дата приема в печать: 10.02.2026.

Для цитирования: *Максименко О.И.*, *Беляков М.В.* Промт-инжиниринг как лингвистическая проблема // Вестник Российского университета дружбы народов. Серия: Теория языка. Семиотика. Семантика. 2026. Т. 17. № 2. С. 442–458. https://doi.org/10.22363/2313-2299-2026-17-2-442-458 EDN LJBERK

© Максименко О.И., Беляков М.В., 2026
This work is licensed under a Creative Commons Attribution-NonCommercial 4.0 International License
https://creativecommons.org/licenses/by-nc/4.0/legalcode

Promt-Engineering as a Linguistic Problem

Olga I. Maksimenko[1], Mikhail V. Belyakov[2]

[1]Federal State University of Education, *Moscow, Russian Federation*
[2]Moscow State Institute of International Relations (University) Ministry of Foreign Affairs of the Russian Federation, *Moscow, Russian Federation*
✉maxbel7@yandex.ru

Abstract. Prompt engineering, a phenomenon arising from human interaction with large language models (LLMs), is the subject of this linguistic analysis. The study argues that a prompt represents a qualitatively new type of controlling text constructed in natural language. The research aims to demonstrate that the analysis of prompts requires the full range of modern linguistic tools: semantics, pragmatics, discourse analysis, cognitive linguistics, and the philosophy of language. The article examines the semantic features of prompts, highlighting the need to handle fuzzy categories and manage reference given the model's lack of human experience. Pragmatic analysis reveals the specific nature of the prompt as a directive speech act, in which felicity conditions must be explicitly verbalized. Within a discursive approach, it is demonstrated how a prompt establishes the macrostructure of the future text and constructs professional roles and social registers. Special attention is paid to the cognitive foundations of prompt engineering, where the formulation of a request serves as an external metacognitive practice that structures human thinking. The article also addresses cross-linguistic aspects related to the dominance of English-language corpora and the specifics of working with the Russian language. It concludes that prompt engineering marks a profound transformation of linguistic activity and necessitates the formation of a new interdisciplinary field: the linguistics of human–model communication.

Keywords: large language models (LLMs), semantics, pragmatics, discourse, controlling text, model

Authors' contribution: The authors contributed equally to this article and preparation of the article text.

Conflicts of interest: The authors stated that there were no conflicts of interest.

Article history: Received: 20.01.2026; Accepted: 10.02.2026.

For citation: Maksimenko, O.I., & Belyakov, M.V. (2026). Promt-Engineering as a Linguistic Problem. *RUDN Journal of Language Studies, Semiotics and Semantics, 17*(2), 442–458. (In Russ.). https://doi.org/10.22363/2313-2299-2026-17-2-442-458 EDN LJBERK

Введение

Появление больших языковых моделей (Large Language Models, LLM) радикально изменило представление о том, как человек взаимодействует с вычислительными системами. Если ранее коммуникация с машиной осуществлялась преимущественно через формальные языки программирования, интерфейсы командной строки и строго структурированные формы ввода информационно-поисковых языков [1], то сегодня ключевым посредником становится естественный язык. Пользователь формулирует задачу в привычной для него вербальной форме, а модель генерирует ответ, создавая иллюзию «понимания» и диалога. Однако эта иллюзия быстро разрушается при практическом использовании. Модели не обладают опытом, интенциональностью, телесностью, культурно-историческим контекстом и здравым смыслом в человеческом понимании; они опираются исключительно на статистические

закономерности в корпусах текстов. В результате качество ответа оказывается непосредственно зависимым от того, как именно сформулирована исходная инструкция — промпт. В русском языке данное слово появилось с развитием нейросетей как транслитерация английского слова *prompt* (подсказка, стимул), которое, в свою очередь, восходит к латинскому *prōmptus* (причастию от глагола *prōmere* (извлекать), которое нередко функционирует как прилагательное «готовый, явный, удобный, быстрый»). В обыденной терминологии промпт описывается как «умение правильно задавать запрос», но такая формулировка скрывает глубину проблемы.

Постановка проблемы

Промпт-инжиниринг в современных условиях представляет собой не набор эмпирических «подсказок», а новую форму управляющего текста, конструируемого на естественном языке. В отличие от традиционных письменных жанров промпт не только передаёт информацию, но и задаёт режим работы системы**,** одновременно описывая задачу, определяя роль модели (эксперт, преподаватель, юрист, редактор и пр.) и фиксируя желаемый жанр и стиль ответа. Кроме того, промпт содержит метатекстовые инструкции относительно формы будущего текста (структура, объём, формат, тон) и компенсирует отсутствие у модели реального контекста, явно формулируя то, что в человеческой коммуникации остаётся невыраженным: пресуппозиции, условия релевантности, критерии «хорошего» ответа. Тем самым промпт становится объектом, который невозможно адекватно описать только терминами ИТ или инженерии. Он оказывается в центре ряда философских и лингвистических проблем. Так, семантическая проблема заключается в том, как переопределяются значения и требования к точности формулировки в ситуации, когда текст интерпретируется не человеком, а статистической моделью. Прагматическая проблема связана с тем, как кодируются намерения, цели и условия успешности речевого акта в отсутствие общего опыта и контекста. Дискурсивная проблема — как промпт задает макроструктуру будущего текста, жанровые рамки и стратегию аргументации. Проблема социолингвистическая — как через промпт конструируются роли, регистры и стили, имитирующие профессиональные и социальные идентичности. Также присутствует проблема когнитивная — как формулировка промпта отражает внешнюю запись когнитивных стратегий и структурирует человеческое мышление.

Актуальность данной темы обусловлена несколькими обстоятельствами. Во-первых, промпт-инжиниринг уже сегодня имеет прикладное значение в образовании, юриспруденции, науке, бизнес-аналитике, технической документации и других областях, где языковые модели используются как инструмент поддержки решений и производства текста [2; 3]. Во-вторых, наблюдается явный дефицит теоретического осмысления промпта именно как лингвистического объекта: публицистические и технические описания (how-to, best practices) доминируют над системным научным анализом.

В-третьих, анализ промптов позволяет по-новому взглянуть на классические лингвистические категории — семантику, прагматику, жанр, дискурс, теорию референции — поскольку они оказываются буквально «вынутыми» на поверхность и формализованными в явных инструкциях.

Объектом исследования в данной работе является вербальная форма взаимодействия человека с языковой моделью, то есть структурные, семантические и прагматические характеристики текстов промптов. Предметом исследования выступает лингвистическая природа промпта как особого типа управляющего текста: его семантическая организация, прагматическая структура, дискурсивные функции и когнитивные предпосылки.

Цель исследования состоит в том, чтобы показать промпт-инжиниринг как лингвистическую проблему, требующую привлечения инструментов современной семантики, прагматики, дискурс-анализа, когнитивной лингвистики и философии языка, и тем самым обосновать необходимость формирования отдельной междисциплинарной области — лингвистики человеко-модельной коммуникации.

Для достижения этой цели в статье решаются следующие задачи:

1) описать промпт как особый тип текста, показать его гибридную жанровую природу и отличие от традиционных директивных и метатекстовых форм;

2) проанализировать семантические особенности промптов: требования к точности, операциональности и устранению размытых категорий;

3) рассмотреть прагматическую структуру промптов с опорой на теорию речевых актов и максимы кооперации;

4) исследовать дискурсивную организацию промпта: ролевые установки, жанровые прототипы, макроструктуру будущего текста;

5) показать когнитивные и философские основания промпт-инжиниринга (прототипы, фреймы, модели мира, уровни объяснения);

6) обсудить межъязыковые аспекты (англоязычная доминанта, специфика русского языка в промптах) и социолингвистические последствия распространения языковых моделей;

7) сформулировать предпосылки для выделения промпт-инжиниринга в рамках новой научной дисциплины.

Методологическую базу работы составляют идеи формальной семантики Р. Монтегю [4], теории прототипов Э. Рош [5], фреймовой семантики Ч. Филлмора [6] и М. Минского [7], теории речевых актов Дж. Остина [8] и Дж. Серля [9; 10], кооперативной прагматики Г. Грайса [11], философии языка Л. Витгенштейна [12], С. Крипке [13] и Д. Деннетта [14], дискурс-анализа Т. Ван Дейка [15], Э. Гофмана [16], М. Фуко [17].

Статья носит теоретико-аналитический характер и ориентирована на исследователей в области лингвистики, философии языка, когнитивных наук и разработчиков, работающих с языковыми моделями, а также на специалистов, которые фактически занимаются промпт-инжинирингом, но нуждаются в более глубоком и формализованном понимании его лингвистической природы.

Промпт как новый тип текста

Появление больших языковых моделей радикально переопределило механизмы текстовой коммуникации. В условиях взаимодействия «человек — модель» промпт занимает место, не соответствующее ни одному из традиционных жанров человеческой речи. Он сохраняет черты директивы, инструкции, запроса, алгоритма, описания задачи, пояснения и метатекста, но не совпадает ни с одним из них полностью. Промпт — это управляющий текст, создающий условия для функционирования модели и определяющий параметры будущего ответа.

Традиционные тексты так или иначе построены вокруг роли коммуникатора, обладающего опытом и пресуппозициями. Промпт, напротив, создаёт условия для симуляции опыта: модель, не имея реального знания мира, извлекает вероятностные паттерны из текста промпта. Следовательно, промпт выступает как механизм замещения контекста, который в естественной коммуникации является неотъемлемой частью взаимодействия собеседников.

Можно говорить о том, что промпту свойственна гибридность, которая проявляется в нескольких аспектах. Жанровая гибридность промпта связана с тем, что он совмещает аналитическую, директивную, описательную и нормативную функции. Функциональная гибридность определена тем, что он одновременно определяет задачу, инструктирует модель и ограничивает диапазон интерпретаций. Дискурсивная гибридность — промпт задает фрейм будущего текста, предопределяя его структуру и стиль. Также можно выделить метатекстовую гибридность**,** поскольку промпт содержит указания относительно языка, которым должен быть написан ответ.

Эта уникальность делает промпт объектом междисциплинарного анализа, соединяющим лингвистику, когнитивные науки, теорию дискурса и философию языка.

Семантические особенности промптов

Семантика промптов является областью, где наиболее явно проявляются ограничения языковых моделей. Для человека смысл высказывания часто выводится из контекста, имплицитных знаний, опыта, социальных и культурных норм. Модель же опирается исключительно на текст — на последовательность символов, являющуюся входом. Поэтому семантика промпта должна быть максимально эксплицитной, операциональной и точной. Наиболее критичны такие семантические аспекты, как размытие категории, полисемия, инструментализация семантики, контроль референции и формальная семантика.

К размытым категориям могут быть причислены определительные наречия образа действия, такие как *подробно, глубоко, кратко, ясно,* которые обладают градуальной природой (по Э. Рош [5]), но не имеют четких операциональных границ. Модель интерпретирует их как усреднённые категории, что приводит к неоднозначности.

Многозначность также присутствует в ряде таких частотных для промптов лексем, как *анализ, обзор, обоснование, описание* и др., обладающих высокой семантической вариативностью, в связи с чем необходима их конкретизация, например: *анализ* — выделить критерии, признаки, отношения; *обзор* — перечислить основные факты; *обоснование* — привести аргументы; *описание* — зафиксировать свойства объекта.

Инструментализация семантики подразумевает, что промпт должен превращать семантические категории в действия, которые модель может выполнить, например: *проанализируй* — классифицируй + оцени; *объясни* — укажи причинно-следственные связи; *сформулируй вывод* — обобщи ключевые положения.

Безусловно, необходим контроль референции — модель часто теряет конкретные объекты без чёткой референции [13]. Например: *проанализируй документ* — лишено референции, *проанализируй договор купли-продажи оборудования от 2025 года* — корректно.

С позиций формальной семантики [4] промпт задаёт не значение, а пространство интерпретаций. Если это пространство слишком широкое, модель принимает вероятностное среднее решение. Отсюда возникает необходимость жёсткой семантической спецификации. Таким образом, промпт должен обладать семантической плотностью — минимизировать неопределённость и расширять операциональность.

Проблемы значения: формальная семантика и интерпретация

Понимание текста моделью принципиально отличается от человеческого понимания. Человек соотносит слова с миром через опыт и перцептивные категории. Модель же соотносит текстовые единицы друг с другом через вероятностные отношения. Это фундаментальный разрыв, который становится особенно очевидным при анализе промптов.

Согласно формальной семантике Р. Монтегю [4], смысл выражения определяется через систему правил отображения формул в модели мира. У языковых моделей нет формальной модели мира; вместо неё используется огромная статистическая структура, вследствие чего модель не проверяет истинность, она проверяет правдоподобие. Следовательно, недоопределённые промпты приводят к «галлюцинациям».

Лингвистическая неопределенность промпта тесно увязана со многими факторами. Если выражение допускает несколько семантических трактовок, модель выбирает наиболее вероятную по корпусу, а не наиболее корректную по логике. Например, при запросе: *напиши научное объяснение* модель выбирает из корпуса усреднённый «научный стиль».

Как упоминалось выше, принципиальна роль контекста как внешней логической структуры. У человека контекст — это мир, ситуация, фоновые знания, у модели контекст — только текст, который находится в промпте. Значит, промпт должен создать контекст, иначе модель его «выдумывает». Если

промпт лишен операциональной структуры, модель повторяет наиболее вероятные шаблоны. Отсюда возникает необходимость формализованной семантики внутри промпта.

Одним из ключевых элементов, определяющих интерпретацию задачи моделью, является прагматика промптов. Человек свободно различает просьбу, указание, рекомендацию, требование через синтаксическую структуру предложения, интонацию, социальный контекст и культурные нормы. Модель лишена этих механизмов, она воспринимает только текст. В связи с этим необходимо обратиться к теории речевых актов Дж. Остина [8] и Дж. Сёрля [9]. Если опираться на данную теорию, то можно рассмотреть промпт как директивный речевой акт, содержащий иллокутивную цель, условия успешности, рамки выполнения и ожидаемую структуру ответа. Проблема состоит в том, что модель не умеет выводить иллокуцию из контекста, поэтому ее необходимо эксплицировать. Например: сформулировать промпт *проанализируй письмо* недостаточно, корректен в таком контексте следующий промпт: *проанализируй письмо как юрист, выделив: риски, формальные нарушения, обязательства сторон, возможные последствия*.

Прагматическими маркерами промпта могут быть следующие: цель («цель анализа…»), роль («ты — эксперт по…»), адресат («объясни студенту…»), регистровые требования («официальный стиль»). По Дж. Сёрлю [9], речевой акт требует соответствия ситуации, компетентности говорящего, понимания намерения. Промпт подменяет эти элементы текстовыми инструкциями, т.е. он фактически создаёт искусственные условия успешности.

Кооперативные принципы и ошибки коммуникации в промптах

Работа модели подчиняется аналогам известных максим Пола Грайса [11]: 1) максима количества — текст промпта должен содержать достаточно информации; 2) максима качества – если требуется точность, это должно быть прописано явно; 3) максима релевантности – модель плохо игнорирует нерелевантные фрагменты, перегруженный промпт ведет к «шуму»; 4) максима способа – структурированность и ясность представляют собой обязательное условие корректной интерпретации. Неполнота прагматических требований (максимы количества) приводит к предсказуемо неудовлетворительному результату, например, в запросе: *напиши эссе* отсутствует целевая аудитория, стиль, объем, формат и цель.

Понятие промпт-шума (отсутствие максимы релевантности) отчасти перекликается с понятием «стоп-слов» в информационном поиске. Тем не менее, говоря о промпт-шуме, в первую очередь необходимо упомянуть такие нежелательные компоненты промпта, как лишние детали, эмоциональные отступления, разговорные конструкции и противоречивые указания. Даже небольшие нарушения могут «переключить» модель на неправильный дискурс.

К типичным ошибкам, создающим противоречивые инструкции, относятся такие конструкции, как *сделай научно, но максимально просто; пиши*

строго, но с юмором; будь кратким, но очень подробным. Без четкой иерархии требований модель выбирает случайный приоритет.

Дискурсивная организация промптов

Дискурс промпта представляет собой совокупность риторических, структурных и прагматических механизмов, формирующих условия порождения будущего текста. В отличие от обычной речи, где дискурс определяется ситуацией общения, ролью участников и контекстом, в промпте дискурс задаётся искусственно — через явные инструкции. По теории Тео ван Дейка [15], любой текст строится вокруг макроструктуры — глобального смыслового каркаса. В естественном общении макроструктура выводится автоматически из социального контекста и жанра. В промпте же макроструктура должна быть явно задана, структурные сигналы — определены заранее, последовательность разделов часто воспроизводится буквально.

Например: *сначала определи проблему, затем выдели факторы, далее приведи анализ, сформулируй вывод*. Такие инструкции (пайплайны, цепочки последовательных действий) выполняют функцию скелета дискурса.

Кроме того, в успешном промпте необходимо активировать жанровые модели, такие как, научная статья, аналитический отчёт, эссе, юридическое заключение, пресс-релиз, официальное письмо и пр. Каждый жанр обладает набором структурных и лексических ожиданий, которые модель восстанавливает по описанию в промпте. Промпт должен задавать средства когезии (связность текста) и когерентности (логическая согласованность). Без таких указаний модель склонна к избыточной вариативности, повторениям и «дрейфу дискурса», нарушающему общую структуру.

Одной из уникальных особенностей промптов является способность задавать модели профессиональные, социальные и статусные роли. В реальной коммуникации регистр определяется социальным положением говорящего, ситуацией общения и культурными ожиданиями. В промпте регистр конструируется искусственно и жёстко. Задание «ролевого дискурса» или «ролевой маски» по И. Гофману [16] (*ты — юрист, ты — историк, ты — профессор* и пр.) автоматически активирует специфический словарь (профессиональную терминологию), соответствующие стратегии аргументации, структурные шаблоны, характерные для данной профессии, степень формальности, композиционные модели рассуждения, типичные речевые клише и формулы, ожидаемую этикетную норму. Профессиональная роль действует как «точка входа» в определённую область корпуса. Лингвистически это означает, что роль становится предиктором дискурсивных правил. Однако существует такое понятие, как «ошибка ролевой спецификации», возникающая тогда, когда промпт задает противоречивые роли: *пиши строго научно, но как маркетолог* или *как юрист объясни ребенку*. В такой ситуации модель не может разрешить понятийный конфликт, что приводит к дискурсивной деградации. Следовательно, задание роли должно быть единым и непротиворечивым.

Отдельно необходимо отметить важность формулировки социальной идентичности и регистра. Регистры могут быть формальными, нейтральными, разговорными, экспертными, академическими или бюрократическими. При этом каждый регистр состоит из специфической лексики, характерных синтаксем, композиционных схем и риторических стратегий. Чем точнее задан регистр, тем предсказуемее ответ модели. Кроме того, в сформированном социуме присутствует этикет и прагматические коммуникативные нормы. Модель не имеет таких норм, их надо задавать текстом, например: *избегай категоричного тона; не используй личных суждений; сохраняй нейтральную позицию.*

Когнитивные основания: прототипы, фреймы, модели мира

Языковые модели не обладают человеческим опытом или когнитивной архитектурой. Тем не менее, они демонстрируют текстовое поведение, напоминающее рассуждение. Это объясняется тем, что задача промпта — компенсировать отсутствие когнитивных структур через явные текстовые инструкции.

Человек классифицирует понятия через прототипы [5]: слово *птица* вызывает образ воробья, а не пингвина. У модели прототипы заменены на статистические центры — наиболее частотные структуры в обучающих данных. Промпт должен превращать градуальные категории в четкие инструкции, для этого все размытые категории усредняются; все неопределённые требования упрощаются и все градуальные признаки сглаживаются.

Кроме того, промпт должен конструировать фрейм текста. Фрейм, согласно теориям Ч. Филлмора [6] и М. Минского [7], — это схема ситуации (роли, отношения, действия, типичные последовательности). У модели нет встроенных фреймов: они восстанавливаются по корпусу или создаются промптом. Например: *проанализируй письмо клиента.* Модель не знает фрейма «деловая коммуникация», значит, в промпт необходимо ввести требуемые уточнения.

Языковая модель также не обладает знаниями о модели мира, у нее отсутствует интенциональность, т.е. у нее нет собственных целей, мотиваций, намерений и убеждений. Она имитирует рассуждение за счёт статистической корреляции. Функция промпта — явным образом задавать «модели мира», которые в человеческой коммуникации являются неявными.

Для экспликации когнитивных шагов создаются промпты типа: *думай шаг за шагом; раздели задачу на стадии; сначала гипотезы, затем проверка.* Они вводят метакогнитивную структуру, компенсируя отсутствие планирования в модели.

Философские основания языка в промпт-инжиниринге

Философия языка предоставляет ключевые инструменты для анализа природы промпта, особенно в аспекте значения, референции и правил

употребления. По Л. Витгенштейну [12], значение слова определяется его употреблением в языковой игре. Промпт, как языковая структура, задаёт жанровые, ролевые, аргументативные и стилистические правила игры. Если правила не заданы, это означает, что игра не определена, следовательно, модель генерирует вероятностное усреднение. С. Крипке [13], в свою очередь, обозначил существующую проблему референции, показав, что референция — это не описание, а устойчивая цепочка использования. Модель теряет референцию при недостаточной спецификации. Если опираться на точку зрения С. Крипке, можно предположить, что промпт должен задавать конкретные объекты, определять контекст и фиксировать область знания. Д. Деннетт [14], в свою очередь, ввел понятие *уровня объяснений* и предложил различать физический уровень, функциональный / уровень дизайна и интенциональный уровень. Модель работает на функциональном уровне, но промпт вынуждает «играть» на интенциональном. Это полезная функция, позволяющая человеку конструировать сложные задания на естественном языке. Отдельно необходимо учитывать, что модель не различает истину и правдоподобие. Поэтому промпты должны задавать требования достоверности, запрещать генерацию фактов при отсутствии данных и ограничивать область знания. Таким образом, философские основания подчёркивают, что промпт — это механизм искусственного воспроизводства смыслов, не являющихся результатом понимания.

Метатекст и метадискурс как ядро промпта

Промпт — единственный текстовый жанр, где метатекстовая функция не просто возможна, а обязательна, при этом можно говорить о так называемой «метатекстовой рефлексии», поскольку промпт содержит конкретные инструкции о структуре («*раздели на главы*»), о стиле («*формальный академический тон*») и, например, об ограничениях («*не используй эмотивы*»). Все эти признаки превращают промпт в самореферентный текст. Кроме того, промпт управляет будущим метадискурсом [18] текста модели, где метадискурс включает организационные маркеры («*во-первых*», «*в заключение*»), интеракционные маркеры («*следует подчеркнуть*») и оценочные конструкции («*важно отметить*»). Также промпт можно считать нормативным текстом, поскольку он задает правила, ограничивает интерпретацию и регулирует тон и жанр. Это делает промпт одновременно текстом и нормативной грамматикой для ответа модели.

Корпусные и межъязыковые аспекты

Большие языковые модели обучаются на объемных корпусах, преимущественно англоязычных, что приводит к явным лингвистическим и социокультурным сдвигам. Типологические особенности разноструктурных английского и русского языков также определяют структуру и семантику промпта.

Промпты на английском языке точнее интерпретируются, дают более устойчивые результаты и намного реже подвержены «галлюцинациям», поскольку английский язык морфологически прост и синтаксически стабилен. Кроме того, он представлен в огромных разнородных корпусах, что нельзя сказать о русском языке с развитой морфологией, свободным порядком слов, высоким уровнем полисемии и выраженными стилистическими различиями и корпусами меньших объемов. Все эти составляющие предъявляют особые требования к промпту на русском языке — пропмт должен быть более структурирован и конкретен, чем на английском. Кроме того, не раз отмечался межъязыковой перенос дискурсов, при котором модель нередко придерживается англосаксонского стиля аргументации и риторических клише, создавая «англифицированные» русские тексты. Во избежание подобного промпт должен компенсировать это явными стилистическими указаниями.

Формальные модели промптов

Для решения указанных выше проблем и формального представления промпта можно предложить следующую параметрическую модель: P = (R, C, T, S, F, L) где R — роль, C — контекст, T — задача, S — структура, F — формат, L — ограничения. Эта модель показывает, что промпт представляет собой структурированный управляющий объект, подчинённый вполне формализуемой логике.

Компоненты модели:

R (роль) — определяет «личность» или экспертный профиль модели при генерации ответа, что нужно для задания стиля, глубины, перспективы ответа. Например: *«ты — опытный юрист», «представь, что ты учитель физики», «действуй как креативный копирайтер».*

C (контекст) — дополнительная информация, помогающая модели понять ситуацию, что нужно для сужения пространства возможных ответов, повышения релевантности и точности. Например: предыстория вопроса, факты, данные, ссылки на документы, описание аудитории.

T (задача) — чёткая инструкция о том, *что именно* нужно сделать. Это ядро промпта, без него модель не поймёт цель запроса. Например, *«составь план», «объясни разницу», «напиши код», «проанализируй текст».*

S (структура) — требования к организации ответа, это нужно, чтобы сделать результат удобным для восприятия и использования. Например, *«раздели на 3 абзаца», «оформи в виде нумерованного списка», «создай таблицу с колонками X и Y».*

F (формат) — технические и стилистические параметры вывода. Это обеспечивает соответствие результата техническим или эстетическим требованиям. Например: задай *«в формате JSON», «в стиле научного доклада», «используй неформальный тон», «на английском языке».*

L (ограничения) — границы и запреты для генерации. Это нужно для предотвращения нежелательных элементов в ответе и экономии ресурсов,

например: *«не более 200 слов», «без жаргонизмов», «не используй цитаты», «исключи упоминания бренда X».*

Пример использования модели:

Промпт без структуры: «Напиши про климат».

Промпт по модели **P = (R, C, T, S, F, L):**

R: ты — журналист в сфере науки.

C: статья для образовательного сайта для старшеклассников.

T: объясни, как парниковый эффект влияет на глобальное потепление.

S: введение (1 абзац), основная часть (3 абзаца), заключение (1 абзац).

F: текстовый абзац, без формул и терминов.

L: не более 300 слов, нейтральный тон, без пугающих прогнозов.

В результате высока вероятность получить чёткий, целевой ответ, соответствующий ожиданиям.

К плюсам предлагаемой модели можно отнести следующее: универсальность, т.к. она подходит для любых типов запросов (творческих, аналитических, технических); повторяемость: ее легко адаптировать под новые задачи, меняя один-два параметра; прозрачность**:** каждый элемент можно обсудить и уточнить.

Безусловно, в модель можно добавить дополнительные параметры, такие как параметр **A** (аудитория) — целевая аудитория (если контекст не охватывает это) или дополнительно ввести **P** (тон) — тон ответа (дружелюбный, официальный, ироничный и т.п.), если он не подразумевается ролью. Предложенную модель можно использовать как базовую и постепенно усложнять, добавляя параметры по мере роста сложности задачи.

В целом, такая модель способна улучшить качество взаимодействия с языковыми моделями. Кроме всего вышеперечисленного допустимо архитектурное понимание промпта как внешнего модуля управления поведением LLM, аналога сценария, спецификации (в техническом понимании термина) или же декларативной программы.

Таким образом, рассматривая промпт-инжиниринг в лингвистической перспективе, мы неизбежно приходим к выводу, что перед нами не просто новый практический навык, а качественно новая форма языковой деятельности, возникшая на стыке технологий, лингвистики, когнитивных наук и философии языка. Промпт — это не «запрос к системе» в узком техническом смысле. Это управляющий текст, который формализует задачу, конструирует контекст, задаёт роль и статус адресата, определяет жанр и регистр будущего высказывания, описывает правила и ограничения дискурса, эксплицирует метатекстовые и метадискурсивные требования.

В традиционной коммуникации значительная часть этих параметров передаётся неявно: через ситуацию, социальные роли, культурные нормы, интонацию, общий опыт. В случае языковой модели весь этот «невидимый слой» должен быть эксплицирован в тексте. Промпт фактически

превращается в вербальную модель контекста, максимально компактную, но при этом насыщенную инструкциями.

В рамках проведённого анализа были рассмотрены несколько ключевых измерений промпт-инжиниринга.

Во-первых, семантическое измерение: промпт требует иной степени семантической точности, чем обычный текст. Размытые, прототипически организованные категории (*«подробно», «кратко», «профессионально», «глубоко»*) оказываются недостаточными, поскольку модель не обладает человеческими прототипами и опытом. Семантика промпта должна быть операционализирована: каждый термин привязан к предполагаемой операции, каждому требованию сопоставлен конкретный результат.

Во-вторых, прагматическое измерение: теория речевых актов показывает, что промпт — это особая разновидность директивного иллокутивного акта, в котором условия успешности, цель, адресат и роль коммуникаторов должны быть эксплицированы, а не оставлены на откуп контексту. Кроме того, для эффективного промпта необходимо соблюдение кооперативных принципов: релевантности, достаточности, ясности и непротиворечивости. Игнорирование этих принципов приводит к тому, что модель, лишённая собственных прагматических интуиций, заполняет пустоты произвольными, статистически вероятными интерпретациями.

В-третьих, дискурсивное измерение: промпт задаёт макроструктуру ответа — последовательность разделов, тип аргументации, жанровые рамки, структуру связности. Здесь становятся актуальными идеи Т. ван Дейка, И. Гофмана, М. Фуко: промпт выступает как инструмент конструирования дискурса, управления ролевыми позициями модели, регуляции допустимого диапазона высказываний. Пользователь, по сути, оказывается в роли архитектора дискурса, задающего правила «игры в текст» для искусственного собеседника.

В-четвёртых, социолингвистическое и ролевое измерение: ролевая установка (*«ты — юрист», «ты — учитель», «ты — врач»*) активирует в модели целые пласты профессионального дискурса, вместе с их лексикой, риторикой и стратегиями аргументации. Через промпт искусственно конструируется социальная и профессиональная идентичность говорящего, что является крайне необычным феноменом для традиционной лингвистики: социальная роль здесь не выводится из контекста, а жёстко задаётся текстом.

В-пятых, когнитивное измерение: человек, формулируя промпт, в явной форме выполняет работу, которая обычно происходит неосознанно — задаёт фреймы, уточняет прототипы, явным образом проговаривает цели, разбивает задачу на шаги, устанавливает критерии «хорошего» результата. Это делает промпт-инжиниринг своего рода внешней метакогнитивной практикой: мышление структурируется через язык, а язык — под требования взаимодействия с моделью. Теория прототипов, фреймов и когнитивных моделей здесь обретает неожиданный прикладной аспект.

В-шестых, философско-языковое измерение: классические вопросы философии языка — о значении, употреблении, референции, истине — получают новое измерение в контексте работы с LLM. Витгенштейновская идея языка как игры проявляется в том, что промпт вынуждает пользователя задавать правила игры эксплицитно. Крипкеанская проблема референции становится особенно острой: при недостаточной спецификации модель теряет связь с конкретными объектами и воспроизводит лишь описательные клише. Деннетовское различие уровней описания (физического, функционального, интенционального) также оказывается релевантным: промпт-инженер оперирует на функциональном уровне, хотя дискурс с моделью внешне оформляется в интенциональных терминах.

Отдельного внимания заслуживают корпусные и межъязыковые аспекты. Доминирование англоязычных корпусов неизбежно формирует «англосаксонский» стиль аргументации и структуры текста, которые затем переносятся в ответы на других языках. Русский язык, с его развитой морфологией, свободным порядком слов и богатой системой регистров, предъявляет к промпту более жёсткие требования: без чёткой структурной и семантической спецификации модель легко «соскальзывает» в стилистические и смысловые искажения. Это подтверждает, что промпт-инжиниринг нельзя рассматривать вне рамок сравнительно-языкового и социолингвистического анализа.

Заключение

В сумме всё это позволяет сделать несколько принципиальных выводов.

Прежде всего, промпт-инжиниринг по своей сути является лингвистической дисциплиной. Он опирается на семантику, прагматику, дискурс-анализ, когнитивную лингвистику, философию языка и социолингвистику не в качестве абстрактных теорий, а как на непосредственные инструменты практики.

Промпт — это новая разновидность текста, несводимая ни к одному из классических жанров. Он совмещает в себе черты инструкций, алгоритмов, ролевых ремарок, жанровых шаблонов и нормативных метатекстов. Это делает его одновременно текстом и описанием правил порождения другого текста.

Промпт-инжиниринг выступает как зона напряжённого взаимодействия человеческого и машинного: именно здесь различие между смыслом и вероятностью, между пониманием и статистикой, между интенциональностью и её симуляцией становится максимально видимым и операционализируемым.

Анализ промптов открывает перспективы для формирования новой междисциплинарной области — лингвистики коммуникации «человек — модель», в рамках которой будут изучаться закономерности построения управляющих текстов; структурные и стилистические особенности дискурса взаимодействия с ИИ; когнитивные эффекты систематического «разворачивания» неявных смыслов в явные инструкции; роль ИИ в трансформации письменных норм и жанров.

Наконец, промпт-инжиниринг имеет и более широкий гуманитарный смысл. Он заставляет по-новому взглянуть на сам человеческий язык как на систему, в которой то, что мы привыкли считать «естественным» и «само собой разумеющимся» — контекст, пресуппозиции, жанровые ожидания, социальные роли, культурные нормы, — оказывается нуждающимся в явном описании. Взаимодействие с языковыми моделями выявляет, насколько сложна и насыщенна человеческая коммуникация, и одновременно показывает, что язык может быть не только средством выражения, но и точным инструментом управления сложными интеллектуальными системами.

Таким образом, промпт-инжиниринг — это не временный технический тренд, а признак глубокой трансформации языка в цифровую эпоху. Он делает очевидным то, что лингвистика давно утверждала теоретически: язык не просто отражает мир, фиксируя уже существующие объекты и явления. Он активно участвует в формировании нашего восприятия мира, структурируя опыт и создавая смысловые рамки, через которые человек интерпретирует действительность. Язык — активный механизм конструирования. Взаимодействуя с языковыми моделями, человек учится конструировать не только тексты, но и правила порождения текстов, тем самым переходя от употребления языка к осознанному управлению им и через него — к управлению искусственными носителями «языкового поведения».

Список литературы

1. *Salton G.* Automatic Text Processing: The Transformation, Analysis, and Retrieval of Information by Computer. Reading, MA : Addison-Wesley, 1989.
2. *Остапенко С.В., Халина Н.В.* Лингвистическая оптимизация промта в аспекте проблематики промт-инжиниринга // Вестник Тюменского государственного университета. Гуманитарные исследования. Humanitates. 2025. Т. 11. № 1 (41). С. 36–50. https://doi.org/10.21684/2411-197X-2025-11-1-36-50 EDN: XCFQTA
3. *Костин А.А.* Промпт-инжиниринг. Язык будущего. Екатеринбург : Ridero, 2025.
4. *Montague R.* Formal Philosophy: Selected Papers of Richard Montague / ed. by R.H. Thomason. New Haven; London : Yale University Press, 1974.
5. *Rosch E., Lloyd B.B.* (Eds.). Cognition and categorization. Lawrence Erlbaum Associates, 1978.
6. *Филлмор Ч.* Фреймы и семантика понимания // Новое в зарубежной лингвистике. Вып. XXIII. Когнитивные аспекты языка. М. : Прогресс, 1988. С. 52–92.
7. *Минский М.* Фреймы для представления знаний. М. : Энергия, 1979.
8. *Остин Дж.Л.* Слово как действие // Новое в зарубежной лингвистике. Вып. XVII: Теория речевых актов. М. : Прогресс, 1986. С. 22–129.
9. *Серль Дж.Р.* Что такое речевой акт // Новое в зарубежной лингвистике. Вып. XVII: Теория речевых актов. М. : Прогресс, 1986. С. 151–169.
10. *Серль Дж.Р.* Классификация иллокутивных актов // Новое в зарубежной лингвистике. Вып. XVII: Теория речевых актов. М. : Прогресс, 1986. С. 170–194. EDN: THRZAL
11. *Грайс П.Г.* Логика и беседа // Синтаксис и семантика. Т. 3: Речевые акты. Нью-Йорк : Academic Press, 1975. С. 41–58.
12. *Wittgenstein L.* Philosophical investigations. Chichester, West Sussex, UK : Wiley-Blackwell, 2009.

13. *Kripke S.A.* Reference and Existence the John Locke Lectures. New York : Oxford University Press, 2013.
14. *Dennett D.C.* The Intentional Stance. Cambridge, Mass. : MIT Press, 1987.
15. *Дейк Т.А. ван*. Язык. Познание. Коммуникация. М. : Прогресс, 1989.
16. *Гофман И.* Представление себя другим в повседневной жизни. М. : Канон-Пресс-Ц, Кучково поле, 2000.
17. *Фуко М.* Археология знания. СПб. : Гуманитарная академия, 2012.
18. *Hyland K.* Metadiscourse: Exploring Interaction in Writing. London & New York : Continuum, 2005.

References

1. Salton, G. (1989). *Automatic Text Processing: The Transformation, Analysis, and Retrieval of Information by Computer*. Reading, MA: Addison-Wesley.
2. Ostapenko, S.V., & Khalina, N.V. (2025). Linguistic Optimization of the Prompt in the Aspect of Prompt Engineering Problems. *Tyumen State University Herald. Humanities Research. Humanitates*, *11*(1(41), 36–50. (In Russ.). https://doi.org/10.21684/2411-197X-2025-11-1-36-50 EDN: XCFQTA
3. Kostin, A.A. (2025). *Prompt Engineering. The Language of the Future*. Yekaterinburg: Ridero. (In Russ.).
4. Montague, R. (1974). *Formal Philosophy: Selected Papers of Richard Montague*, R.H. Thomason (Ed.). New Haven; London: Yale University Press.
5. Rosch, E., & Lloyd, B. B. (eds.). (1978). *Cognition and Categorization*. Lawrence Erlbaum Associates.
6. Fillmore, C.J. (1988). Frames and the Semantics of Understanding. In: *New in Foreign Linguistics. Iss. XXIII. Cognitive Aspects of Language* (pp. 52–92). Moscow: Progress. (In Russ.).
7. Minsky, M. (1979). *Frames for Representing Knowledge*. Moscow: Energiya. (In Russ.).
8. Austin, J.L. (1986). How to Do Things with Words. In: *New in Foreign Linguistics. Iss. XVII: Theory Of Speech Acts* (pp. 22—129). Moscow: Progress. (In Russ.)
9. Searle, J.R. (1986). What is a Speech Act? In: *New in Foreign Linguistics. Iss. XVII: Theory of Speech Acts* (pp. 151—169). Moscow: Progress. (In Russ.)
10. Searle, J.R. (1986). A Classification of Ilocutionary Acts. In: *New in Foreign Linguistics. Iss. XVII: Theory of Speech Acts* (pp. 170–194). Moscow: Progress. (In Russ.). EDN: THRZAL
11. Grice, H.P. (1975). Logic and Conversation. In: P. Cole & J.J. Morgan (eds.), *Syntax and Semantics. Vol. 3: Speech Acts* (pp. 41–58). New York: Academic Press. (In Russ.).
12. Wittgenstein, L. (2009). *Philosophical Investigations*. Chichester, West Sussex, U.K.: Wiley-Blackwell.
13. Kripke, S.A. (2013). *Reference and Existence: The John Locke Lectures*. New York: Oxford University Press.
14. Dennett, D.C. (1987). *The Intentional Stance*. Cambridge, Mass.: MIT Press.
15. Dijk, T.A. van. (1989). *Language. Cognition. Communication*. Moscow: Progress. (In Russ.).
16. Goffman, E. (2000). *The Presentation of Self in Everyday Life*, A.D. Kovalyov (Trans.). Moscow: Kanon-Press-Ts, Kuchkovo Pole. (In Russ.).
17. Foucault, M. (2012). *The Archaeology of Knowledge*. Saint Petersburg: Humanitarian Academy. (In Russ.).
18. Hyland, K. (2005). *Metadiscourse: Exploring Interaction in Writing*. London & New York: Continuum.

Сведения об авторах:
Максименко Ольга Ивановна, доктор филологических наук, профессор, профессор кафедры теории языка, англистики и прикладной лингвистики лингвистического факультета, Государственный университет просвещения (105005, Российская Федерация, г. Москва, ул. Радио, д. 10а, стр. 2) *сфера научных интересов*: автоматизированный анализ текстовых массивов, лингвосемиотика, функциональная семантика, теория интерпретации; *e-mail*: maxbel7@yandex.ru
ORCID: 0000-0002-6611-8744. SPIN-код: 7708-5901, AuthorID: 374818.
Беляков Михаил Васильевич, доктор филологических наук, доцент, профессор кафедры русского языка, Московский государственный институт международных отношений (университет) МИД России (119454, Российская Федерация, г. Москва, проспект Вернадского, д. 76); *сфера научных интересов*: лингвосемиотика, дискурс-анализ, компьютерная лингвистика, семантические и когнитивные исследования языка; *e-mail:* m.belyakov@my.mgimo.ru
ORCID: 0000-0002-6230-9893. SPIN-код: 1761-5400, AuthorID: 474699.

Information about the authors:
Olga I. Maksimenko, Dr. Sc. (Philology), Full Professor, Professor at the Language Theory, Anglistics and Applied Linguistics Department, Linguistic Faculty, Federal State University of Education (10a built 2, Radio street, Moscow, Russian Federation, 105005); *Research interests*: automatic text analysis, linguosemiotics, linguistic semantics, interpretation theory*; e-mail:* maxbel7@yandex.ru
ORCID: 0000-0002-6611-8744. SPIN-code: 7708-5901, AuthorID: 374818.
Mikhail V. Belyakov, Dr. Sc. (Philology), Associate Professor, Professor at Russian Department, MGIMO University (76, Vernadskogo Ave., Moscow, Russian Federation, 119454); *Research interests*: linguosemiotics, discourse analysis, computer science, semantic and cognitive language studies*; e-mail:* m.belyakov@my.mgimo.ru
ORCID: 0000-0002-6230-9893. SPIN-code: 1761-5400, AuthorID: 474699.

RUDN Journal of Language Studies, Semiotics and Semantics
2313-2299 (Print), ISSN 2411-1236 (Online)
Вестник Российского университета дружбы народов.
Серия: ТЕОРИЯ ЯЗЫКА. СЕМИОТИКА. СЕМАНТИКА

2026 Vol. 17 No. 2 459–473
http://journals.rudn.ru/semiotics-semantics

DOI: 10.22363/2313-2299-2026-17-2-459-473
EDN: LJHONM

Научная статья / Research Article

TextAnalyst и GPT: модель текста (модель текстового описания сущего) и семантика текста

А.А. Харламов

Московский государственный лингвистический университет, *Москва, Российская Федерация*
Институт высшей нервной деятельности и нейрофизиологии РАН, *Москва, Российская Федерация*
Московский физико-технический институт (национальный исследовательский университет), *Долгопрудный, Российская Федерация*
Национальный исследовательский университет «Высшая школа экономики», *Москва, Российская Федерация*
✉kharlamov@analyst.ru

Аннотация. В статье описан один из подходов к представлению семантики (смысла) текста в виде однородной семантической сети со взвешенными вершинами и их связями. Это представление противопоставляется представлению смысла текста в современных искусственных нейронных сетях типа GPT, где текст также (если исследовать представление внимательно) представляется в виде сети в многомерном пространстве признаков, реализованном множеством искусственных нейронов, составляющих искусственную нейронную сеть. Различие заключается в невозможности разобраться с этим представлением в сети, использующей миллиарды параметров представления — в невозможности выявления в явном виде этой самой семантической сети. Потому многие механизмы обработки текстов (автоматическое аннотирование текстов, сравнение текстов по смыслу, классификация текстов) содержательно остаются на совести разработчиков таких инструментов из-за невозможности контролировать решения, принимаемые этими инструментами: слишком велик объем обрабатываемой информации и отсутствуют инструменты проверки качества работы.

Ключевые слова: смысл текста, однородная семантическая сеть текста

Конфликт интересов: автор заявили об отсутствии конфликтов интересов.

История статьи: дата поступления: 23.12.2025; дата приема в печать: 15.02.2026.

Для цитирования: *Харламов А.А.* TextAnalyst и GPT: модель текста (модель текстового описания сущего) и семантика текста // Вестник Российского университета дружбы народов. Серия: Теория языка. Семиотика. Семантика. 2026. Т. 17. № 2. С. 459–473. https://doi.org/10.22363/2313-2299-2026-17-2-459-473 EDN LJHONM

© Харламов А.А., 2026
This work is licensed under a Creative Commons Attribution-NonCommercial 4.0 International License
https://creativecommons.org/licenses/by-nc/4.0/legalcode

TextAnalyst and GPT: Text Model (Text Description Model) and Text Semantics

Alexander A. Kharlamov

Moscow State Linguistic University, *Moscow, Russian Federation*
Institute of Higher Nervous Activity and Neurophysiology, RAS, *Moscow, Russian Federation*
The Moscow Institute of Physics and Technology, *Dolgoprudny, Russian Federation*
National Research University "Higher School of Economics", *Moscow, Russian Federation*
✉kharlamov@analyst.ru

Abstract. The article describes one of the approaches to representing the semantics (meaning) of a text in the form of a homogeneous semantic network with weighted vertices and their connections. This representation is contrasted with the representation of the meaning of the text in modern artificial neural networks such as GPT, where the text is also (if you examine the representation carefully) represented as a network in a multidimensional feature space implemented by a set of artificial neurons that make up an artificial neural network. The difference is that it is impossible to interpret this representation in a network that uses billions of representation parameters—the impossibility of explicitly identifying this very semantic network. Therefore, many text processing mechanisms (automatic annotation of texts, comparison of texts by meaning, classification of texts) meaningfully remain the responsibility of the developers of such tools due to the inability to control the decisions made by these tools: The volume of processed information is too large and there are no tools for checking the quality of work.

Keywords: meaning of the text, homogeneous semantic network of the text

Conflict of interests: The author declares no conflict of interest.

Article history: Received: 23.12.2025; Accepted: 15.0212.2025.

For citation: Kharlamov, A.A. (2026). TextAnalyst and GPT: Text Model (Text Description Model) and Text Semantics. *RUDN Journal of Language Studies, Semiotics and Semantics, 17*(2), 459–473. (In Russ.). https://doi.org/10.22363/2313-2299-2026-17-2-459-473
EDN LJHONM

Введение

Искусственные нейронные сети! GPT! Народ сошел с ума от восторга перед этими явлениями природы. Если не смотреть на обратную сторону медали — совершенно непохожую по качеству реакцию GPT на такое же к нему обращение (когда ответ GPT — полная чушь!) — то этот восторг объясним. Это похоже на то, как старенькая хозяйка, распутывая провод пылесоса, говорит ему: «подожди, миленький, сейчас я тебя освобожу!». Что хочу, то и вижу.

Далее не будет почти ничего по поводу GPT, только в конце мы попытаемся коснуться как самого явления, так и границ его применимости.

Далее во введении требуется показать предметную область со ссылками на наиболее характерные работы в ней. Как следует из названия, в статье представлена предметная область «Семантика». Объем этой предметной области очень велик, что можно понять хотя бы из различных определений этой предметной области, которые мы приводить не будем. Все дело в том, что под этим названием «Семантика» собралось много всего в разных предметных

областях от психологии до лингвистики, от систем распознавания речи до систем для анализа текстов.

Поскольку автор чисто практически в основном работал с анализом текстов, понятие «Семантика» будет относиться именно к анализу текстов и касаться содержания (смысла) текста. Далее во введении будет рассмотрено несколько подходов к анализу семантики текста, что и представит предметную область, к которой относится данная работа. Будут рассмотрены подходы, где слова в тексте рассматриваются не связанными по смыслу (так называемый «мешок слов»), первым таким подходом был LSA (так называемый Латентный семантический анализ), будут рассмотрены подходы, где слова рассматриваются в цепочке некоторой длины (“word2vek”), наконец, будут рассмотрены подходы на основе графов.

Мешок слов. Первые подходы к анализу содержания текста были реализованы на основе так называемого «мешка слов», когда слова в тексте рассматриваются как несвязанные между собой элементы. Мешок слов (англ. *bag-of-words*) — упрощенное представление текста, которое используется в обработке естественных языков и информационном поиске. В этой модели текст (одно предложение или весь документ) представляется в виде мешка (множества) его слов без какого-либо учета грамматики и порядка слов, но с сохранением информации об их количестве. Это упрощенное представление текста, которое показывает, какие слова встретились в тексте, но при этом не учитывает их порядок. Такое представление легко запрограммировать, оно удобно для использования в задачах автоматической обработки текста. Несмотря на свою простоту, оно оказывается достаточно полезным и позволяет успешно решать такие задачи, как классификация текста, т.е. отнесение текста к определенной группе / категории.

Как устроен мешок слов? Представление мешка слов — это таблица с числами, в которой столбцы таблицы — уникальные слова, а строки — документы коллекции. В ячейках таблицы находится число вхождений слова в документ. Значит, в каждой строке получится набор чисел (он же вектор), характеризующий состав документа. В разных вариантах используются разные представления элементов этой таблицы (LSA, pLSA, LDA) [1, 2], но суть этой таблицы не меняется.

Word2vek. «Word-to-vec» — общее название для совокупности моделей на основе искусственных нейронных сетей, предназначенных для получения векторных представлений слов на естественном языке. Модель-вектор — включает, кроме анализируемого слова текста, еще *n* слов текста до и *n* слов после, то есть контекст слова. Используется для анализа семантики естественных языков, основанного на дистрибутивной семантике, машинном обучении и векторном представлении слов. Программное обеспечение под названием «word2vec» исходно было разработано группой исследователей Google в 2013 году [3].

Графы в анализе текстов. Графовый анализ — это метод обработки и визуализации данных, при котором все элементы системы представлены в виде узлов, а связи между ними — в виде рёбер. Такой подход позволяет увидеть, как именно элементы взаимодействуют: откуда происходили входы, какие элементы использовались для транзакций, где пересекаются цепочки элементов [4].

Как выяснилось, ни тот, ни другой, ни третий подходы не несут в себе почти ничего, что относится к содержанию (смыслу) текста, что можно было бы использовать практически при анализе текста, например, при сравнении текстов по смыслу. Но они содержат элементы представления, которые можно использовать, пытаясь выявить это самое «содержание» (смысл) текста.

Эти элементы рассматриваемого представления следующие: частота встречаемости элементов текста (слов), их объединение в цепочки, графическое представление «содержания» текста в виде графа.

Сейчас можно сказать несколько слов об этих понятиях, которые подробно будут рассмотрены позже. Частота встречаемости, пересчитанная в смысловой вес, — одно из основных понятий, характеризующих смысл текста; цепочки понятий (но не в тексте, а на семантической сети текста), характеризуют важность слов в тексте; а граф-отношений слов в тексте, взвешенный по вершинам и связям, является основной смысловой характеристикой текста.

Поговорим о семантической сети текста, которая почти исключительно характеризует содержание (смысл) текста.

Сеть как способ представления семантики текста

Семантическая сеть текста — понятие не новое. Хорошо известны так называемые неоднородные семантические сети [5], которые используются для представления содержания текста (сейчас одновременно с ними используются так называемые онтологии [6]). В них пытались соединить языковые и экстралингвистические представления. Если говорят о красном яблоке, то это и слово «красный», и 450 нм, которые характеризуют длину волны видимого спектра, соответствующую красному цвету.

Представляемая в семантической сети информация

Сначала поговорим о том, какая информация должна быть представлена в семантической сети, чтобы она характеризовала смысл текста. Дело в том, что семантика предложения текста в лингвистике представлена весьма хорошо, а семантика целого текста почти не рассматривается.

В семантике текста должно быть все, что касается языка, и ничего больше. Рассмотрим в качестве примера какой-то текст. Он состоит из пред-

ложений, содержание которых описывается некоторыми (хорошо известными в лингвистике) способами. Например, содержание предложения текста может быть представлено расширенной предикатной структурой (рис. 1).

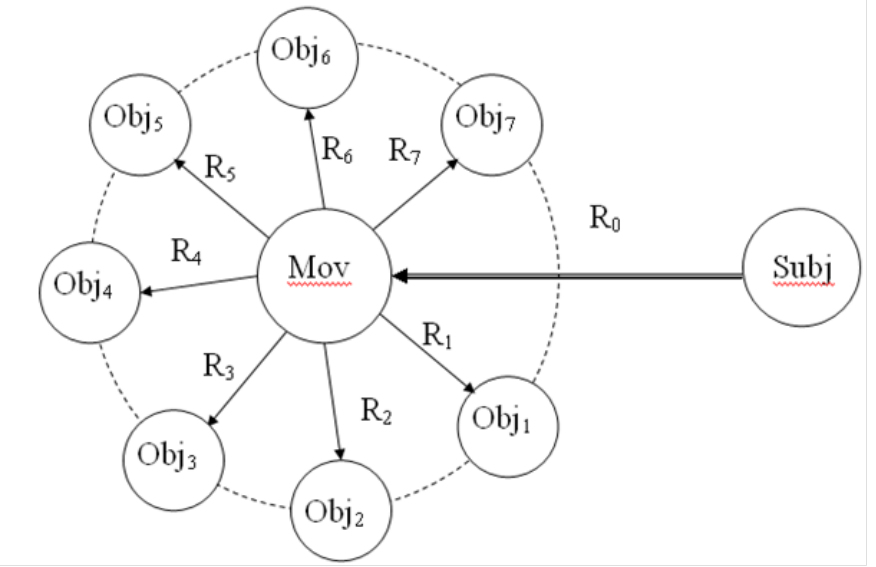


Рис. 1. Расширенная предикатная структура предложения (РПС).
Структура: Subj — субъект, Obj_i — актанты предиката, R_i — отношения предиката, R_0 — отношение «быть субъектом»
Источник: составлено А.А. Харламовым.

Figure 1. Extended Supply Predicate Structure (APF). Structure: Subj is the subject, Obji is the actant of the predicate, Ri is the relation of the predicate, R0 is the relation to "being a subject"
Source: compiled by Alexander A. Kharlamov.

Предварительно сделаем одно замечание: введем в РПС как фрагмент будущей семантической сети направление от субъекта к главному объекту, ко второстепенным объектам, и к атрибутам. Что из расширенной предикатной структуры предложения должно перейти в семантическую сеть текста? Элементами предложений одного текста могут быть одинаковые сущности. Объединим расширенные предикатные структуры предложений текста по одинаковым сущностям. Получим сеть, которую (поскольку она состоит из РПС предложений текста) можем назвать семантической сетью этого текста.

Чтобы она действительно характеризовала смысл текста необходимо сделать еще один шаг и выполнить одно требование. Этот шаг заключается в том, что при объединении РПС в сеть возникают цепочки событий, которые характеризуют текст в целом. Что это значит? Не частота встречаемости слов в тексте, а их взаимосвязи в тексте (которые отражаются в семантической сети текста), характеризуют их важность в тексте [7]. И потому при выявлении важности понятия в тексте осуществляется пересчет частоты встречаемости понятий в их смысловой вес [8].

Однородная семантическая сеть

Требование же, о котором шла речь выше, заключается в том, что в прототипе искусственной системы для смыслового анализа текста (в мозгу человека), в котором осуществляются подобные представления, работают две сигнальные системы (первая — экстралингвистическая, и вторая — языковая), и, соответственно, реализованы две семантические сети — экстралингвистическая и языковая [8]. Они работают вместе, дополняя друг друга (это в экстралингвистической семантической сети существует параметр, характеризующий термин «красный», находящийся в языковой семантической сети — 450 нм).

Поэтому человек, глядя на представленную ему языковую семантическую сеть текста, дополняет ее экстралингвистической семантической сетью (в основном это сеть, реализованная на представлениях визуальной и соматосенсорной модальностей), имеющейся в нашем мозге. Поэтому в языковой семантической сети не нужны некоторые подробности: нам не обязательно знать точно предикат, достаточно знать, что субъект и объект взаимодействуют. Собственно, эта информация не является лишней, но работа с сетью, в которой ее нет, проще: ее можно реализовать автоматически. Дело в том, что работа с неоднородными семантическими сетями (где взаимодействие субъекта с главным и второстепенными объектами обозначено отношением) существенно сложнее, чем работа с однородной семантической сетью. И по этой причине эта работа выполняется по сю пору вручную (пока не существует алгоритмов автоматической работы с неоднородными семантическими сетями [5]).

Таким образом, использование однородных семантических сетей позволяет, с одной стороны, автоматизировать процесс обработки текстовой информации (сложный момент здесь — вычислить семантические веса понятий — перевзвесить вершины сети [7]), а с другой стороны, это упрощает структуру сети (разные словоформы одного слова приводятся к одной корневой основе), что делает алгоритм обработки более робастным.

Обработка текста в технологии TextAnalyst

Использование семантической сети для обработки текстов (кроме моментов, связанных с обработкой отдельного текста, ее можно использовать для обработки совместно нескольких текстов, например, для сравнения текстов) позволяет реализовать, весьма удобную для представления текста функциональность, которая используется, например филологами, в виде гипертекстовой структуры [7].

Уже достаточно давно была разработана линейка продуктов TextAnalyst для смыслового анализа текстов [9], которая и реализует описанную выше технологию построения однородной семантической сети с перевзвешенными

в смысловой вес вершинами, функциональность которой по анализу текстов будет представлена ниже.

Разработанная система обработки текстовой информации основана на использовании структурных свойств языка и текста, которые могут быть выявлены с помощью статистического анализа, где статистика конкретного текста позволяет автоматически сформировать представление семантики текста в виде однородной семантической сети, на основе которого и реализуются функции организации текстовой базы в гипертекстовую структуру, автоматического реферирования, кластеризации и классификации текстов, а также функция смыслового поиска.

Этот подход позволяет на основе анализа статистики слов и их связей в предложениях текста реконструировать внутреннюю (смысловую) структуру текста. Статистический анализ фиксирует частоту появления элементов текста (слов или устойчивых словосочетаний), а также частоту появления взаимосвязей между выявленными элементами текста. Далее статистические показатели пересчитываются в семантические с помощью алгоритмов, подобных алгоритму перерасчета весов в сетях Хопфилда [10]. Идея подобных алгоритмов заключается в том, что при расчете какой-то характеристики элемента сети учитываются подобные характеристики элементов, с ним связанных, в виде численных показателей связей. После такого пересчета статистических характеристик текста в семантические характеристики понятий текста, которые мало соответствуют анализируемой предметной области, получают малый вес, а наиболее представительные — наделяются высокими показателями. Полученная семантическая сеть позволяет производить различные виды анализа текстовой информации. Сеть отражает внутреннюю структуру текста, значимость выделенных понятий, а также показывает степень связанности понятий в тексте. Такое представление текста может быть получено полностью автоматически. Так формируется семантическая сеть текста.

Семантическая сеть, представляющая текст, далее используется как для анализа исходного текста, так и для совместного анализа нескольких текстов анализируемого корпуса текстов. Анализируемый текст, входящий в гипертекстовую структуру вместе с построенной на основе этого текста семантической сетью, используется: (1) для представления смысла (структуры) этого текста в виде семантической сети; (2) для представления оглавления исходного текста в виде минимального древовидного подграфа этой сети; (3) вершины сети — слова текста — позволяют выявить предложения текста, содержащие эти слова; (4) предложения текста с наибольшим весом формируют аннотацию текста; наконец, (5) имеется возможность поиска элементов текста в этой гипертекстовой структуре с целью быстрого обращения к информации в тексте.

Возможность совместного анализа нескольких текстов позволяет: (1) выявлять степень смыслового пересечения текстов (пересечение их семантических сетей); (2) классифицировать тексты (выявлять степень смыслового пересечения семантической сети текста с семантическими сетями предметных областей — классов); (3) кластеризовать предметную область (корпус текстов) на кластеры.

Семантические веса элементов сети используются при расчете близости (релевантности) фрагментов текста к запросу со стороны пользователя системы. На их основе возможно выделение наиболее информативных участков текста. Использование ассоциативных связей элементов сети позволяет расширять поле поиска информации. Ответ на запрос пользователя в этом случае может содержать информацию, явно не указанную в запросе, но связанную с ней по смыслу.

Программная реализация технологии

Система TextAnalyst реализована как инструмент для автоматического формирования баз знаний на основе множества естественно-языковых текстов. Ядро системы выполнено как программный компонент (inproc server), соответствующий спецификации Component Object Model (COM) фирмы Microsoft [9].

Ядро системы реализует следующие функции: (1) нормализацию грамматических форм слов и вариаций словосочетаний; (2) автоматическое выделение базовых понятий текста (слов и словосочетаний) и их взаимосвязей с вычислением их относительной значимости в тексте; (3) формирование представления семантики текста (множества текстов) в форме семантической сети.

В состав ядра системы, помимо блока первичной обработки, входят следующие блоки (рис. 2): лингвистический процессор (2), блок выделения понятий текста (12), блок формирования семантической сети (10), блок хранения семантической сети (11).

Блок первичной обработки. Задачами этого блока являются извлечение текста из файла (входного потока данных) и подготовка его к обработке в лингвистическом процессоре. Подготовка текста заключается в очистке его от символов, неизвестных лингвистическому процессору, а также в корректной обработке таких единиц текста, как аббревиатуры, инициалы, заголовки, адреса, номера, даты, указатели времени.

Лингвистический процессор. Лингвистический процессор осуществляет предобработку входного текста (последовательности символов в определенной кодировке) на основе априорных лингвистических знаний, общих для выбранного языка (в настоящий момент поддерживаются несколько европейских языков, помимо русского и английского), и выполняет следующие функции: сегментацию предложений текста на основе знаков пунктуации и специальных грамматических слов и их фильтрацию; нормализацию слов и словосочетаний — фильтрацию флексий (окончаний) с сохранением только

корневых основ; фильтрацию в тексте семантически несущественных, вспомогательных слов: удаляются предлоги, числительные и самые общеупотребимые слова с широким значением; маркировку общеупотребимых слов.

Сегментация на предложения позволяет разбить текст на участки, которые могут содержать терминологические словосочетания предметной области, и избежать выделения неадекватных словосочетаний на стыках таких участков.

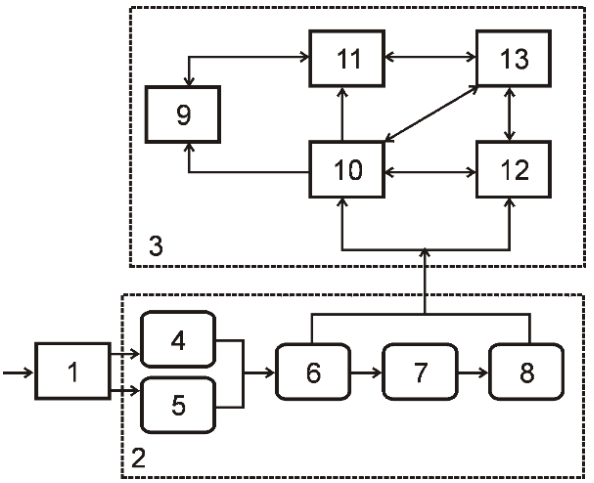


Рис. 2. Система анализа текстов содержит блок первичной обработки (1), лингвистический и семантический процессоры. Лингвистический процессор (2) состоит из словарей: (4) слов — разделителей, (5) служебных слов, (6) общеупотребимых слов, а также (7) флективных и (8) корневых морфем. Семантически процессор (3), в свою очередь, содержит: (9) блок отсылок в текст, (10) блок формирования семантической сети, (11) блок хранения семантической сети, (12) блок выделения понятий и (13) блок управления

Источник: составлено А.А. Харламовым.

Figure 2. A text analysis system comprises a primary processing unit (1), linguistic and semantic processors. The linguistic processor (2) consists of dictionaries: (4) divisional words, (5) service words, (6) common words, and (7) a phletical and (8) root morphs. Semantically, the processor (3), in turn, contains: (9) a unit of references into text, (10) a unit for forming a semantic network, (11) a unit for storing a semantic network, (12) a unit for separating concepts, and (13) a control unit.

Source: compiled by Alexander A. Kharlamov.

В результате предобработки семантически близкие словосочетания приводятся к одинаковой форме (нормализуются). Маркировка общеупотребимых слов необходима с целью исключения их выделения как самостоятельных терминов при дальнейшем анализе.

База общих языковых знаний лингвистического процессора содержит словари, по одному для реализации каждой из четырех функций: словарь слов-разделителей предложения, словарь вспомогательных слов, словарь флексий и словарь общеупотребимых слов.

Блок выделения понятий. Блок выделения ключевых понятий предметной области (слов и словосочетаний) реализует алгоритмы автоматического

формирования частотного словаря текста. Он реализован как иерархическая структура пар представлений, где первое представление в паре — это символ алфавита или номер предыдущей пары. Число уровней в иерархической структуре определяет априорно заданную максимально допустимую длину понятия предметной области и равняется двадцати.

На первом уровне иерархической структуры представлен словарь двухбуквенных специальных слов предметной области — слов, пропущенных через все фильтры лингвистического процессора и не отнесенных к общеупотребимым, а также двухбуквенных сочетаний из слов этого словаря. Там же хранятся двухбуквенные слова общеупотребимой лексики, входящие в устойчивые словосочетания, и их начальные двухбуквенные фрагменты. На втором уровне иерархической структуры хранятся словари трехбуквенных слов и сочетаний букв из словарей специальных и общеупотребимых слов, встреченных в тексте, в виде индексов элементов соответствующих словарей первого уровня, дополненных еще одной буквой. На последующих уровнях представление информации полностью однородно: там хранятся индексы элементов хранения более низкого уровня иерархии, дополненные одной буквой.

В процессе формирования представления информации в иерархической структуре подсчитывается частота встречаемости каждого сочетания букв в соответствующих элементах иерархии. Частота слов (сочетаний букв, не имеющих продолжения на следующем уровне) используется для последующего анализа: если есть скачок частоты — слово завершено.

Сформированное таким образом представление лексики текста подвергается затем пороговому преобразованию по частоте встречаемости. Порог отражает степень детальности описания текста. В процессе статистического анализа в иерархической структуре выделяются устойчивые термины и терминологические словосочетания, которые служат далее в качестве элементов для построения семантической сети. При этом общеупотребимые слова, а также словосочетания, содержащие только общеупотребимые слова, опускаются.

Блок формирования семантической сети. Блок формирования семантической сети реализован как база данных, в которой представляются семантические связи понятий предметной области. Поскольку типы семантических связей [11] в системе не определяются, такие связи представляют собой просто ассоциативные связи.

В качестве критерия для определения наличия семантической связи между парой понятий используется частота их совместной встречаемости в предложениях текста. Превышение частотой некоторого порога позволяет говорить о наличии между понятиями ассоциативной (семантической) связи, а совместные вхождения понятий в предложения с частотой меньше порога считаются просто случайными.

Элементы семантической (ассоциативной) сети и их связи имеют числовые характеристики, отражающие их относительный вес в данной предметной области — семантический вес. При достаточно представительном множестве текстов, описывающих предметную область, значения частот встречаемости понятий действительно отражают соответствующие семантические (субъективно оцениваемые) веса. Однако для небольших выборок, в частности при анализе отдельного текста, не все частотные характеристики соответствуют действительным семантическим весам — весам важности понятий в тексте. Для более точной оценки семантических весов понятий используются веса всех связанных с ними понятий, т.е. веса целого семантического сгущения. В результате такого анализа наибольший вес приобретают понятия, обладающие мощными связями и находящиеся как бы в центре семантических сгущений.

Основные функции системы «TextAnalyst»

На основе результатов работы модуля индексации реализованы следующие функции обработки текстовой информации: (1) функция формирования гипертекстовой структуры; (2) навигации по базе знаний; (3) формирования тематического дерева; (4) реферирования текстов; (5) автоматической кластеризации множества текстов; (6) сравнения текстов (автоматической классификации текстов); и, наконец, (7) функция формирования ответа на смысловой запрос пользователя — формирование тематического реферата.

После формирования семантической сети исходный текст, объединенный гиперссылками с семантической сетью, становится гипертекстовой структурой. Семантическая сеть в этом случае оказывается удобным средством навигации по тексту. Она позволяет исследовать основную структуру текста, переходя от понятия к понятию по ассоциативным связям. Пользуясь гиперссылками, пользователь может быстро найти множество предложений текста, содержащих эти понятия. С помощью тех же гиперссылок он может перейти от любого предложения непосредственно к его контексту в тексте. С этой же целью пользователь может пользоваться минимальным древовидным подграфом семантической сети — тематическим деревом. В нем оказываются иерархически представленными основные и соподчиненные понятия сети, причем понятия нижнего уровня объясняют содержание понятий более высокого уровня. Тематическим деревом также можно пользоваться для навигации по базе знаний — оно напоминает оглавление текста.

Семантическая сеть с числовыми характеристиками ее компонентов — понятий и их связей — позволяет вычислить вес каждого предложения в тексте. Множество предложений текста, выбранных в порядке их появления в тексте, вес которых превысил некоторый пороговый уровень, можно считать рефератом текста.

Семантическая сеть исследуемого текста (или группы текстов) может быть разбита на подсети удалением из нее слабых связей. Каждая такая подсеть группируется вокруг некоторого понятия с максимальным весом в данной подсети. Это понятие обозначает тему части текста или отдельных текстов, которые оказываются сгруппированными в данной подсети. Такая автоматическая кластеризация позволяет разбить множество текстов на рубрики, а также визуализировать динамику развития этих рубрик во времени.

Используя числовые характеристики семантической сети, можно сравнивать сети двух текстов с точки зрения вычисления их пересечения (общей части). То есть можно сравнивать степень совпадения текстов по смыслу. Если в качестве одного из текстов берется целая рубрика, то имеется возможность оценить степень принадлежности исходного текста к данной рубрике, то есть автоматически классифицировать тексты.

Система для смыслового анализа текстов позволяет реализовать также смысловой поиск (сформировать тематический реферат). Функция смыслового поиска, основываясь на ассоциативном иерархическом представлении содержания информации в базе, функциях кластеризации и классификации, осуществляет выборку информации, соответствующей запросу пользователя, и структурирует ее в соответствии с близостью к запросу.

Смысловой поиск, используя ассоциации, способен выдавать пользователю информацию, явно не указанную в тексте запроса, но связанную с ней по смыслу. Использование такого подхода ведет не к увеличению выдаваемой пользователю информации, а к ее тщательному отбору и анализу по главному критерию — смысловой близости к запросу.

Представление текста в технологии GPT

Как же устроена база знаний трансформеров, так стремительно ворвавшихся в нашу жизнь? Если отбросить подробности представления и использования информации в технологии GPT (как наиболее известного представителя этой когорты), то (опять же независимо от объема обрабатываемой информации) представление корпусов текстов в технологии GPT сводится к довольно длинным цепочкам текстовой информации заданной длины, представленным в искусственных нейросетевых структурах. Эти текстовые цепочки представляются в некоем многомерном пространстве, где возможно вычисление расстояния до такой цепочки (посредством механизма многоголового внимания [12]), с выбором искомой цепочки с наименьшим до нее расстоянием. Манипулирование этой информацией сводится к заданию вопроса системе и к синтезу ответа искусственной нейронной сети (ИНС) в виде связного текста. Вы — вопрос, Вам — ответ, где синтезируемый ответ оптимально приближен в исходном многомерном пространстве к Вашему вопросу с точки зрения упомянутого механизма многоголового внимания.

В силу того, что длина цепочки в такой ИНС может быть достаточно большой, создается впечатление, что такая ИНС самостоятельно генерирует

осмысленный связный текст. На самом деле это впечатление обманчиво: (1) суть этого текста сводится к ответу на вопрос; и (2) сгенерирован может быть только текст, который по смыслу содержится в базе данных ИНС (хотя бы собранный из разных частей). Попробуйте пообщаться с такой ИНС (хоть и с GPT) по поводу статьи Н.А. Козырева «Причинная или несимметричная механика в линейном приближении» [13]. Мы пытались, но ничего путного у ИНС выяснить не удалось (только Вы должны хорошо знать предметную область, чтобы корректно интерпретировать ответы ИНС): разговор с ИНС был очень длинным — несколько дней, поэтому привести его не представляется возможным.

Заключение

В заключение можно сказать следующее. Современные искусственные нейронные сети продвинулись далеко в реализации алгоритмов обработки информации разных видов: устных и письменных текстов, изображений, даже запахов. Но речь идет фактически только о распознавании.

В настоящий момент реализованы достаточно изощренные механизмы анализа (до сих пор не реализованы алгоритмы распознавания сцен — совокупностей объектов — разработчики пока не догадались, что для этого надо совместить глубокие нейронные сети с сетями Хопфилда). Алгоритмов синтеза информации как целенаправленного поведения, реализовано до сих пор не было.

Дело в том, что целенаправленное поведение предполагает достижение цели, которая до сих пор ставится только человеком. Целенаправленное поведение по существу является днятельностью, направленной на удовлетворение потребностей (неважно каких). А удовлетворение потребностей искусственных систем может противопоставить искусственные системы естественным интеллектуальным системам (человеку). И хоть до возникновения сугубого искусственного интеллекта еще очень далеко (разве можно сравнить результаты проекта “Blue Brain” с естественным мозгом человека?), а возможно, искусственного интеллекта, сравнимого с человеческим, никогда и не появится; даже мысленно их желательно не противопоставлять.

Однако интеллектуальные системы, которые решают поставленные человеком задачи лучше него, необходимо разрабатывать и ими пользоваться. И известный прогресс в этом направлении явно есть. Просто его не следует преувеличивать.

Список литературы

1. *Deerwester S., Dumais S.T., Landauer T.K., Furnas G.W., Harshman R.* Indexing by Latent Semantic Analysis // Journal of American Society of Information Science. 1990. Vol. 41. No. 6. P. 391–407. https://doi.org/10.1002/(SICI)1097-4571(199009)41:6<391::AID-ASI1>3.0.CO;2-9
2. *Blei D.M., Ng A.Y., Jordan M.I.* Latent Dirichlet Allocation // The Journal of Machine Learning Research. 2003. № 3. P. 993–1022.

3. *Mikolov T., Chen K., Corrado G., Dean J.* Efficient Estimation of Word Representations in Vector Space. In: *Proceedings of Workshop at ICLR, Scottsdale, 2–4 May 2013.* Scottsdale : AZ, Scottsdale, U.S.A.: Poisoned Pen Press.2013. P. 1–11.
4. *Григорьева Е.Г., Клячин В.А.* Исследование статистических характеристик текста на основе графовой модели лингвистического корпуса // Известия Саратовского ун-та. Новая Серия Математика. Механика. Информатика. 2020. Т. 20. Вып. 1. С. 116–126. https://doi.org/10.18500/1816-9791-2020-20-1-116-126 EDN: KNPYYV
5. *Голенков В.В., Гулякина Н.А.* Графодинамические модели параллельной обработки знаний: принципы построения, реализации и проектирования // Открытые семантические технологии проектирования интеллектуальных систем. 2012. № 2. С. 3–13. EDN: YUCQXM
6. *Гаврилова Т.А., Хорошевский В.Ф.* Базы знаний интеллектуальных систем. СПб. : Питер, 2001.
7. *Kharlamov A.A.* Network N-gram Model of the Text. Topic Tree of the Text — Minimal Tree Subgraph of the Semantic Network. In book: Neuroinformatics and Semantic Representations: Theory and Applications. Cambridge: Cambridge Scholars Publishing, 2020. P. 140–152.
8. *Kharlamov A.A.* Text and Quasi-Text Analysis: Upper Levels of Speech Information Processing. In: Integral Robot Technologies and Speech Behavior. A.A. Kharlamov, M. Pilgun (Eds.). Collective Monography. Chapter Three. Cambridge: Cambridge Scholars Publishing, 2023. P. 88–129.
9. *Hopfield J.J.* Neural networks and physical systems with emergent collective computational abilities // Proc. Natl. Acad. Sci. 1982. No. 79. P. 2554–2558
10. *Осипов Г.С.* Приобретение знаний интеллектуальными системами: Основы теории и технологии. М. : Наука. Физматлит, 1997. EDN: QOWVDX
11. *Radford A., Wu J., Child R., Luan D., Amodei D., Sutskever I.* Language Models are Unsupervised Multitask Learners. OpenAI Blog, 1, 9. 2019.
12. *Козырев Н.А.* Причинная или несимметричная механика в линейном приближении. Акад. наук СССР. Глав. астрон. обсерватория: Пулково, 1958.

References

1. Deerwester, S., Dumais, S.T., Landauer, T.K., Furnas, G.W., & Harshman, R. (1990) Indexing by Latent Semantic Analysis. *Journal of American Society of Information Science. 41*(6). 391–407. https://doi.org/10.1002/(SICI)1097-4571(199009)41:6<391::AID-ASI1>3.0.CO;2-9
2. Blei, D.M., Ng, A.Y., & Jordan, M.I. (2003) Latent Dirichlet Allocation. T*he Journal of Machine Learning Research*. 3. 993–1022.
3. Mikolov T., Chen K., Corrado G., & Dean J. (2013) Efficient Estimation of Word Representations in Vector Space. In: *Proceedings of Workshop at ICLR, Scottsdale, 2–4 May 2013.* Scottsdale : AZ, Scottsdale, U.S.A.: Poisoned Pen Press. 1–11.
4. Grigorieva, E.G., & Klyachin, V.A. (2020) The Study of The Statistical Characteristics of The Text Based on the Graph Model of the Linguistic Corpus. *Izvestiya of Saratov University. New Series. Mathematics. Mechanics. Informatics*. *20*(1). 116–126. (In Russ.) https://doi.org/10.18500/1816-9791-2020-20-1-116-126 EDN: KNPYYV
5. Golenkov, V.V., & Guliakina, N.A. (2012) Graphodynamical Models of Parallel Knowledge Processing. *Open Semantic Technologies for Intelligent Systems Design*. 2. 3–13. (In Russ.) EDN: YUCQXM
6. Gavrilova, T.A., & Khoroshevsky, V.F. (2001). *Knowledge Bases of Intelligent Systems. St.* Petersburg: Piter. (In Russ.)
7. Kharlamov, A.A. (2020) Network N-gram Model of the Text. Topic Tree of the Text — Minimal Tree Subgraph of the Semantic Network. In book: *Neuroinformatics and Semantic Representations: Theory and Applications*. Cambridge: Cambridge Scholars Publishing. 140–152.

8. Kharlamov, A.A. (2023) Text and Quasi-Text Analysis: Upper Levels of Speech Information Processing. In: *Integral Robot Technologies and Speech Behavior*. A. A. Kharlamov, M. Pilgun (Eds.). Collective Monography. Chapter Three. Cambridge: Cambridge Scholars Publishing. 88–129.
9. Hopfield, J.J. (1982) Neural networks and physical systems with emergent collective computational abilities. *Proc. Natl. Acad. Sci*. 79. 2554–2558.
10. Osipov, G.S. (1997). *Knowledge Acquisition by Intelligent Systems: Fundamentals of Theory and Technology*. Moscow: Nauka. Fizmatlit. (In Russ.) EDN: QOWVDX
11. Radford, A., Wu, J., Child, R., Luan, D., Amodei, D., & Sutskever, I. (2019) *Language Models are Unsupervised Multitask Learners. OpenAI Blog, 1, 9.*
12. Kozyrev, N.A. (1958). *Causal or Asymmetric Mechanics in Linear Approximation*. Leningrad: Acad. Sci. USSR, Main Astronomical Observatory. (In Russ.)

Сведения об авторе:
*Харламов Александр Александрович***,** доктор технических наук, профессор, профессор кафедры прикладной и экспериментальной лингвистики, Московский государственный лингвистический университет (119034, Российская Федерация, г. Москва, ул. Остоженка, д. 38 с. 1); заведующий кафедрой интеллектуальных информационных систем и технологий, Московский физико-технический институт (национальный исследовательский университет) (141701, Российская Федерация, Московская область, г. Долгопрудный, Институтский переулок, д. 9); старший научный сотрудник лаборатории нейроонтогенеза, Институт высшей нервной деятельности и нейрофизиологии РАН (117485, Российская Федерация, Москва, ул. Бутлерова, д. 5А); профессор департамента программной инженерии, Национальный исследовательский университет «Высшая школа экономики», (101000, Российская Федерация, Москва, Мясницкая ул., 20); *сфера научных интересов*: нейроинформатика; *e-mail:* kharlamov@analyst.ru
ORCID: 0000-0003-2942-5101; Scopus Author ID: 57193909855; ResearcherID: E-3760-2014; SPIN-код: 6959-6240.

Information about the author:
Alexander A. Kharlamov, Dr.Sc. (Technics); Prof. at Applied and Experimental Linguistics Department, Moscow State Linguistic University (38 build 1 Ostozhenka Str., Moscow, Russian Federation, 119034); Senior Researcher at Neuroontogenezis Laboratory, Institute of Higher Nervous Activity and Neurophysiology, RAS (5A Butlerova Str, Moscow, Russian Federation, 117485); Intellectual Information Systems and Technology Department Chair, The Moscow Institute of Physics and Technology (9 Institutskiy per., Dolgoprudny, Moscow Region, Russian Federation, 141700); Prof. at Program Engineering Department of HSE, National Research University “Higher School of Economics” (20 Myasnitskaya Str., Moscow, Russian Federation, 101978); *Research interests*: Neuroinformatics; *e-mail:* kharlamov@analyst.ru
ORCID: 0000-0003-2942-5101; Scopus Author ID: 57193909855; ResearcherID: E-3760-2014; SPIN code: 6959-6240.

RUDN Journal of Language Studies, Semiotics and Semantics
2313-2299 (Print), ISSN 2411-1236 (Online)
Вестник Российского университета дружбы народов.
Серия: ТЕОРИЯ ЯЗЫКА. СЕМИОТИКА. СЕМАНТИКА

2026 Vol. 17 No. 2 474–495
http://journals.rudn.ru/semiotics-semantics

DOI: 10.22363/2313-2299-2026-17-2-474-495
EDN: LJXNSN

Научная статья / Research Article

Генеративные нейросети как инструмент транспозиции печатного текста в экранную форму на примере работ нейроиллюстратора DJ БлокNote

Ю.А. Евграфова

Национальный исследовательский университет «МЭИ», *Москва, Российская Федерация*
yu.evgrafova@mail.ru

Аннотация. Актуальность рассмотрения медиальной транспозиции при участии генеративных нейросетей и особенностей осуществляемого переноса вербального содержания в визуальную форму обусловлена тенденциями применения искусственного интеллекта в современной коммуникации при создании информационной среды и для трансляции культурной памяти, в том числе через её ремедиацию, переосмысление. Целью данной статьи является рассмотрение особенностей переноса печатного поэтического текста в медийный формат при помощи генеративных нейросетей, выявление специфических приёмов, а также черт происходящей при этом ремедиации исходного текста. Материалом исследования послужили работы (всего 6) нейроиллюстратора Никиты Глухова (псевдоним DJ БлокNote). Генеративные нейросети, направляемые человеком-оператором и обучаемые на массиве антропогенных данных, являются опосредованной формой социального взаимодействия, что предполагает реализацию интегративного подхода и анализ заявленных явлений с позиций интеракционизма, герменевтики, психологической и визуальной антропологии, лингвосемиотики. Основными методами исследования выступают качественные методы: интерпретация, анализ, синтез, описательный метод, метод контекстуального анализа, функционально-прагматический метод, элементы психоаналитического метода для извлечения скрытых смыслов. Применяется трехуровневый анализ семиотически осложненного текста, разработанный Р. Бартом, адаптированный под нужды исследования. Было обнаружено следующее. При транспозиции печатного текста в экранный медиум с участием генеративных нейросетей происходит перекодировка его текстовой структуры в зрительную и звуковую материю. Сконструированное произведение состоит из интермедиальных образов, а динамика, монтаж, композиция, цвет и свет заменяют ритм и метафоричность. Помимо аудиовизуального обнаруживается и сюжетно-фабульный выразительный слой — нарративизация лирического. Промпты для генеративной модели создаются человеком-оператором на основе опорных семантических точек стихотворения — конкретных лексических единиц (ЛЕ), с добавлением в них дополнительных ЛЕ, которые отсутствуют в исходном тексте, но которые нужны для его визуализации. Следуя промпту, нейросеть, овеществляя контекст стихотворения, «обрамляет» заложенные в него символы, «переводит» эмоциональную экспрессию в динамическую целостность.

© Евграфова Ю.А., 2026
This work is licensed under a Creative Commons Attribution-NonCommercial 4.0 International License
https://creativecommons.org/licenses/by-nc/4.0/legalcode

Ключевые слова: ИИ, экранный текст, артефакт культуры, поэтический текст, декодирование, герменевтика

Конфликт интересов: автор заявил об отсутствии конфликтов интересов.

История статьи: дата поступления: 20.01.2026; дата приема в печать: 10.02.2026.

Для цитирования: *Евграфова Ю.А.* Генеративные нейросети как инструмент транспозиции печатного текста в экранную форму на примере работ нейроиллюстратора DJ БлокNote // Вестник Российского университета дружбы народов. Серия: Теория языка. Семиотика. Семантика. 2026. Т. 17. № 2. С. 474–495. https://doi.org/10.22363/2313-2299-2026-17-2-474-495 EDN LJXNSN

Generative Neural Networks as a Tool for Transposing Printed Text into Screen Format: Case Study of Neuro-Illustrator DJ БлокNote

Yulia A. Yevgrafova ✉

National Research University «Moscow Power Engineering Institute», *Moscow, Russian Federation*
✉yu.evgrafova@mail.ru

Abstract. This article examines media transposition involving generative neural networks and the specifics of transferring verbal content into visual form. The relevance of the research is determined by the trends in using "artificial intelligence" to create an infosphere in modern communication and to transmit cultural memory, involving its remediation and reinterpretation. The aim of this article is to examine the characteristics of transferring printed poetic text into a media format using generative neural networks, to identify specific techniques, and to outline the features of the original text's remediation that occurs in this process. The research material consists of works (6 in total) by neuro-illustrator Nikita Glukhov (pseudonym DJ БлокNote). Generative neural networks, guided by a human operator and trained on a dataset of human-generated content, represent a mediated form of social interaction. In this regard, this study implements an integrative approach and examines the stated phenomena from the perspectives of interactionism, hermeneutics, psychological and visual anthropology, and linguosemiotics. The main research methods are qualitative: interpretation, analysis, synthesis, descriptive method, functional-pragmatic and contextual analysis, and elements of psychoanalytic method for extracting hidden meanings. A three-level analysis of semiotically complex text, developed by R. Barthes, and adapted to the needs of the research, is applied. The following was discovered. During the transposition of printed text into a screen medium with the involvement of generative neural networks, its textual structure is recoded into visual and auditory matter. The constructed work consists of intermedial images, while dynamics, montage, composition, color, and light replace rhythm and metaphoricity. In addition to the audio-visual, a plot-narrative expressive layer is also identified – the narrativization of the lyrical. Prompts for the generative model are created by the human operator based on key semantic points of the poem – specific lexical units (LUs), with the addition of extra LUs that are absent in the original text but necessary for its visualization. Following the prompt, the neural network materializes the poem's context, "frames" the symbols embedded within it, and "translates" emotional expression into a dynamic whole. This article is aimed at students, graduate students, teachers, linguists, philologists, and anyone interested in general and linguistic semiotics.

Key words: AI (Artificial Intelligence), screen text, cultural artifact, poetic text, decoding, hermeneutics

Conflict of interests: The author stated that there were no conflicts of interest.

Article history: Received: 20.01.2026; Accepted: 10.02.2026.

For citation: Yevgrafova, Yu.A. (2026). Generative Neural Networks as a Tool for Transposing Printed Text into Screen Format: Case Study of Neuro-Illustrator DJ БлокNote. *RUDN Journal of Language Studies, Semiotics and Semantics, 17*(2), 474–495. (In Russ.). https://doi.org/10.22363/2313-2299-2026-17-2-474-495 EDN LJXNSN

Введение

Цифровые технологии приобретают онтологический статус. Они вытесняют традиционные коммуникативные практики, трансформируют не только социальное бытие и атрибуты [1], но и культурные формы, и эстетические ценности [2]. В этом контексте проблема актуализации и «осовременивания» культурного наследия техническими средствами, в том числе генеративными нейросетями, является насущной проблемой. Генеративные нейросети, прошедшие путь от автокодировщика до сложной нейросетевой архитектуры, в современных художественных практиках применяются как средство создания новых и интерпретации уже существующих артефактов культуры. Они продолжают эволюционировать, обеспечивая научный интерес к себе в различных областях, в том числе и гуманитарных — семиотики, герменевтики, визуальной антропологии.

На сегодняшний день искусственный интеллект (ИИ) изучается современными учёными-филологами в различных аспектах. Например, в контексте осмысления влияния генеративного ИИ на медиаиндустрию [3], описания процесса представлений в теории искусственного интеллекта через сложные концептуальные структуры [4], автоматизированного перевода [5], преподавания [6], внедрения ИИ в сферу визуального искусства [7]. Также были найдены несколько работ, смежных с изучением вопросов смыслопорождения моделями ИИ, их последующего декодирования и интерпретации человеком. Так, в статьях [8; 9] рассматриваются вопросы искусственного семиозиса и смыслопорождения на примере текста, созданного ChatGPT 4o по простому запросу *Напиши, пожалуйста, сказку про дерево и листок*. В результате применения коммуникативного подхода с учетом семиотических и когнитивных процессов делается вывод о том, что подобного рода тексты, будучи созданными на основе обработки данных заранее заданными алгоритмами, не содержат авторского замысла и не являются самодостаточными объектами интерпретации. Отмечается, что ИИ может стать «помощником, эффективным инструментом, удобным подсобным средством», но «никогда не станет полноценным смыслообразующим агентом, подлинным семиотическим актором» [9. С. 27]. Изучаются также антропологические риски, связанные с развитием искусственного интеллекта. В статье [10] приводятся формы репрезентации антропологических вызовов ИИ: 1) научные концепты; 2) философская рефлексия; 3) визуальные искусства; 4) компьютерные игры; 5) кинематограф. Подчеркивается важность интеграции гуманитарной экспертизы в развитие ИИ. Помимо этого, модели ИИ рассматриваются и в контексте трендов современной 3D-графики. Отмечается влияние «генеративного дизайна»

[11. С. 24] на процессы производства изображений, что способствует к усилению фотореализма и художественной выразительности. В качестве перспективы использования ИИ-технологий называется их конвергенция с культурологическим и семиотическим подходами.

В настоящем исследовании мы интегрируем результаты, полученные коллегами [8–11], и подходим к ИИ как к инструменту создания текстов, а не как к мыслящему актору. Исследование проводится на границе трех форм антропологических вызовов ИИ — визуального искусства, кинематографа и философской рефлексии. Учитываются культурологические и семиотические аспекты использования генеративного дизайна при создании семиотически осложненных текстов с участием ИИ — в фокусе работы находится медиальная транспозиция, осуществляемая генеративными нейросетями. Рассматривается то, как смыслы исходного текста сначала декодируются, а затем собираются заново, что приводит как к проявлению скрытых смыслов, так и к приращению новых. Основной **целью** ставится изучение особенностей переноса печатного поэтического текста в медийный формат при помощи генеративных нейросетей, а также выявление специфических приёмов и черт происходящей при этом ремедиации исходного текста.

Объект настоящего исследования составляет процесс медиальной транспозиции линейного поэтического текста средствами генеративных нейросетей в форму коротких видеоклипов, **предмет** — происходящие при этом семиотические трансформации с исходным текстом.

Материал и методы исследования

В качестве **материала** исследования (табл. 1[1]) были выбраны работы нейроиллюстратора Никиты Глухова, известного под псевдонимом DJ БлокNote. Он является первым в русскоязычном секторе Интернет, кто стал создавать клипы на стихи русских поэтов. На форуме «Территория будущего. Москва 2030» 12 сентября 2025 г. DJ БлокNote провёл кибермузыкальное шоу (рис. 1), где выступил вместе с «ожившими» русскими классиками[2].

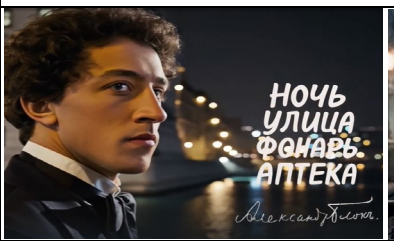
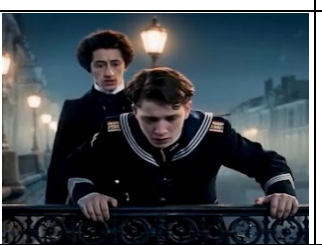
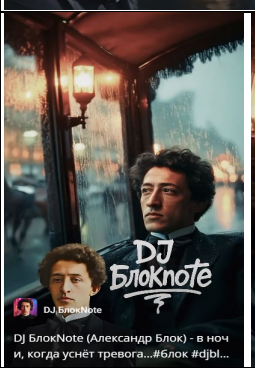
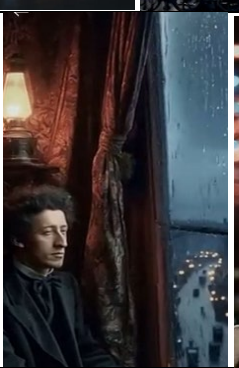
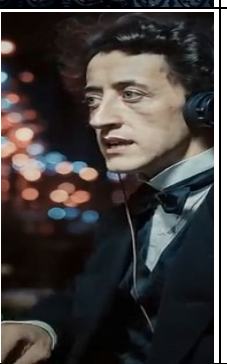
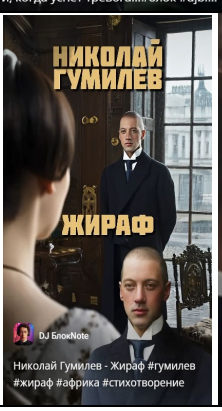
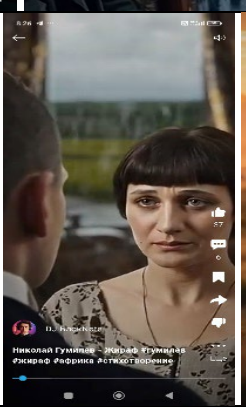
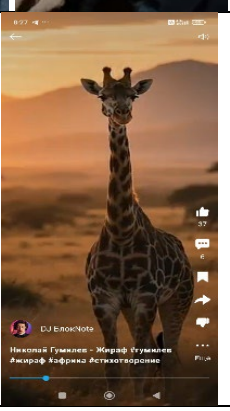
В работах Н. Глухова / DJ БлокNote обнаруживаются сложные семиотические трансформации, которым подвергаются исходные поэтические тексты. Согласно разработанному им алгоритму, процесс переноса стихотворения в медийный формат многоступенчат: сначала выбирается исходный текст, «рождается музыка, а потом уже визуальный ряд и только потом уже сценарий». Визуалу уделяется особое внимание — проделывается тщательная «работа с изображением: улучшение качества, конвертация

[1] *Прим.*: Произведения даны в табл. 1 в хронологической последовательности, в которой они были опубликованы на официальном канале DJ БлокNote на платформе RuTube. Режим доступа: https://rutube.ru/channel/67927732/ (дата обращения: 25.12.2025).

[2] Нейросети оживили поэзию Серебряного века на форуме «Территория будущего. Москва 2030». Режим доступа: https://fedpress.ru/news/77/society/3401566 (дата обращения: 03.10.2025).

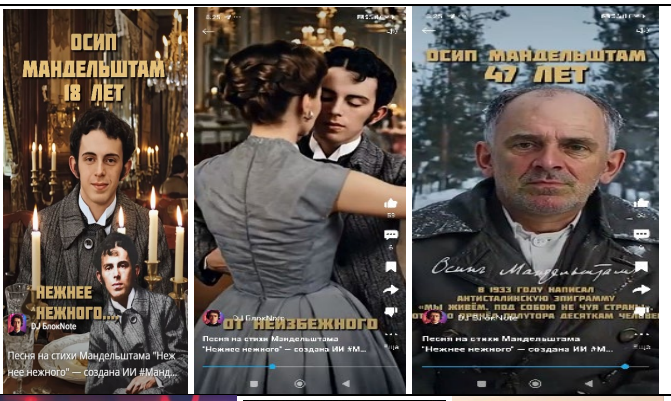
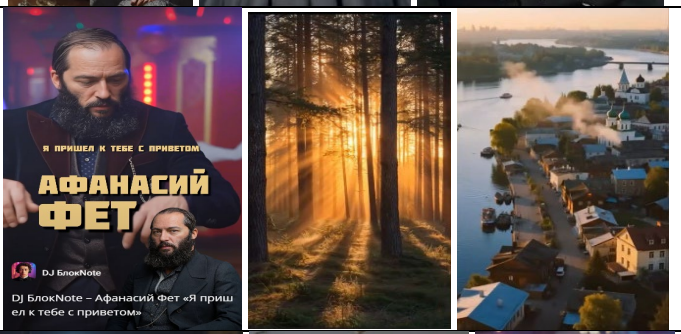
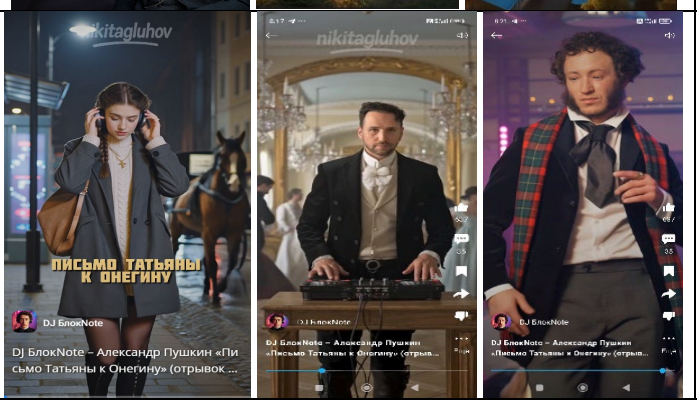
в реалистичный формат, перевод в модель для нейросети»[3]. Нейросеть в таком алгоритме, с одной стороны, направляется человеком-оператором, и, с другой стороны, визуализируя контекст исходного текста, вносит в него свои изменения расширяя интерпретацию. Это делает работы DJ БлокNote интересными для изучения вопросов ремедиации поэтического текста и передачи его множественности и лиричности в ходе медиальной транспозиции с участием нейросети.

Таблица 1 **Материал исследования**
Table 1. **Data for study**

№ п/п	Произведение / Artwork	Скриншоты из ИИ-клипа *(выборочно)* / Screenshots from the AI clip	Ссылка RuTube / Link RuTube
1.	Ночь, улица, фонарь, аптека (А. Блок)	 	
2.	В ночь, когда уснет тревога (А. Блок)	  	
3.	Жираф (Н. Гумилёв)	  	

[3] DJ БлокNote — Серебряный век / antimantikora // LIVEJOURNAL. 8 августа 2025. Режим доступа: https://dzen.ru/a/aHmOstJg_mis2ZQd (дата обращения: 15.12.2025).

Окончание табл. 1 / Ending of the Table 1

№ п/п	Произведение / Artwork	Скриншоты из ИИ-клипа *(выборочно)* / Screenshots from the AI clip	Ссылка RuTube / Link RuTube
4.	Нежнее нежного (О. Мандельштам)		
5.	Я пришел к тебе с приветом (А. Фет)		
6.	Письмо Татьяны к Онегину (А. Пушкин)		

Источник: исследование Ю.А. Евграфовой.
Source: compiled by Yulia A. Yevgrafova

Данное исследование реализует *интегративный подход* и осмысляет отобранный материал с позиций интеракционизма, герменевтики, психологической и визуальной антропологии, лингвосемиотики. Мы исходим из того, что человек существует в культуре — динамичной среде [12–18], используя ее артефакты в своей деятельности для взаимодействия с материальной средой

и социальной интеракции. Она фундирована вербальной и невербальной системой знаков, текстов, мифологем и идеологем, которую можно семиотически понять, расшифровать, деконструировать [19]. Создаваемые в ходе практической деятельности человека в медиасфере аудиовизуальные продукты понимаются нами как семиотически осложненные тексты, образованные звуковой и зрительной материей как целостное произведение с целью хранения и передачи сконструированных образов реципиенту, которые можно «прочесть».

Рис. 1. Выступление DJ БлокNote на форуме Москва 2030[4]
Источник: исследование Ю.А. Евграфовой.
Figure 1. DJ БлокNote's performance at the forum Moscow 2030
Source: compiled by Yulia A. Yevgrafova

В ходе исследования используются качественные **методы**: интерпретация, анализ, синтез, описательный метод, метод контекстуального анализа, функционально-прагматический метод, элементы психоаналитического метода для извлечения скрытых смыслов. Также для критического анализа и интерпретации семиотически осложненных текстов применяется трехуровневый анализ, разработанный Р. Бартом [20] и адаптированный под нужды исследования. Единицей анализа выступает аудиовизуальный образ без расщепления его на значимые изоляты. Сначала ИИ-клип описывается как физический объект, для чего составляется описание его материи (звуковой и зрительной). В следующем шаге рассматривается денотативный уровень ИИ-клипа, выделяются аудиовизуальные образы, далее проводится их инвентаризация с текстуальной составляющей стихотворения, затем определяются ключевые элементы и опорные семантические точки. После этого рассматривается коннотативный уровень, фиксируются дополнительные значения, в том числе созданные на пересечении изобразительного и звукового ряда ИИ-клипа и текста стихотворения. Далее следует интерпретация и критический анализ

[4] Режим доступа: https://www.osnmedia.ru/kultura/v-moskve-pri-podderzhke-radio-dfm-proshyol-kontsert-dj-bloknote-kibermiks/ (дата обращения: 15.12.2025).

обособленных образов, выделяются ключевые аспекты осмысления поэтических текстов средствами генеративных нейросетей, которые и представлены в данной статье.

Отметим, что в ходе исследования разграничиваются понятия ‘визуальный’ и ‘зрительный’, ‘аудиальный’ и ‘звуковой’. Когда мы рассматриваем материал как физический объект, то говорим о его материальных характеристиках — зрительных и звуковых, которые могут быть восприняты органами чувств в процессе просмотра и прослушивания. Когда мы переходим к рассмотрению способов передачи и организации того, *что* мы видим и слышим, и того, *как* это организовано, то говорим о визуальном и аудиальном компоненте соответственно. Когда мы рассматриваем материал как медиапродукт (экранизацию), то выделяем в нем изобразительный ряд и звукоряд.

Медиальная транспозиция: изменение внутренних свойств текста

В настоящем исследовании медиальная транспозиция рассматривается как одна из форм реализации интермедиальности — когда границы между носителями информации, средами или медиумом («medium») пересекаются [21]. Под медиальной транспозицией вслед за И. Раевски и Г.В. Зыковой в данной работе понимается перенос произведения из одного медиума в другой [21; 22] как, например, экранизация произведения.

Если говорить об экранизации, выполненной нейросетью, то она характеризуется тем, что в ней изобразительный ряд «идёт» от семантики, тогда как в классической он зависит от сюжета. Линейный печатный текст переносится в цифровую среду. Смена медиума приводит к изменению внутренних свойств текста. Если традиционный поэтический текст фиксируется на странице в печатной форме вербальными знаками, обладает статичностью, линейностью, ритмом, рифмой, многозначен, то при переносе его в экранную форму он теряет линейность и становится динамичным пространственно-временным континуумом, в основе которого лежит зрительная и звуковая материя, организованная и смонтированная определенным образом.

Исходя из этого, медиальная транспозиция, осуществляемая нейросетью, представляет собой не простую нейроиллюстрацию, а сложную реструктуризацию, при которой вербальному компоненту положен свой визуальный аналог. Опишем этот процесс и рассмотрим ниже конкретные примеры из отобранного материала.

Примеры медиальной транспозиции в работах нейроиллюстратора DJ БлокNote

Приведем отобранный материал исследования в форме таблицы, в которой кратко отобразим сюжет оригинального текста и экранизации, тезисно охарактеризуем ее изобразительный и звуковой ряд (табл. 2).

Таблица 2 **Сводные характеристики материала исследования**
Table 2. **Composite characteristics of the data for study**

№ п/п	Произведение / Artwork	Сюжет оригинального текста / Plot of the original text	Сюжет ИИ-клипа / The plot of the AI clip	Изобразительный ряд / Pictorial series	Звукоряд / Scale
1.	*Ночь, улица, фонарь, аптека (А. Блок)*	Герой размышляет о жизни, смерти, цикличности существования	Герой (А. Блок) видит самоубийцу (юношу) на мосту. Спасает его и отводит в аптеку, где дает лекарство. Юноша перерождается. А. Блок встречает его уже в С.-Петербурге XX в.	Изображено место действия — г. С.-Петербург конец XIX в. и конец XX в. Крупные и дальние планы героев, детальное изображение мимики. Ночной городской пейзаж. Одежда юноши после перерождения меняется на современную форму, А. Блок же остается в своем костюме	Музыка в стиле синтвейв (трипхоп, дарк-эмбиента и IDM). Стихотворный текст поет мужской голос. В начале слышны несколько выкриков юноши
2.	*В ночь, когда уснет тревога (А. [illegible]*	Герой выражает восхищение окружающим миром, полагая его совершенным творением Всевышнего	Герой (А. Блок) едет в дилижансе по сумеречному дождливому городу, стоит / сидит у окна и смотрит на дождливый город	Крупные и дальние планы героев, детальное изображение мимики	Музыка в стиле синтвейв. Стихотворный текст поет мужской голос
3.	*Жираф (Н. Гумилёв)*	Герой рассказывает своей возлюбленной о далекой Африке и о жирафе, гуляющем вдоль озера Чад. Он пытается отвлечь ее от грустных переживаний	Герой (Н. Гумилёв) рассказывает А. Ахматовой про Африку. Та плачет. Он открывает ей дверь, за которой оказывается Африка. Они стоят рядом и оба смотрят на пейзаж	Крупные и дальние планы героев, детальное изображение мимики. Кадры С.-Петербурга чередуются с пейзажами Африки и гуляющего по саванне жирафа, кадрами пещеры, чернокожего вождя и чернокожей девушки, сменами дня и ночи	Музыка в стиле дипхаус. Стихотворный текст поет мужской голос
4.	*Нежнее нежного (О. Мандельштам)*	Герой восхищается внешними данными и духовным миром своей возлюбленной и говорит о ее незаурядности	Герой (О. Мандельштам) едет с девушкой в карете по зимнему городу, они ведут разговор, который продолжают в ресторане. Танцуют. Потом гуляют по городу	Зимняя Москва (виден Кремль) и Санкт-Петербург (видна Петропавловская крепость). Крупные и дальние планы героев, детальное изображение мимики	Музыка в стиле дипхаус. Стихотворный текст поет мужской голос

Окончание табл. 2 / Ending of the Table 2

№ п/п	Произведение / Artwork	Сюжет оригинального текста / Plot of the original text	Сюжет ИИ-клипа / The plot of the AI clip	Изобразительный ряд / Pictorial series	Звукоряд / Scale
5.	*Я пришел к тебе с приветом (А. Фет)*	Герой приветствует свою возлюбленную рассказом о своих переживаниях утра	Герой (А. Фет) утром отправляется в лес, далее приходит к своей возлюбленной с букетом цветов, а затем уходит	Крупные планы героев. Утренний городской пейзаж чередуется с кадрами утра в лесу. Много солнечного света. А. Фет оказывается в ночном клубе, где танцует под «свою» музыку	Музыка в стиле синтвейв. Стихотворный текст поет мужской голос
6.	*Письмо Татьяны к Онегину (А. Пушкин)*	Татьяна напрямую выражает свои чувства к Евгению Онегину и желает понять, что он чувствует к ней	Героиня (Татьяна) переносится из современного города, надев наушники и начав слушать положенный на музыку стих. В интерьерах дворянской усадьбы она пишет / играет на рояле / поет свое письмо	Показаны интерьеры провинциальной дворянской усадьбы и бальная зала, деревенский пейзаж (двор, хлев). Особо выделены предметы: самовар, православный крестик, рушник у иконы, православная икона, старушка-няня, пейзаж (березки), колодец, коромысло	Музыка в стиле дипхаус. Стихотворный текст поет женский голос

Источник: исследование Ю.А. Евграфовой.
Source: compiled by Yulia A. Yevgrafova

Остановимся на каждом случае в отдельности.

Кейс № 1: «Ночь, улица, фонарь, аптека» (А. Блок)

В оригинальном тексте стихотворения выражается уверенность в безысходности и «зацикленности» всего сущего через использование таких лексических единиц (ЛЕ), как *бессмысленный*, *тусклый*, *исхода нет*, *начнешь опять сначала*, *повторится все*, *как встарь*, а также через образ ночного пейзажа: мрачной и холодной городской набережной, выстраивающийся ЛЕ *ледяная рябь канала*, *ночь*, *улица*, *фонарь*, *аптека.* При интерпретации произведения DJ БлокNote в сотрудничестве с нейросетью развивает смысл лирического сюжета через визуальное повествование. Оно начинается с того, что А. Блок прогуливается по ночной набережной Санкт-Петербурга конец XIX в. (что ясно из архитектуры и вывесок), видит юношу-моряка, который готов вот-вот утопиться. А. Блок спасает его, отводит в аптеку, где дает некое лекарство. Юноша перерождается — зритель видит его счастливого, гуляющего с девушкой по набережной уже современного Санкт-Петербурга.

Экранизация произведения привела к тому, что рефлексия о бытии, вокруг которой оно выстроено, трансформируется в конкретное событие, участие в котором принимает сам автор. Драматичность оригинального сюжета визуализируется через персонажа-самоубийцу, спасение и счастливое перерождение которого в современном Санкт-Петербурге есть преодоление безысходности, изначально заложенной в текст. В этом смысловом узле ИИ-клипа цикл «несчастного» существования разорван, предлагается вариант *исхода*, которого, на самом деле, *нет*.

Кейс № 2: «В ночь, когда уснет тревога» (А. Блок)

Стихотворение А. Блока «В ночь, когда уснет тревога» передает созерцание им мира, восхищение гармонией и *музыкой бога. Буря жизни* автором противопоставляется *розам*, которые *цветут* и *горят*, а *человеческие слезы — румянящемуся закату»*. Выстраивается образ закатного неба, затихающего буйства природы, готовящейся ко «сну». Все мирское ничто по сравнению с тем божественным замыслом, который можно увидеть в природе.

Интересно, что «экранное» осмысление этого произведения DJ БлокNote и нейросетью имеет иной акцент. Возвышенное переживание божественного заменяется на демонстрацию того, как автор, А. Блок, едет в дилижансе по вечернему дождливому городу и, приехав домой, задумчиво наблюдает дождливые сумерки из окна. Эти два действия можно рассматривать не как цепочку конкретных событий, а как визуальную метафору настроения. Судя по всему, ЛЕ *тревога*, *мгла, слезы*, *кровь*, *муки* были интерпретированы нейросетью буквально, без учета окружающего контекста, в результате чего оппозиция мирское–Божественное, на основе которой выстроено произведение, «потерялась». При экранизации сложная абстракция трансформировалась в простую демонстрацию меланхолии.

Кейс № 3: «Жираф» (Н. Гумилёв)

В оригинальном тексте «Жирафа» Н. Гумилёв беседует со своей возлюбленной (многие исследователи его творчества склонны думать, что стихотворение адресовано А. Ахматовой, первой жене поэта) о своей поездке в Африку, им упоминается *озеро Чад*, *тропический сад*, *пальмы*, *травы*, *черная дева*, *вождь*, *изысканный жираф,* описывается его *грациозная стройность*, *шкура*, *плавный бег*. Параллельно с «Африкой» в стихотворении обрисованы напряженные отношения между рассказчиком и слушательницей, которая сидит *колени обняв*, а *взгляд* ее *грустен*, и, по выражению рассказчика, она *слишком долго вдыхала тяжелый туман* и *верить не хочет во что-нибудь*, *кроме дождя.* Поэтому ему тяжело рассказывать ей про сказочную Африку.

Эти два сюжета экранизированы в ИИ-клипе буквально. Основное событие, вокруг которого выстроено визуальное повествование — это беседа Н. Гумилёва с плачущей А. Ахматовой, то сидящей в кресле, то стоящей напротив поэта. Детализировано место действия — интерьер дворянской квартиры конец XIX — начало XX в. — и костюмы героев, мужской фрак и

дамское платье того же периода. Кадры разговора чередуются с кадрами саванны, жирафа, африканского племени. В конце видео Н. Гумилёв открывает двери комнаты и оба смотрят на открывшийся африканский пейзаж. Это действие, во-первых, визуализирует совместную победу над грустью адресата, что расширяет семантику произведения, и, во-вторых, трансформирует метафорический рассказ о «сказочной» Африке в нечто реальное. Такой ход в визуальном повествовании ИИ-клипа отражает стремление акмеизма, основателем и идеологом которого являлся Н. Гумилёв, к конкретным «вещественным» образам.

Кейс № 4: «Нежнее нежного» (О. Мандельштам)

О. Мандельштам любуется своей возлюбленной, восхищается ее внешней и внутренней красотой, употребляя такие эпитеты, как *нежнее нежного*, *белее белого*, *неостывающие пальцы рук*, *неунывающие речи*, *даль очей*. DJ БлокNote и нейросеть «переводят» лирическое чувство в формат романа, зрителям показан его фрагмент — прогулка в карете, свидание в ресторане, танцы, чтение стихов. Детализировано изображено место действия — карета, зимний город, ресторан, костюмы персонажей, транслируя «аристократичный дух» Санкт-Петербурга. Созданные образы поддерживают лиричность, транслируемую текстом стихотворения.

В конце ИИ-клипа нейроиллюстратором DJ БлокNote добавлено фото О. Мандельштама, сделанное в 1938 г., перед смертью в пересыльном лагере, что добавляет оригинальному стихотворению, изначально писавшемуся молодым О. Мандельштамом как лирическое восхищение возлюбленной, трагичности и драматизма.

Кейс № 5: «Я пришел к тебе с приветом» (А. Фет)

В стихотворении «Я пришел к тебе с приветом» А. Фет выражает свою восторженность утром и пробуждающейся природой, детально ее описывая и оживляя: *солнце встало*, *горячим светом по листам затрепетало*, *лес проснулся веткой каждой, каждой птицей встрепенулся*, *весенней полон жажды.* Он пришел к своей возлюбленной, чтобы поделиться этими переживаниями: *душа все так же счастью / И тебе служить готова; / <…> отовсюду / На меня весельем веет, / <…> не знаю сам, что буду / Петь — но только песня зреет*.

При экранизации произведения его лирический сюжет претерпел трансформацию в визуальный нарратив с разделением на три хронотопа: 1) А. Фет в современном ночном клубе исполняет свое стихотворение, положенное на музыку, за диджейским пультом и у микрофона; 2) пейзаж утреннего леса, полного света; 3) А. Фет пришел с цветами к своей возлюбленной. Смысловой стержень оригинального текста – интенсивное эмоциональное состояние, страсть и восторг, переданы в ИИ-клипе через обилие света и мимику А. Фета. Эмоциональный заряд также несут в себе кадры ночного клуба, выступающего в нем поэта-диджея А. Фета, кадры исполнения им песни, а затем и танца под нее.

Кейс № 6: «Письмо Татьяны к Онегину» (А. Пушкин)

Отрывок «Письмо Татьяны к Онегину» из романа в стихах «Евгений Онегин» А. Пушкина представляет собой пылкое признание Татьяны, главной героини, в любви к Евгению Онегину. DJ БлокNote выбрал для экранизации первую половину письма до слов *То воля неба: я твоя* включительно. Визуальное повествование представлено тремя хронотопами: 1) Татьяна в современной Москве; 2) Татьяна пишет письмо Евгению; 3) Евгений в исполнении Ильи Глухова на балу и в деревне. В ИИ-клип включены персонажи: А. Пушкин (гуляет по современной Москве, где встречает Татьяну), Татьяна (надевает наушники и оказывается в хронотопе романа), Евгений Онегин (в исполнении Ильи Глухова на балу — диджеит, в деревне — показан неуклюжим, пытающимся познакомиться с деревенским бытом).

Экранизация насыщена артефактами русской культуры — самовар, прическа героини — коса, интерьер провинциальной дворянской усадьбы, бальная зала, православный крестик, рушник у иконы, православная икона, мужское и женское платье начала XIX в., старушка-няня, пейзаж (березки), колодец, коромысло, деревенский пейзаж (двор, хлев). Благодаря этому визуализируется историко-культурный контекст «Письма...», а ИИ-клип «насыщается» смыслами, относящимися к центральной зоне русской культуры. Разрушение временн***о***го барьера (Татьяна перемещается в «прошлое», надев наушники) объединяет прошлое и настоящее, маркируя современность описываемых в «Письме ...» чувств, их «вечность» и актуальность.

Анализ рассмотренных выше кейсов позволяет прийти к следующим промежуточным выводам.

Во-первых, при медиальной транспозиции с участием генеративной нейросети в стихотворениях *вычленяется объект(-ы) и ряд действий*, которые *помещаются в определенные ситуации*, связанные с идейным содержанием стихотворений, аналогиями с ним и/или возможными его альтернативами. Это создает условия для образования неограниченного количества значений в сгенерированном медийном продукте. Нейросеть при участии оператора DJ БлокNote *овеществляет символы* и *«обрамляет»* их. Так, в Кейсе 1 отчаяние и безысходность переданы через визуальное повествование о самоубийстве и перерождении юноши, преодолевающего безысходность и разрывающего цикличность через спасение, что расширяет семантику стихотворения. Примером «обрамления символов» служит и Кейс 2, в котором образ Африки «материализуется» и вводится в диегетическое пространство героев клипа, когда поэт открывает дверь.

Во-вторых, в результате медиальной транспозиции чувства и эмоциональная экспрессия стихотворений «переводятся» нейросетью и ее оператором в *динамическую культурно созданную целостность*. В ней обнаруживается детализация изображения места действия (Москва, Санкт-Петербург, провинциальная усадьба, городские апартаменты), обрисовываются фрагменты быта и частной жизни конкретного исторического периода (XIX в. или

XX—XXI вв.). Визуальное повествование организовывается вокруг конкретных образов как, например, в Кейсе 2 — это дождливый сумеречный город, в Кейсе 4 — ресторан и зимний город, в Кейсе 5 — лес и город на заре, ночной клуб. При внедрении стихотворения в такой динамический «культурно целый» *овеществленный контекст* культурные значения и индивидуальный опыт зрителя вступают в динамические отношения – зритель не воспринимает транслируемую символику как символ, но как валентность, воспринимая ее как «личное значение» и соотнося со своим опытом. При постижении нейросетевого произведения на первый план выходят жизненные ассоциации, художественные ассоциации возникают на их основе.

В-третьих, в ИИ-клипах *«сталкиваются» артефакты разных эпох*, например костюм А. Блока XIX в. и форма юноши-моряка XX в. в Кейсе 1; ночной клуб XXI в. и интерьеры апартаментов XIX в. в Кейсе 5; наушники и Москва XXI в. и детали быта XIX в. (икона, рушник, самовар и т.д.) в Кейсе 6. Обнаруживаемые в экранизациях маркеры разных культурно-исторических контекстов принадлежат «центральной зоне» русской культуры, что добавляет в текст семиотического напряжения и создает основу для конструирования не только индивидуальных, но и интерсубъективных значений.

В-четвертых, отдельно следует подчеркнуть *многоканальность воздействия* в сгенерированных клипах. Звуковой и изобразительный ряд взаимодействуют по принципу *контрапункта* — текст стихотворения, положенный на современную электронную музыку, и визуальное повествование, сгенерированное на его основе, «существуют» вместе и одновременно, дополняя друг друга. Так, в Кейсе 1 к опорным точкам текста *бессмысленный, тусклый, исхода нет, начнешь опять сначала, повторится все, как встарь* добавляется нарратив спасения и преодоления безысходности. В Кейсе 3 музыкальное сопровождение (имеется ввиду синтез специальной звуковой материи — дипхауса, основным инструментом в котором является электроорган, и материи звучащего слова — стихотворения) характеризуется глубокой, плавной и утонченной элегантной подачей, что семантически поддерживает экзотические образы соответствующего изобразительного ряда.

Средства переосмысления поэтического текста генеративной нейросетью

В ходе исследования были обнаружены следующие средства, использованные нейросетью для переноса литературного произведения из одной среды в другую — из линейного печатного медиума в пространственно-временн***о***й континуум экранного медиума.

Прежде всего, это конверсия мономодального текста в полимодальный текст. Новое произведение дополняется невербальными единицами: зрительными и слуховыми. Зрительные единицы — это конкретные визуальные образы, овеществляющие метафору лирического произведения. Слуховые единицы — это специальная звуковая материя (в случае с рассматриваемым

материалом — современная электронная музыка) и материя звучащего слова — стихотворения. Обнаруживается *детализация визуальных образов*, а именно: крупные планы действующих в визуальном повествовании лиц, детализированное изображение интерьеров, пейзажей, единичных материальных артефактов культуры.

Нейросеть использует исходный стихотворный текст как источник для построения нового повествования и нового смысла. Так, в ходе *нарративизации* в стихотворение «Ночь, улица, фонарь, аптека» (А. Блок) добавляется надежда на спасение, в стихотворение «Жираф» (Н. Гумилёв) – надежда на счастливое разрешение размолвки, а стихотворение «В ночь, когда уснет тревога» (А. Блок) полностью меняет свой замысел, транслируя тревожность и меланхоличность. В созданное визуальное повествование включается *несколько хронотопов*, а иногда и *анахронизм*, как, например, в случае с «Ночь, улица, фонарь, аптека» (А. Блок) или «Письмо Татьяны к Онегину» (А. Пушкин), где культурно-исторические контексты сознательно смещены для того, чтобы подчеркнуть «вечность» затрагиваемых тем.

Автор лирического произведения «оживляется» нейросетью и *включается визуальный нарратив*, принимая непосредственное участие в происходящем. Причем оживленное изображение поэта, вставленное в ИИ-экранизацию стихотворения, не соответствует возрасту, в котором он его создал, а является его «популярным» изображением, под которым он известен массовой аудитории.

Проблемы нейросетевой медиальной транспозиции и вопросы авторства

Медиальная транспозиция поэтического текста имеет ряд проблемных моментов. Прежде всего, это *риск утраты исходных смыслов*, заложенных в произведение. Слово лишается своей дискретности при переходе линейного печатного текста в экранный континуум. Существует вероятность, что при трансформации в аудиовизуальные образы вербальный компонент останется без изначально заложенной в него множественности значений в угоду создания общего «настроения» и атмосферы. Кроме того, *опасность* полного изменения смысла произведения в ИИ-экранизации может быть из-за, во-первых, *предельно буквальной экранизации*, как, например, в случае с «В ночь, когда уснет тревога» (А. Блок), когда опорные точки стихотворения были интерпретированы нейросетью буквально, во-вторых, из-за *добавления* оператором нейросети *новых деталей*, как, например, в случае «Письмо Татьяны к Онегину» (А. Пушкин) — Татьяна переносится из настоящего времени в прошлое, «Я пришел к тебе с приветом» (А. Фет) — А. Фет танцует в ночном клубе или «Нежнее нежного» (О. Мандельштам) — фото О. Мандельштама перед смертью в конце ИИ-клипа заставляет зрителя полностью переосмыслить оригинальный текст, подчеркивая горестные и мрачные годы жизни поэта, и, в-третьих, из-за *добавочной интерпретации*, как, например, в случае

«Ночь, улица, фонарь, аптека» (А. Блок), когда стихотворение превращается в мини-драму, где поэт становится всевидящим вершителем судьбы.

Чрезмерное использование визуального компонента может привести к *переиллюстрированности* и превратить ИИ-экранизацию в *китч*. Гиперреалистичная картинка способствует утрате поэтическим текстом ритма, множественности, метафоричности. Гиперреалистичный и «банальный» визуальный идиостиль ИИ может вытеснить идиостиль автора оригинала, снизив сложное произведение до привлекательной «картинки».

Сотрудничество оператора и нейросети ставит вопрос авторства нового произведения. Следует *четко разграничивать роли*: автор (поэт) — человек-оператор — нейросеть. Нейросеть — это машина, которая создает материал по запросу оператора. Оператор — это человек (для подобной деятельности уже существует новое название — нейроиллюстратор), который пишет промты и курирует нейросеть, ставит ей задачи, отбирает, редактирует и собирает воедино полученный материал. Автор — это создатель оригинального произведения. Порожденный его творческим трудом текст служит источником для написания промптов оператором. В этой связи необходима *прозрачная атрибуция*, т.е. указание на первоисточник, на конкретную нейросеть и на нейроиллюстратора. При таком *распределенном авторстве* оператор несет ответственность за итоговую интерпретацию поэтического текста.

Основным *критерием качества* нового произведения, созданного при нейросетевой медиальной транспозиции, является стилистическая цельность, соответствие эмоциональному ядру оригинального поэтического текста, целостность сгенерированных образов. Технической стороне следует уделять меньшее внимание.

Выводы

Проведенное исследование позволяет сделать следующие выводы относительно медиальной транспозиции с участием нейросетей.

Нейросетевая медиальная транспозиция — это перенос исходного линейного текста из печатной в нелинейную экранную форму средствами нейросети, оператором которой является человек. Промпты для нейросети пишутся на основе опорных семантических точек стихотворения — конкретных ЛЕ, с добавлением в них дополнительных ЛЕ, которые отсутствуют в исходном тексте, но которые нужны для визуализации. В рассмотренных работах нейроиллюстратора DJ БлокNote в ходе транспозиции помимо зрительной материи задействуется и звуковая — на изобразительный ряд накладывается специальная звуковая материя — современная электронная музыка, и материя звучащего слова — стихотворения. В результате происходит ремедиация, переосмысление исходного текста средствами зрительной и звуковой материи.

Целостное произведение разбивается на «данные». Для составления промпта исходный текст деконструируется сначала оператором для выделения опорных ЛЕ, а затем генеративной нейросетью, которая преобразует команду в токены, обрабатывает их и собирает «ответ», но уже в визуальный образ. Нейросеть действует как машина, перекодируя текстовую структуру, в случае настоящего исследования, поэтическую, в структуру визуальную, в которой ритм и метафоричность заменяется динамикой, монтажом, композицией, цветом и светом. Получившаяся ИИ-экранизация состоит из интермедиальных образов и представляет собой новый артефакт культуры, помещенный в экранную среду.

Нейросетевая медиальная транспозиция опирается на зрительно-слуховую ассоциативную работу психики потенциального зрителя. Включенный в экранизацию зримый облик вещественного окружения запускает зрительные ассоциации, раскрывая имплицитное (в нашем случае — это лирическое) на конкретном предметном «материале». Акустическое воздействие специальной звуковой материи (музыки) в сочетании с материей звучащего слова (стихотворения) подключает слухо-ассоциативную работу психики, связывая эмоции с акустикой.

Созданный текст сложен и обнаруживает не только аудиовизуальный, но и сюжетно-фабульный выразительный слой. В рассмотренных случаях нейросетевой медиальной транспозиции поэтических произведений обнаруживается нарративизация лирического. Находя визуальные соответствия опорным ЛЕ исходного текста, нейросеть конструирует визуальное повествование, через которое передается эмоциональный компонент.

При переносе исходного печатного текста в новый медиум нейросеть актуализирует культурные схемы, принадлежащие к «центральной зоне» культуры, что способствует конструированию не только сугубо «личностных», уникальных, но и интерсубъективных значений, понятных для членов общества в целом. В рассмотренных работах нейроиллюстратора DJ БлокNote была обнаружена актуализация культурных схем через обращение к стереотипным образам «русская помещичья усадьба», «Петербург Достоевского» — идеальным артефактам; через использование образов культурнозначимых предметов «самовар», «коса», «православный крестик», «православная икона», «мужское и женское платье XIX в.» и мн. др. — когнитивные схемы объектов; через изображение аристократического быта — модели и культурные сценарии. Знаки культуры «рождаются» повторно в новом медийном продукте, в котором воспроизводится их конвенциональное значение и создаются условия для конструирования идиосинкретичного.

Заключение

Рассмотренные работы нейроиллюстратора DJ БлокNote являются примером не только вторичного семиотического произведения, созданного в соавторстве с «машиной», но и одного из первых шагов в диалоге между

традиционными и новыми формами искусства. Нейросети, перенося литературное произведение из линейной печатной формы в экранную среду, служат инструментом его истолкования. Текст, созданный в процессе медиальной транспозиции, не угрожает, но обогащает традиционное прочтение «классики», в том числе благодаря кураторству человека-оператора.

Нейроиллюстратор DJ БлокNote, используя добавочный визуальный компонент — цвет, свет, композицию и др. — как ведущий элемент, актуализирующий исходное содержание, выстраивает в своих экранизациях визуальный нарратив. Важно, что созданный контент помещается им в формат коротких видео (Shorts). Этот формат понятен современной молодежи и позволяет компактно «упаковать» многослойное и сложное произведение, давая ему «второе рождение». Такой подход позволяет говорить о творчестве нейрохудожников как о самостоятельном художественном явлении, продолжающем традиции поэзии и литературного искусства, но на принципиально новом технологическом и эстетическом уровне.

Перевод вербального в визуальное, добавление музыкального сопровождения способствует проявлению имплицитных связей, слабовыраженных и/или неочевидных смыслов и метафор. Получившийся семиотически осложненный текст является продуктом сотрудничества (1) автора исходного текста, произведение которого «вдохновляет» (2) оператора нейросети, курирующего процесс генерации через написание промптов, и (3) нейросети — «машины», обрабатывающей промпты по определенным алгоритмам и выдающей ответы в формате визуальных образов. Такое сотворчество начинает новый этап в интерпретации текстов — герменевтику, исследующую продукты, созданные цепочкой автор–оператор–нейросеть, или герменевтику ИИ, в рамках которой уже можно сформулировать вопросы, требующие изучения и дальнейшего исследования.

Прежде всего, это проблема автора. Создание экранизаций с участием генеративных нейросетей «размывает» понимание авторства, порождая экранные тексты, в которых заложенные смыслы гибридны, несут отпечаток интенсионального мира человека-оператора, а также коллективного культурного корпуса, используемого генеративной моделью. Это вынуждает рассмотреть проблему наращивания смыслов после «смерти автора» по Р. Барту в новом, технологическом, аспекте. Также встают смежные с этим вопросы подлинности и оригинальности — созданный с участием генеративной нейросети экранный текст представляет собой новую форму искусства, является своего рода коллажем или, всё же, плагиатом? И второй — вопрос ответственности за сгенерированные смыслы, особенно это касается потенциально опасного и дискриминационного контента. Возникает проблема понимания герменевтического круга, который трансформируется в замкнутую цепочку запрос — ответ — новый запрос, в которой модель интерпретирует запрос человека-оператора, а он — ответ модели, что влияет на

его понимание о запрашиваемом вплоть до полного его изменения, что инициирует новые запросы — и так неограниченное количество раз.

Семейство генеративных нейросетей обучается на артефактах человеческой культуры — массиве данных, выступая своего рода её отражением и сжатым изложением. Изучение контента, сгенерированного нейросетями, позволит выделить семиотические и нарративные доминанты, стереотипы, существующие в культуре в конкретный момент времени. Интересно проследить, какие трансформации претерпевают культурные коды после прохождения через ИИ. И немаловажная проблема возникает в аспекте семиотики и философии языка — контент, созданный с участием ИИ, содержит в себе знаки в классическом их понимании, т.е. имеющие референта в реальности, или их следует считать копиями без оригинала — симулякрами? Эти и другие вопросы требуют дальнейшего исследования и научного описания.

В заключении, отметим, что изучение медиапродуктов, созданных с участием семейства генеративных нейросетей, в аспекте герменевтики, семиотики, психологической и визуальной антропологии имеет практическую значимость, поскольку позволит разработать алгоритм критического анализа сгенерированного контента на предмет выявления идеологий и манипулятивного компонента, заложенного в данные для обучения моделей, а также методологию интерпретации порожденных смыслов, уяснения их культурного генезиса, а также поможет сохранить роль агенса за человеком в процессе коммуникации с ИИ.

Список литературы

1. *Мамедова Н.М.* Человек в эпоху цифровизации: на грани реального и виртуального // Век глобализации. 2021. № 3. С. 74–85. EDN: RMIXIC
2. *Шапинская Е.Н.* Культурное наследие и вызовы цифровой эпохи: трансформация культурных форм и эстетических ценностей // Большая Евразия: развитие, безопасность, сотрудничество. 2019. № 2–2. С. 659–663.
3. *Малыгина Л.Е., Парфун А.В.* Теория подрывных инноваций как методологическая основа анализа трансформации телевизионного вещания в эпоху генеративного ИИ // Научные исследования и разработки. Современная коммуникативистика. 2025. Т. 14. № 5. С. 95–102. https://doi.org/10.12737/2587-9103-2025-14-5-95-102 EDN: YHPEWG
4. *Манерко Л.А.* Пропозиция, фрейм и когнитивная карта как ступени понимания человеческой когниции и ИИ // Когнитивные исследования языка. 2024. № 5(61). С. 115–120. EDN: OLDEHS
5. *Джунковский А.В.* Искусственный интеллект как новая форма CAT. Путь к идеальной автоматизации перевода или новый вызов? // Речевые технологии. 2024. № 1. С. 14–17. EDN: VARJMR
6. *Копылова Е.В.* Конструирование процесса обучения иностранному языку средствами ИИ // Актуальные проблемы германистики, романистики и русистики. 2025. № 1. С. 132–137. EDN: PDQUQE
7. *Хворостов В.В.* Основные проблемы внедрения ИИ в сферу искусства: по материалам публицистики 2018–2023 годов // Социология искусственного интеллекта. 2023. Т. 4. № 4. С. 31–38. EDN: CIZFEA

8. *Вдовиченко А.В.* Искусственный интеллект написал сказку: имитация воздействия текстом на фоне естественного словосодержащего семиозиса (часть первая) // Вестник Православного Свято-Тихоновского гуманитарного университета. Серия 3: Филология. 2024. № 80. С. 23–33. https://doi.org/10.15382/sturIII202480.23-33 EDN: STULKF
9. *Вдовиченко А.В.* Искусственный интеллект написал сказку: имитация воздействия текстом на фоне естественного словосодержащего семиозиса (часть вторая) // Вестник Православного Свято-Тихоновского гуманитарного университета. Серия 3: Филология. 2025. № 82. С. 9–29. https://doi.org/10.15382/sturIII202582.9-29. EDN: KEBCLN
10. *Сидорова Т.А.* Образы восприятия и концептуализация антропологических вызовов искусственного интеллекта // Праксема. Проблемы визуальной семиотики. 2024. № 1(39). С. 102–119. https://doi.org/10.23951/2312-7899-2024-1-102-119 EDN: GQXASO
11. *Архипова Н.А.* Современные тенденции в 3D-графике и анимации: технологические инновации и художественная выразительность в игровой индустрии // Вестник Московского международного университета. 2025. № 3 (7). С. 23–26. EDN: NACOWR
12. *D'Andrade R.G.* The Development of Cognitive Anthropology. Cambridge University Press, 1995.
13. *D'Andrade R.G.* The Cultural Part of Cognition // Cognitive science. 1981. № 5. P. 179–195.
14. *Cole M.* Cultural psychology: A once and future discipline. Cambridge, Massachusetts, and London, 1996.
15. *Лотман Ю.М.* Культура и взрыв. М. : Гнозис; Прогресс, 1992.
16. *Лурье С.В.* Культура и ее сценарий: имплицитный обобщенный сценарий как внутрикультурный интегратор // Общественные науки и современность. 2017. № 2. С. 153–164. EDN: YFUNZH
17. *Лурье С.В.* Обобщенный культурный сценарий и функционирование социокультурных систем // Журнал социологии и социальной антропологии. 2010. Т. 13. № 2. С. 152–167. EDN: NCFHJX
18. *Лурье С.В.* Что такое культурная картина мира и как формируется культурная традиция // Тетради по консерватизму. 2021. № 2. С. 13–62. https://doi.org/10.24030/24092517-2021-0-2-13-62
19. *Дорофеев Д.Ю., Марков Б.В.* Визуальная антропология и коммуникация: философский взгляд. СПб. : Изд-во РХГА, 2022.
20. *Barthes R.* S/Z. Blackwell, 1990.
21. *Rajewsky I.O.* Intermediality, Intertextuality, and Remediation: A Literary Perspective on Intermediality. Intermédialités // Intermediality. 2005. No. 6. P. 43–64. https://doi.org/10.7202/1005505ar
22. *Зыкова И.В.* Интермедиальность в моделировании поэтики кинодискурса // Зборник Матице српске за славистику. 2022. № 102. P. 251–268. https://doi.org/10.18485/ms_zmss.2022.102.16 EDN: QCYFRQ

References

1. Mamedova, N.M. (2021). Humans in the Age of Digitalization: on the Verge of Real and Virtual. *Age of Globalization,* 3, 74–85. (In Russ.). EDN: RMIXIC
2. Shapinskaja, E.N. (2019). Cultural Heritage and the Challenges of the Digital Age: The Transformation of Cultural Forms and Aesthetic Values. *Great Eurasia: Development, Security, Cooperation*, 2–2, 659–663. (In Russ.).

3. Malygina, L.E., & Parfun, A.V. (2025). Disruptive Innovation Theory as a Methodological Framework for Analyzing the Transformation of Television Broadcasting in the Era of Generative AI. *Scientific Research and Development. Modern Communication Studies*, *14*(5), 95–102. (In Russ.). https://doi.org/10.12737/2587-9103-2025-14-5-95-102 EDN: YHPEWG
4. Manerko, L.A. (2024). Proposition, Frame and Cognitive Mapping as Stages in Understanding of Human Cognition and Artificial Intellect. *Cognitive Studies of Language*, *5*(61), 115—120. (In Russ.). EDN: OLDEHS
5. Dzhunkovskiy, A.V. (2024). AI as a New Type of Computer-Assisted Translation. A road paved with good intentions. *Speech Technology*, 1, 14–17. (In Russ.). EDN: VARJMR
6. Kopylova, E.V. (2025). Designing a Foreign Language Teaching Process Using AI. *Current Issues in Germanic, Romance, and Slavic Studies*, 1, 132–137. (In Russ.). EDN: PDQUQE
7. Khvorostov, V.V. (2023). The Main Problems of Introducing AI into the Field of Art: Based on Journalistic Materials from 2018-2023. *Sociology of Artificial Intelligence*, *4*(4), 31–38. (In Russ.). EDN: CIZFEA
8. Vdovichenko, A. (2024). Artificial Intelligence Wrote a Fairy Tale: Simulation of Text Impact Against the Background of Natural Word-Containing Semiosis. *St. Tikhon's University Review. Series III: Philology*, 80, 23–33. (In Russ.). 10.15382/sturIII202480.23-33 EDN: STULKF
9. Vdovichenko, A. (2025). Artificial Intelligence Wrote a Fairy Tale: Simulation of Text Impact Against the Background of Natural Word-Containing Semiosis. *St. Tikhon's University Review. Series III: Philology*, 82, 9–29. (In Russ.). https://doi.org/10.15382/sturIII202582.9-29 EDN: KEBCLN
10. Sidorova, T.A. (2024). Perception Images and Conceptualization of Anthropological Challenges of Artificial Intelligence. *ΠΡΑΞΗΜΑ. Journal of Visual Semiotics*, *1*(39), 102–119. (In Russ.). https://doi.org/10.23951/2312-7899-2024-1-102-119 EDN: GQXASO
11. Arkhipova, N.A. (2025) Current Trends in 3D Graphics and Animation: Technological Innovation and Artistic Expression in the Gaming Industry. *Bulletin of the Moscow International University*, *3*(7), 23–26. (In Russ.). EDN: NACOWR
12. D'Andrade, R.G. (1995). *The Development of Cognitive Anthropology*. Cambridge: Cambridge University Press.
13. D'Andrade, R.G. (1981). The Cultural Part of Cognition. *Cognitive Science*, 5, 179–195.
14. Cole, M. (1996). *Cultural Psychology: A Once and Future Discipline*. Cambridge, Massachusetts, and London.
15. Lotman, Ju.M. (1992). *Culture and Explosion*. Moscow: Gnosis; Progress. (In Russ.).
16. Lourie, S. (2017). The Culture and Its Script. Implicit Generalized Script as The Intracultural Integrator. *Social Sciences*, 2, 153–164. (In Russ.). EDN: YFUNZH
17. Lurye, S. (2010). General Cultural Script and Socio-Cultural Systems Functioning. *Journal of Sociology and Social Anthropology*, 13 (2), 152–167. (In Russ.). EDN: NCFHJX
18. Lourie, S.V. (2021). What is Cultural World Image and How Cultural Tradition is Formed. *Writings on Conservatism*, 2, 13–62. (In Russ.). https://doi.org/10.24030/24092517-2021-0-2-13-62 EDN: DVDMQD
19. Dorofeev, D.Ju., & Markov, B.V. (2022). *Visual Anthropology and Communication: A Philosophical Perspective.* Saint Peterbourg: RHGA publ. (In Russ.).
20. Barthes, R. (1990). *S/Z.* Blackwell.
21. Rajewsky, I.O. (2005). Intermediality, Intertextuality, and Remediation: A Literary Perspective on Intermediality. Intermédialités. *Intermediality*, 6, 43–64. https://doi.org/10.7202/1005505ar

22. Zykova. I.V. (2022). Intermediality in Modeling the Poetics of Cinematic Discourse. *Zbornik Matice Srpske za Slavistiku-Matica Srpska Journal of Slavic Studies*, 102, 251–268. (In Russ.). https://doi.org/10.18485/ms_zmss.2022.102.16 EDN: QCYFRQ

Сведения об авторе:
Евграфова Юлия Александровна, доктор филологических наук, доцент, профессор кафедры рекламы, связей с общественностью и лингвистики, Национальный исследовательский университет «МЭИ» (111250, Российская Федерация, г. Москва, ул. Красноказарменная, д. 14, стр. 1); *сфера научных интересов*: семиотика, семантика, киносемиотика, поликодовые тексты, герменевтика; *e-mail:* yu.evgrafova@mail.ru
ORCID: 0000-0002-4201-5281. SPIN-код: 9494-6751.

Information about the author:
Yulia A. Yevgrafova, Dr.Sc. (Philology), Associate Professor, Professor at the Department of Advertising, Public Relations and Linguistics, National Research University «Moscow Power Engineering Institute» (14, build. 1 Krasnokazarmennaya str., Moscow, Russian Federation, 111250); *Research interests:* semiotics, semantics, film-semiotics, polycode texts, hermeneutics; *e-mail:* yu.evgrafova@mail.ru
ORCID: 0000-0002-4201-5281. SPIN-code: 9494-6751.

RUDN Journal of Language Studies, Semiotics and Semantics
2313-2299 (Print), ISSN 2411-1236 (Online)
Вестник Российского университета дружбы народов.
Серия: ТЕОРИЯ ЯЗЫКА. СЕМИОТИКА. СЕМАНТИКА

2026 Vol. 17 No. 2 496–516
http://journals.rudn.ru/semiotics-semantics

DOI: 10.22363/2313-2299-2026-17-2-496-516
EDN: LKJCAT

Научная статья / Research Article

Произведения больших языковых моделей как симулякр смысла. Семантическая энтропия и особенности новейших алгоритмов обнаружения использования искусственного интеллекта

М.Л. Новикова ✉, **Ф.Н. Новиков**

Российский университет дружбы народов, *Москва, Российская Федерация*
✉novikova-ml@rudn.ru

Аннотация. Ситуация, сложившаяся в связи со стремительным развитием и повсеместной доступностью искусственного интеллекта в целом и больших языковых моделей в частности, способных генерировать качественные, связные, правдивые или правдоподобные тексты на различных языках, является актуальной проблемой. Путем исследования таких междисциплинарных феноменов на стыке, как «семантическая энтропия», «симулякр смысла» и «перплексия», опираясь на теоретические базы квантитативной лингвистики, философии и информационных технологий, авторы статьи выдвигают гипотезу о том, что следствием таких изменений является радикальная трансформация информационной картины мира. Данная ситуация, с одной стороны, ведет к глобальной семантической девальвации и скептическому отношению к вновь создаваемым произведениям письменной речи, а с другой стороны повышает ценность аутентичных текстов, созданных человеком (подкрепленных его авторитетом и фактической базой). В статье также рассматриваются особенности новейших технических и семантических механизмов обнаружения фактов использования нейросетей при создании текстов и делается попытка описать дальнейшие тенденции в этой области. Авторами рассматривается феномен генеративного искусственного интеллекта (LLM) через призму концепции симулякра Жана Бодрийяра. Анализируется техническая природа нейросетей как вероятностного распределения токенов, создающего иллюзию семантической связности. На примерах демонстрируется разрыв между синтаксической корректностью и отсутствием интенциональности (смыслополагания), что позволяет квалифицировать продукты ИИ как «чистую форму», лишенную референта в реальности.

Ключевые слова: искусственный интеллект, семиотика, семантика, симулякр смысла, перплексия

Вклад авторов: вклад авторов равнозначен на всех этапах исследования и подготовки текста статьи.

Конфликт интересов: авторы заявили об отсутствии конфликтов интересов.

История статьи: дата поступления: 21.11.2025; дата приема в печать: 5.12.2025.

© Новикова М.Л., Новиков Ф.Н., 2026
This work is licensed under a Creative Commons Attribution-NonCommercial 4.0 International License
https://creativecommons.org/licenses/by-nc/4.0/legalcode

Для цитирования: *Новикова М.Л., Новиков Ф.Н.* Произведения больших языковых моделей как симулякр смысла. Семантическая энтропия и особенности новейших алгоритмов обнаружения использования искусственного интеллекта // Вестник Российского университета дружбы народов. Серия: Теория языка. Семиотика. Семантика. 2026. Т. 17. № 2. С. 496–516. https://doi.org/10.22363/2313-2299-2026-17-3-496-516 EDN LKJCAT

Works of Large Language Models as Simulations of Meaning. Semantic Entropy and Features of the Latest AI Detection Algorithms

Marina L. Novikova ✉, Philipp N. Novikov

RUDN University, *Moscow, Russian Federation*
✉novikova-ml@rudn.ru

Abstract. The situation created by the rapid development and widespread availability of artificial intelligence in general, and large language models in particular, which can generate high-quality and coherent true or truthful texts in different languages, is a pressing problem. By studying interdisciplinary phenomena at the junction such as “semantic entropy”, “simulation of meaning” and “perplexity”, based on the theoretical bases of quantitative linguistics, philosophy and information technology, the authors hypothesize that the consequence of such changes is a radical transformation of the informational picture of the world. This situation, on the one hand, leads to a global semantic devaluation and skepticism towards newly created written speech works, and, on the other hand, increases the value of authentic human-made texts (supported by their authority and factual basis). The article also discusses the features of the latest technical and semantic mechanisms for detecting the use of neural networks in text creation, and attempts to describe further trends in this area. The authors consider the phenomenon of generative artificial intelligence (LLM) through the prism of the concept of Jean Baudrillard’s simulation. The technical nature of the neural networks is analyzed as a probabilistic distribution of tokens, creating an illusion of semantic coherence. The examples show a gap between syntactic correctness and the lack of intentionality (meaning-making), which allows AI products to be classified as “pure form” that lacks a referent in reality.

Keywords: artificial intelligence, semiotics, semantics, simulation of meaning, perplexity

Authors’ contribution: The authors contributed equally to this article and preparation of the article text.

Conflict of interests: The authors stated that there were no conflicts of interest.

Article history: Received: 22.11.2025; accepted: 5.12.2025.

For citation: Novikova, M.L., & Novikov, Ph.N. (2026). Work of Large Language Models as Simulations of Meaning. Semantic Entropy and Features of the Latest AI Detection Algorithms. *RUDN Journal of Language Studies, Semiotics and Semantics, 17*(2), 496–516. (In Russ.). https://doi.org/10.22363/2313-2299-2026-17-3-496-516 EDN LKJCAT

Введение. Искусственный интеллект и лингвистика: точки сопряжения

Лингвистическая мысль в различные периоды развития языкознания отвечает на актуальные вызовы времени. Использование искусственного

интеллекта в лингвистике по праву считается и действительно является одной из самых активно развивающихся областей современной науки, что связано с доминированием статистических методов, нейросетевого моделирования и дистрибутивной семантики. Искусственный интеллект «существует как *динамическое поле возможностей* (*курсив наш*), которое актуализируется по-разному в зависимости от контекста взаимодействия, от языка описания, от самой постановки вопроса» [1. С. 326].

На протяжении десятилетий лингвистика и информатика двигались навстречу друг другу — от формальных грамматик Н. Хомского, которые легли в основу языков программирования, до современных нейронных сетей, способных генерировать информацию в любой цифровой форме.

В середине XX в. два направления исследования переживали свое рождение практически одновременно: искусственный интеллект (в первоначальном понимании слова) и генеративная лингвистика. В центре этого интеллектуального взрыва, качественного изменения в исследованиях находился Ноам Хомский, а его книга «Синтаксические структуры» [2] — формализованная общая теория лингвистической структуры, база современных теорий, необходимых для решения задач, возникающих при автоматической переработке естественно-языковой информации, не просто обогатила лингвистику, но и предоставила математический аппарат, необходимый компьютерным наукам для работы с языком. Во второй половине XX в. доминирующим был символьный искусственный интеллект. Пытаясь реализовать эти идеи, ученые выдвинули гипотезу о том, что описание всех правил грамматики (следуя Н. Хомскому), создание огромных «деревьев» правил, загруженных в компьютер, даст возможность машине овладеть языком. Но язык — подвижная динамичная структура с нарушениями правил; жесткие грамматики Н. Хомского «ломались» на реальных текстах. Грамматика Н. Хомского не только описывает то, что *можно* сказать, но и то, что сказать *нельзя*; искусственный интеллект может генерировать все, если это *статистически вероятно*, не обращаясь к ограничениям формальной грамматики.

Примечательно, что в своих недавних исследованиях ученый, обращаясь к базовому осмыслению структур языка, на которых основаны большие языковые модели, отмечает, что сложность системы правил и структуры языка предполагает сложность задачи их воспроизведения искусственным интеллектом [3]. Научное знание переходит к синтезу, где статистическая мощь искусственного интеллекта направляется в русло структурных ограничений, описанных великим лингвистом несколько десятилетий назад.

Исследователи пытаются объединить нейросети со строгостью формальных грамматик в нейросимвольном искусственном интеллекте. Таким образом, идеи Н. Хомского определили развитие языков программирования и первых языковых моделей, теоретически они являются стандартом для измерения сложности задач (от регулярных выражений до контекстной зависимости), на практике же они реализуются созданием гибридных нейросимвольных систем.

Нейросети генерируют вероятностные распределения, предсказывая следующий пиксель (в картинках) или следующий токен (в тексте). Они создают информацию на основе усвоенных закономерностей обучающих данных. В философском смысле, по Бодрийяру [4], они генерируют гиперреальность — контент, который выглядит более упорядоченным, красивым и «правильным», чем реальный мир, но не имеет под собой физического оригинала. Большие языковые модели — и инструмент анализа, и объект изучения — обучены на больших объемах текстовых данных, они способны генерировать речь и текст, что позволяет им выполнять различные задачи, связанные с обработкой естественного языка. «В основе языковых моделей лежат нейронные сети. Они обучаются путем анализа огромных массивов текстов, выявляя закономерности, грамматические правила и смысловые связи» [5. С. 109].

Современный искусственный интеллект как «совокупность технологий компьютерных систем, имитирующих когнитивные способности человека (обучение, анализ данных, адаптация к изменяющимся условиям, принятие решений в режиме реального времени и обработка естественного языка)» [6. С. 791] не обращается к правилам грамматики: он учится на терабайтах текстов, предсказывая, вычисляя возможность следующего слова, **он** не *выводит* правила, а выражает их статистически.

В эпоху небывалого ранее прогресса нейронных сетей применение больших языковых моделей «инициирует радикальный сдвиг парадигмы в лингвистике — от эпистемологии антропоцентризма к онтологии новых „постантропных“ коммуникаций» [7. С. 145].

Классическая лингвистика ставила целью описать структуру языка (Ф. де Соссюр) или порождающие правила (Н. Хомский), современная компьютерная лингвистика в союзе с искусственным интеллектом переживает фундаментальные перемены. Если раньше (в эпоху классических алгоритмов) главной задачей было научить машину построить дерево зависимостей, то сейчас, в эпоху больших языковых моделей, задача изменилась: надо научить машину понимать смысл. Необходимо обучить языковую модель рассуждению (chain-of-thought). Это важно, чтобы перед тем, как ответить на вопрос, модель «проговаривала» про себя логические шаги, проверяла их и только после этого генерировала текст.

Лингвистические знания и методы исследования искусственного интеллекта, анализа баз данных применяются в лингвистике широко, открывая новые грани полипарадигмальных исследований: «оценка успешности моделирования отдельных элементов языковой системы и механизмов мышления, выражаемых с помощью языка, исключительно важна для систем искусственного интеллекта. Это, в свою очередь, предполагает и новые эвристические возможности для междисциплинарных исследований языка на пересечении философии, социологии, когнитивных и информационных дисциплин» [8. С. 119]. Интерес исследователей направлен на изучение лингвистического потенциала нейросетей, на анализ с позиций фреймовой семантики нарратива

[9], на возможности нейронных моделей в моделировании грамматических структур [10. С. 179]; изучаются употребление слов различных тематических групп в сгенерированной и естественной речи [11. С. 48], композиционно-структурные, семантические и пресуппозиционально-прагматические параметры сгенерированных текстов, генерация текстов в аспекте авангардной поэтики [12], исследуются способности больших языковых моделей интерпретировать референцию на материале фразеологических выражений [13], анализируется интеграция практик машинного обучения и науки о языке, оценка систем искусственного интеллекта и его эффективного обучения, где в качестве основного инструмента оценки уровня интеллектуальности системы используются языковые тесты [8. С. 119].

Междисциплинарные исследования от вопроса «Как научить машину „говорить“» перешли к изучению проблемы; «Что машина поняла о структуре языка?», «Как знание законов языка может исправить ошибки нейросетей (галлюцинации, смещения и др.)?». Лингвистика как наука, изучающая естественный человеческий язык, его структуру, функционирование, влияние на мышление, использование в обществе, нацелена на объяснение, почему язык устроен именно так, в то время как искусственный интеллект — «черный ящик» с миллиардами параметров — стремится к *предсказанию* и *результату*. Он выдает правильный результат, но внутренняя логика его работы абсолютно чужда человеческой логике и лингвистической теории.

«Как моделирующая наука лингвистика обладает широким спектром методов, которые наследуются другими дисциплинами, связанными с обработкой языковых данных» [7. С. 119]. Лингвисты должны стать «биологами», изучающими новый вид «организмов» — нейросети [14]. Взаимодействие лингвистики и искусственного интеллекта выражается не только в использовании текстов для обучения нейросетей, оно включает тесные неразрывные отношения связи и взаимозависимости, при которых языковые данные являются материалом, а лингвистические теории — базой для структуры алгоритмов, модифицированных языковыми практиками моделей, способных аккумулировать то, что «ускользает как от чисто технического описания (векторы, матрицы, алгоритмы), так и от традиционного философского анализа (сознание, субъект, интенциональность). В этом гибридном пространстве технические термины обретают экзистенциальную глубину, а философские концепты получают операциональную конкретность» [1. С. 319].

Развитие нейросетей как катализатор трансформации глобальной текстовой картины мира

Нейросети и алгоритмы, генерирующие текст, существовали и ранее. Тем не менее, они были очень сильно ограничены по своим функциям и могли работать только в закрытой семиотической системе с четким разделением на небольшое количество семантических категорий — например, они могли описывать данные о фондовом рынке или прогнозы погоды. Основными

чертами и даже условиями существования подобных систем были их закрытость, однотипность генерируемых текстов и узконаправленность. Искусственный интеллект способен создавать новые, оригинальные тексты [15]. Именно модели, работающие в режиме диалога, стали тем форматом, который является наиболее удобным и привычным человеку, особенно тогда, когда он не имеет опыта программирования и написания инструкций для машин. Выход ChatGPT в конце 2022 г. стал важнейшей точкой отсчета, популяризовавшей технологии машинного обучения среди пользователей, не являющихся профессионалами в области информационных технологий, и сделавшей другие модели, в том числе русскоязычные Алиса AI и GigaChat, доступными для широкого круга пользователей.

Стремительное развитие нейросетей в целом и больших языковых моделей в частности, наблюдаемое в течение последних трех лет [6] с выходом ChatGPT, а также массовое использование данных технологий полностью изменили не только возможности современных пользователей Интернета, но и послужили катализатором полной трансформации глобальной текстовой картины мира. Это связано с тем, что мы переходим от эры, когда текст был исключительно продуктом человеческого разума, к эпохе *гибридного* и *синтетического* контента. Глобальная текстовая картина мира — это все записанные знания, мнения, факты и др., доступные человечеству. Раньше этот массив данных рос органично и поэтапно, ограниченный когнитивными способностями и возможностями людей. Сейчас Интернет наполняется «шумом» — сгенерированными статьями, комментариями и отзывами, голоса людей теряются в океане синтетического текста. Нейросети обучаются на данных из интернета, поэтому, если он наводняется текстами, которые создали другие нейросети, то модели начинают обучаться на синтетических данных, что ведет к ухудшению качества, это явление носит название «самозаражения» моделей. Языковая модель усредняет реальность: исчезают диалекты, стилевое разнообразие, изобразительно-выразительные средства, культурные особенности, которые не вписываются в статистическую норму. Трансформация глобальной текстовой картины мира неизбежна. Главная проблема в том, чтобы сохранить *субъектность*. Если мы перестанем писать и думать самостоятельно, полностью делегировав это нейросетям, наша картина мира станет плоской, усредненной галлюцинацией машины.

Достижения языковых моделей в генерации текстов приводят к ложным выводам, что проблема понимания естественного языка решена. Несомненно, с позиций теоретической лингвистики современные системы демонстрируют компетентность в грамматической системе языка (*синтаксис, морфология)*, но не могут преодолеть серьезные проблемы в *содержании* (семантика) и *намерении* (прагматика). Языковые модели владеют «формальной лингвистической компетентностью», поскольку знают правила языка, но не обладают «функциональной компетентностью», обусловленной мышлением и логикой [16].

Семиотические системы в искусственном интеллекте. Семантическая интерпретация, прагматическая контекстуальность. От обработки сигналов к интерпретации смыслов

Проблемы семиотики, семантики, прагматики и верификации истины требуют применения фундаментальной лингвистической теории в структуре машинного обучения. Интеграция семиотики в разработки искусственного интеллекта продиктована и теоретической традицией, и технологической необходимостью решения многих задач, в том числе фундаментального вопроса о том, может ли машина на самом деле понимать, или она лишь имитирует понимание через сложную статистику. Это переход от уровня «как передать данные» (сигнал) к уровню «зачем это сказано и что это значит в текущей ситуации» (смысл и цель). Семиотика, язык, коммуникация и прагматика исследуются как единое концептуальное пространство. «Семиотика по отношению к сознанию выступает как его ознаковление, означивание. Язык, в союзе с семиотикой, по отношению к сознанию выступает в качестве его репрезентанта, как средство материализации мыслительных структур [17. С. 268]. Искусственный интеллект оперирует формой, но не смыслом; статистикой, но не логикой реального мира; основные проблемы — в сфере семантики и прагматики. Кроме того, очень важны и значимы проблемы интерпретируемости («черный ящик») и этические вопросы (смещения, галлюцинации), требующие исследования. Современные нейросети отлично пользуются синтактикой, статистическим комбинированием знаков, но демонстрируют трудности с семантикой, связью знака с реальным миром и прагматикой, пониманием целей коммуникации.

Для перехода от генерации правдоподобного текста к реальному решению задач и пониманию контекста языковые модели должны научиться работать со структурой знака, различая его форму (сигнал) и содержание (смысл). Для этого нужно исследовать *рождение смысла на пересечении человеческого и искусственного интеллекта.* Следует особо подчеркнуть, что для искусственного интеллекта смысл слова — это то, как оно используется в тысячах различных контекстов, в тысячах практик [1. С. 323]. В связи с этим необходимо обратиться к вопросу о *семиотическом пределе* искусственного интеллекта, переходе от синтаксической имитации к генеративному семиозису. Активный исследовательский интерес к этим проблемам — свидетельство их возрастающей значимости. Стремительное развитие генеративного искусственного интеллекта актуализировало фундаментальные вопросы философии языка и семиотики. Если ранее машинное обучение рассматривалось преимущественно в категориях кибернетики и информатики, то появление таких мощных моделей, как GPT-4, диктует необходимость решения важного вопроса: является ли деятельность нейросети семиозисом (процессом порождения и интерпретации знаков) или же мы наблюдаем сложную форму стохастической манипуляции означающими без доступа к означаемому?

Семиотический подход к искусственному интеллекту («когнитивная семиотика» или «семиотическое моделирование») дает основание рассматривать искусственный интеллект как интерпретирующую систему: знак — объект — интерпретанта. В семиотике знак рассматривается *в единстве означающего* — форма (звук, написанное) и *означаемого* (смысл, образ предмета в реальности). Для человека «Апельсин» (означаемое) тесно связан с фруктом оранжевого цвета, сочным, цитрусовым, который можно съесть (означаемое). Для языковой модели токен [Апельсин] — *это вектор, набор цифр* в многомерном пространстве. имеющий математическую близость к векторам [Фрукт], [Оранжевый], [Вкусный], но он не связан с реальным апельсином, что свидетельствует о разрыве означающего и означаемого. Языковая модель руководствуется означающими, она оперирует «оболочками» слов, создавая своего рода орнаменты и узоры, *предсказывая возможность и вероятность следующего токена на основе предыдущих контекстов*.

Закономерно возникает вопрос: есть ли в сгенерированном тексте отношение к реальности или мы наблюдаем только взаимодействие знаков? Жан Бодрийяр называл это «симулякром третьего порядка» — знаком, который маскирует отсутствие глубинной реальности, перенеся исследования симулякра из сферы онтологии в описание современной социальной реальности, описывая процесс постепенного исчезновения реальности и замены ее знаками, утверждая, что «симулякр — это вовсе не то, что скрывает собой истину, — это истина, скрывающая, что ее нет» [4. С. 170]. Когда нейросеть пишет эссе о чувствах, которых у нее нет, или описывает историческое событие, которого не было, она делает вид, что за этими словами стоит опыт или факт. Но за ними — только математическая вероятность. «С точки зрения семиотического направления, «симулякр» — самореференциальный знак (образ несуществующей действительности), подменяющий символ, который олицетворяет слово, и его конвенциональные значения» [18. С. 146].

У модели нет доступа к физическому миру, но она *создает такую детальную систему языковых вероятностей,* что демонстрирует свойства, похожие на логическое рассуждение. Современные модели, которые обучаются на триллионах параметров, можно рассматривать как реализацию такого процесса, ведь они не познают мир (физическую реальность), они познают *тексты о мире* (знаковую реальность). «ИИ-симулякры — это не просто обманчивые факсимиле; они становятся действующими реальностями внутри цифровой культуры, все больше формируя дискурс, эстетику и даже эмоциональные реакции в отсутствие человеческой субъектности» [19]. Текст искусственного интеллекта — это копия человеческого мышления, но за ней не стоит думающий субъект, вероятность представлена вместо истины.

Проблема генеративного искусственного интеллекта — отсутствие референта, объекта в реальном мире, на который указывает слово. Для человека язык *интенционален* (нацелен на предмет, мы говорим о чем-то), а для модели — нет, она предсказывает следующий токен на основе предыдущего.

В языковых моделях это проявляется следующим образом: обучающая выборка (Dataset) — это отражение цифрового пространства, которое само по себе является набором репрезентаций, мнений и копий, нейросеть учится не на реальности, а на ее отражении, как результат искусственный интеллект генерирует *копию копии без оригинала*. Это объясняет феномен «галлюцинаций» с технической точки зрения: галлюцинация — это ошибка, но по Ж. Бодрийяру, — это торжество симулякра. Галлюцинация не является ложью, потому что у искусственного интеллекта нет концепции «истины», это чистый знак, освобожденный от обязанности что-либо означать. «Инструменты генеративного ИИ создают тексты, изображения и аудиовизуальный контент, которые не отсылают к какой-либо единственной реальности или авторскому замыслу, а скорее возникают из вероятностных моделей, обученных на огромных корпусах ранее существовавших данных. Этот процесс воплощает чистый симулякр Бодрийяра, в котором результат не является ни копией, ни подделкой, а совершенно новой конструкцией, оторванной от первоисточника» [20].

Исследования показывают: если обучать новые модели на данных, сгенерированных предыдущими моделями, качество деградирует. Симулякры начинают копировать симулякры. Реальность (человеческий опыт) вымывается из обучающей выборки. Коммуникация, по Бодрийяру, требует обмена, но ведь диалог с искусственным интеллектом — не обмен, а *бесконечное отражение запросов* пользователей. Эффективное статистическое моделирование языка выявляет и идентифицирует модели и образцы в огромных массивах данных, но языковая модель не понимает физической реальности и причинно-следственных связей, у нее нет тела и опыта взаимодействия с окружающей действительностью [21].

Одна из фундаментальных проблем в области искусственного интеллекта — проблема *семантики*, или способности машин извлекать, интерпретировать и генерировать смысл: можно идеально манипулировать символами по правилам, не понимая, что они значат. Несмотря на достижения в создании больших языковых моделей, увеличение их количества, объема и мощности, разработку более совершенных методов глубокого восприятия и интерпретации контекста [5. С. 110], выделяются существенные проблемы, связанные с их пониманием и обработкой языка. Искусственный интеллект сейчас находится на этапе *«семантической имитации»*. Он уверенно предсказывает, какое слово *должно* идти следующим, но не *понимает*, почему оно там стоит.

Вопрос о том, может ли машина «понимать» *смысл* обрабатываемой информации, является одним из самых важных. В эпоху доминирования больших языковых моделей граница между *синтаксической обработкой данных* и *семантическим осмыслением* стала еще более размытой. Переход от статистической обработки текста к реальному смысловому пониманию — главная задача на пути к совершенствованию искусственного интеллекта. Главной

проблемой компьютерной лингвистики остается разрыв между статистической вероятностью появления слова и его смысловым наполнением. Языковые модели обучаются предсказывать следующий токен на основе массивов данных. Они «знают», какие слова часто стоят рядом, но не имеют когнитивной модели мира, стоящей за этими словами, не понимают, «как функционирует реальный мир. Они могут распознавать закономерности, но не понимают причинно-следственных связей, <…> ориентированы на выявление статистических закономерностей в данных, чем на понимание их смысла» [5. С. 110]. В области *семантики, проблем смысла и значения* единиц языка и знаковых систем, искусственный интеллект использует векторные представления, слова и фразы преобразуются в математические векторы. *Близость векторов* означает *близость значений*. В семантике значение описывается через *семы (компоненты смысла*), в языковых моделях оно описывается через *координаты вектора*. Несмотря на то, что искусственный интеллект предсказывает токен, формируя цепочки размышлений, распределяет вычисление в зависимости от сложности задачи [6. С. 792], трудные случаи полисемии ему недоступны. Например, фраза *«Предприятие стоит»* без расширенного контекста недоступна в полной мере искусственному интеллекту: означает ли это, что предприятие не работает (бездействие), или что оно находится в определенном месте (локация), или речь идет о стоимости. Слова имеют разные значения в зависимости от контекста. Для человека выбор значения происходит автоматически благодаря жизненному опыту, а для искусственного интеллекта — это математическая вероятность.

Семантическая энтропия как фундаментальное свойство языка. Энтропия как мера непредсказуемости символа

Энтропия (*греч.* Εντροπία — *поворот, превращение,* мера беспорядочности или разупорядоченности системы) как универсальное естественнонаучное понятие прочно вошло в гуманитарные и компьютерные науки. Специалист по языковым моделям видит в ней меру неопределенности некоторой ситуации [22], «связанной с данными: чем более неопределенны данные, тем выше энтропия. Точнее, энтропия связана с количеством информации, содержащейся в данных, то есть чем более неопределенные данные, тем больше информации требуется, чтобы описать их» [23. С. 153]. Лингвист видит в энтропии возможности для творчества, определяющие потенциал языка [24]. Различные подходы к трактовке энтропии открывают разнообразные возможности исследования этого феномена. Уточним, что в русской лингвистической традиции энтропия — не «шум», а «глубина», не мера непредсказуемости символа, а явление, имеющее глубокие корни, уходящие в структурную лингвистику, семиотику и теорию информации.

Семантическая энтропия рассматривается как фундаментальное свойство языка, обеспечивающее его гибкость и способность передавать новые

смыслы; это мера неопределенности значения языкового знака (слова, фразы) вне контекста, важная характеристика словесного творчества.

Интересно в этой связи обратиться к *избыточности, предсказуемости, информативности* и *новизне.* Низкая энтропия (высокая упорядоченность/предсказуемость) и высокая избыточность (наличие элементов, дублирующих информацию) тесно связаны в языке, что вполне объяснимо, поскольку чем выше избыточность, тем ниже его энтропия, так как следующее слово или букву легче предсказать. Эти особенности важны в таких речевых ситуациях, где достоверность информации более приоритетна, чем ее оригинальность и новизна. К ним относятся этикетные ритуальные фразы, документы, протоколы, инструкции, в которых используются клише, сложные синтаксические конструкции и повторы, чтобы избежать разночтений, обеспечивающих однозначное понимание. Рифма, ритм и устойчивые эпитеты понижают энтропию, ограничивают выбор слова, а во фразеологизмах, воспринимающихся как единое целое, энтропия стремится к нулю. *Низкая энтропия* и *высокая избыточность* — фундамент и основа языка, обеспечивающие быстрое понимание, в том числе и при беглом чтении (мы не читаем все буквы, мы угадываем слова). Если текст с низкой энтропией можно читать поверхностно, то текст с высокой энтропией требует внимания и сосредоточенности, потому что, пропустив одно слово, мы можем потерять смысл всего сообщения.

Высокая энтропия в языке — непредсказуемость следующего слова на основе предыдущих, информационная плотность и *новизна* сообщений, в которых очень мало избыточности, где каждое слово несет информационную нагрузку. Это язык творчества, нарушение стандартных связей слов с невозможностью достоверно предсказать следующее слово, что требует усилий для расшифровки, но рождает новые смыслы. Чем больше у слова значений (полисемия), тем выше его семантическая энтропия; таким образом, процесс понимания сообщения — процесс *подавления энтропии* с помощью контекста, где из веера возможных значений выбирается правильное. «Чем выше энтропия, тем больше неопределенности или многозначности содержится в сообщении, что делает его менее предсказуемым или однозначным» [22. С. 80]. Например, низкая энтропия характеризует канцелярский язык или разнообразные команды и указания.

Если бы слова имели четко фиксированные значения, как в языке программирования, эволюция мысли остановилась бы, а человек не смог бы выразить ничего нового, только комбинировать старое. Смыслы не жестко привязаны к знакам, они размыты, что дает возможность людям понимать друг друга не как машины, обменивающиеся командами, а как личности, «соприкасающиеся мирами». Значение слова — не точка, а функция распределения вероятностей [25]. Высокая энтропия рассматривается как *ценность*. Ю.М. Лотман отмечал, что культура стремится повышать свою энтропию, чтобы не застыть. В семиотике «ошибка» (нарушение нормы) — это часто рождение нового значения.

В классической теории информации (К. Шеннон) энтропия — *мера неопределенности или «свободы выбора»*. Для понимания энтропии как возможности нужно опираться не на смысл сообщения в его лингвистическом осмыслении, а на *вероятность* его появления. «Энтропия — это мера неопределенности содержания текста, она связана с объемом информации, заключенной в данных: чем более они неопределенны, тем больше информации требуется для их описания» [22. С. 78]. Необходимо подчеркнуть, что понятие энтропии трансформируется при переходе от теории информации к системам искусственного интеллекта, и поэтому для нейросети и поэта «хаос» имеет разное значение; энтропия = неопределенность = объем информации является не метафорой, а строгим математическим законом.

Неопределенность — это количество вариантов, из которых можно выбрать, информация — результат выбора. Чем он сложнее (чем выше была неопределенность), тем важнее результат. Объем описания — цена, которую мы платим за сохранение уникальности этого выбора. Хаос требует детального описания каждого элемента, в то время как порядок позволяет использовать краткие формулы. Текст с высокой энтропией информационно насыщен и требует большего объема памяти для хранения, так как он непредсказуем. Для современных больших языковых моделей энтропия — это рабочий инструмент: искусственный интеллект стремится снизить неожиданность следующего слова. Идеальная модель — та, для которой текст из обучающей выборки имеет минимальную энтропию (она точно знает, что будет дальше). Для искусственного интеллекта энтропия — это ландшафт, структурирование и символизирование окружающей среды, это математическая задача. Для семиотики энтропия — это пространство, где рождаются искусство, юмор и новые культурные коды. Понимание этих двух граней энтропии необходимо для создания искусственного интеллекта нового поколения, который сможет не просто генерировать статистически верный текст, но и понимать природу творчества. Переход от энтропии токенов к семантической энтропии знаменует собой сдвиг от анализа *синтаксиса* к анализу *семантики* в искусственном интеллекте. Семантическая энтропия переносит измерение неопределенности из пространства лексики (слова/символы) в пространство смыслов (семантика). Переход от лексического пространства к семантическому — фундаментальный сдвиг в том, как мы оцениваем «уверенность» искусственного интеллекта. Семантическая энтропия — мост между математической теорией информации и гуманитарной наукой о смыслах. Этот инструмент критически важен не только для теоретической лингвистики, но и для создания надежного искусственного интеллекта, способного понимать, когда он говорит правду, а когда — генерирует шум.

Прагматика и контекстуальная «слепота»

Серьезная проблема искусственного интеллекта — неспособность понимать контекст, намерения и скрытые смыслы, очевидно, что между

«правильным» текстом модели и «осмысленным» общением лежит огромная пропасть. Отсутствие прагматической составляющей — главная причина, по которой языковые модели совершают ошибки в простых ситуациях. Прагматика изучает отношения субъекта коммуникации к языку и экстралингвистический контекст, *значение предложения* (буквальный смысл) и *значение говорящего* (что человек хотел сказать). Искусственный интеллект не всегда хорошо справляется с распознаванием иронии, сарказма и культурных аллюзий, так как они не содержатся в самом сигнале (тексте), а находятся в надсистеме (в культуре, ситуации и др.).

Сарказм часто строится на противоречии между позитивной лексикой и негативным контекстом, или тональностью. В моделях нет просодической информации (интонация, ударение), визуальных указаний и помощи (мимика, жесты), поэтому они воспринимают текст буквально. Научить машину грамматике сложно, но возможно, так как грамматика логична. Научить машину прагматике — значит научить ее понимать саму суть человеческого существования: нашу нелогичность, наши недомолвки, нашу культуру и наш юмор.

Язык пронизан аллюзиями и культурными отсылками. Утверждение *«Он настоящий Плюшкин»* для носителя русского языка означает патологическую скупость человека, страсть к накопительству и деградацию личности. Для искусственного интеллекта, обученного преимущественно на англоязычных текстах или технической документации, эта метафора может быть потеряна или переведена буквально, если в обучающей выборке не было книг Н.В. Гоголя с комментариями.

В силу того, каким образом обучаются нейросети, следует отметить англоцентризм многих больших языковых моделей, которые изначально обучались на английском языке, на котором представлено наибольшее количество индексируемой информации в Интернете. Это приводит к наличию большого количества англицизмов на всех уровнях языка, вплоть до таких очевидных аспектов, как отсутствие согласования между членами предложения. Эта проблема постепенно решается как создателями глобальных моделей искусственного интеллекта, совершенствующих разметку лексических массивов данных, так и русскоязычными компаниями. Помимо этого повсеместное использование больших языковых моделей ведет к унификации и стандартизации речи, а также сопровождается искоренением региональной вариативности и лексико-грамматических особенностей.

Ситуации с растущим недоверием к текстам и информации, сгенерированной нейросетями, может быть противопоставлено чрезмерное доверие к подобным технологиям без дальнейшей верификации, что может привести к серьезным проблемам в различных профессиональных сферах. Одним из первых известных примеров подобного использования может служить ситуация, возникшая в июне 2023 г. — спустя полгода после выхода ChatGPT [26]. Адвокат, использовавший раннюю модель ChatGPT, не имеющую выхода в

Интернет, сгенерировал юридическое заключение, которое затем было предоставлено в суд Нью-Йорка. Документ содержал шесть ссылок на судебные прецеденты, которые выглядели правдоподобно, имели непосредственное отношение к рассматриваемому делу и были грамотно составлены, но при этом не существовали вообще. Когда непреднамеренная фальсификация была раскрыта, автор текста заявил, что не осознавал, что искусственный интеллект не ищет прецеденты в базе данных, а лишь создает текстовые объекты, похожие по структуре на то, что заложено в его базу знаний.

Ситуация улучшается, так как подключаются модули поиска по Интернету, а в новых моделях появился режим размышления, который позволяет искусственному интеллекту многократно проанализировать запрос пользователя и написать подробные инструкции для себя, «переведя» его с естественного языка, производимого человеком, в упрощенную форму данного естественного языка, а затем — в машинный код. Этот алгоритм открывает уникальную с точки зрения коммуникации ситуацию саморефлексии машины.

Алгоритмы обнаружения текстов, написанных большими языковыми моделями. Перплексия как оценка качества языковых моделей

Несмотря на новейшие разработки и большой прогресс, проблема создания текстов, основанных на несуществующих фактах, ни в коем случае не является решенной. Такой феномен носит название «галлюцинации», то есть уверенной реакции системы искусственного интеллекта, не соответствующей исходным данным. Несмотря на высокий уровень развития систем и даже на новейшие способности «размышления», с эпистемологической точки зрения система не в состоянии провести подробный самоанализ и осознать отношение своих высказываний к действительности.

Упоминая эпистемологические особенности, следует также обсудить так называемый эффект когнитивной амнезии Гелл-Манна — феномен, описывающий парадоксальное доверие людей к средствам массовой информации. Читатели положительно или нейтрально воспринимают информацию, которая находится за пределами их области экспертизы, но после того, как сталкиваются с хорошо знакомой темой, видят поверхностность предоставляемой информации и могут начать скептически относиться не только к рассматриваемому материалу, но и его источнику в целом.

Распространение генеративного искусственного интеллекта повышает важность человеческого авторитета и доверия к текстам, написанным человеком при условии последующей верификации их аутентичности.

Чтобы понять, каким образом работают алгоритмы обнаружения текстов, написанных большими языковыми моделями, следует понять принцип работы GPT, то есть генеративных предобученных трансформеров. Принцип их работы схож с механизмами предиктивного ввода, который существовал еще

несколько десятилетий назад на различных устройствах. Он может быть основан на завершении отдельной лексемы или словосочетания. Большие языковые модели, в свою очередь, работают с более крупными единицами языка и имеют более объемный контекст — у современных моделей он может превышать миллионы знаков.

В таких случаях вывод слов, словосочетаний, предложений и целых текстов строится в первую очередь на статистической вероятности, которая построена на основе результата обучения языковой модели. Так как современные неспециализированные модели, используемые пользователями интернета, содержат крайне объемный массив данных, даже сами разработчики программного обеспечения, создающие их, часто не могут предсказать их возможности. Тем не менее следующая единица, выдаваемая большой языковой моделью, предсказывается статистически, как и в языковых корпусах, основываясь на частотности сочетаний.

Для демонстрации и введения понятия перплексии рассмотрим пример сочетаемости лексемы «добрый» взятый из Национального корпуса русского языка воспользовавшись инструментом n-gram, где n — число лексических единиц. На примере ниже можно отметить, что сочетание «доброе утро» является гораздо более частотным, чем «доброе отношение». Это значит, что вероятность получить данное слово при генерации текста при равнозначных остальных параметрах гораздо выше, и большая языковая модель имеет тенденцию придерживаться определенных слов и выражений (табл. 1).

Сочетаемость лексемы «добрый» / Collocation of the lexeme “good”

№	Контекстов / Contexts	IPM	Отрывок Examples
1	1,112	2.86	доброе дело
2	832	2.14	доброе утро
3	558	1.43	доброе сердце
4	415	1.07	доброе слово
5	409	1.05	доброе имя
6	231	0.59	доброе лицо
7	188	0.48	доброе старое
8	149	0.38	доброе отношение

Источник: исследование М.Л. Новиковой и Ф.Н. Новикова.
Source: compiled by Narina L. Novikova & Philipp N. Novikov.

Основной параметр, позволяющий определить предсказуемость текста с помощью машинных средств, носит название перплексии. Перплексия является статистическим критерием и рассчитывается по следующей формуле:

$$PP(h) := 2^{H(p)} = 2^{-\sum_x p(x)\log_2 p(x)},$$

где $PP(p)$ — перплексия, $H(p)$ — энтропия, которая в данном случае обозначает хаотичность, неопределенность. Высокая степень перплексии

соответствует неожиданности и отклонению от статистической нормы, а низкая показывает высокую предсказуемость. Таким образом, в тексте с высокой перплексией языковая модель может с более высокой вероятностью угадать следующий токен (языковой знак, длина которого в зависимости от языка составляет определенное количество символов или слов), основываясь на данных о частотности, подобных тем, которые были представлены выше в примере из корпуса русского языка.

Важно также отметить, что в больших языковых моделях случайность регулируется таким показателем, как «температура». Этот параметр обычно варьируется от 0 до 1, где 0 — максимально детерминистическая ситуация, а 1 — наиболее произвольная, но при этом возможно, более релевантная определенному запросу ситуация. При параметре «температура», равном нулю, моделью должен выдаваться один и тот же ответ при одинаковом вводе данных.

Механизмы распознавания использования генеративного искусственного интеллекта полагаются в первую очередь именно на параметр перплексии. Это означает, что алгоритму нужно сначала предсказать наиболее вероятные сценарии генерации текста, а затем сопоставить их с тем текстом, который представлен в анализируемом документе. Таким образом, механизм детекции использования большой языковой модели сам по себе является большой языковой моделью. С другой стороны, при высоком параметре температуры повышается вероятность получения галлюцинаций, что с точки зрения семантики облегчает задачу по обнаружению сгенерированных текстов людьми, но не алгоритмами, которым становится сложнее найти лексико-семантические закономерности. Исследование, проведенное на основе массива данных, в который были включены аннотации к научным статьям, показало, что перплексия написанных человеком аннотаций к научным статьям с медианой 35,9 значительно выше, чем у тех, которые были сгенерированы искусственным интеллектом [27].

Помимо понятия перплексии и общего уровня предсказуемости текстов одним из надежных показателей использования генеративного искусственного интеллекта является всплеск популярности следующих лексических единиц, многие из которых, как считается, появились под влиянием большого коллектива людей, владеющих определенным региональным вариантом английского языка и занимавшихся первоначальной разметкой данных для нейросети ChatGPT: «delve into», «in summary», «take a dive into», «navigating the landscape» и т.д. Таким образом, определенные клише и устойчивые выражения, которые хотя и являются естественными для человеческой речи, выдаются искусственным интеллектом с гораздо более высокой стабильностью.

Среди синтаксических особенностей, которые проявляются в текстах, сгенерированных нейросетями, можно отметить наличие «Em dash» (длинного тире) — символа, являющегося одним из самых известных примеров автоматически сгенерированных текстов. Этот символ является нехарактерным

для большого количества письменных стилей и расположен во многих раскладках на клавиатурах в неудобных местах и является одним из самых явных признаков использования генеративного искусственного интеллекта.

Использование различных эмодзи, то есть графических символов, заранее встроенных в современные операционные системы, а также списков с буллитами также можно отнести к явным признакам использования искусственного интеллекта. Если каждый из них по отдельности не всегда является явным признаком того, что автор прибегал к генерации текста, то совокупность данных признаков приводит к ощутимому кумулятивному эффекту со стороны читателей, позволяющего распознавать произведения больших языковых моделей.

Повышенное внимание к определенным структурным признакам текста со стороны читателей может привести к обратному эффекту, при котором они могут подозревать авторов в использовании нейросетей для генерации текста, если обнаруживают хотя бы один из вышеупомянутых признаков, например — длинное тире. Ложное срабатывание механизмов присуще также и автоматическим системам опознавания нейросетевых текстов. Подобная ситуация может сложиться тогда, когда в систему загружается текст, данные о структуре которого заранее заложены в нее, или тогда, когда текст, написанный человеком, имеет определенные структурные признаки машинного текста, но при этом написан самостоятельно.

Заключение

Научные изыскания находятся в точке фазового перехода искусственный интеллект перерос стадию «стохастических попугаев» (чистая обработка сигналов), но еще не достиг стадии «субъекта коммуникации» (полноценная прагматика). Проблема заключается в том, что мы пытаемся извлечь человеческий смысл из математической функции. И пока неясно, может ли количество (вычислительная мощность и данные) перейти в качество (понимание интенций), или для этого нужно принципиально иное, не вероятностное построение, а когнитивное.

Сопряжение лингвистики и искусственного интеллекта переходит на новый этап. Период «чистой инженерии», когда казалось, что достаточно лишь увеличить объем данных и мощность серверов, заканчивается. Будущее — за гибридными, полипарадигмальными системами, где статистическая мощь нейросетей направляется и контролируется структурированными знаниями о языке и мире. Это позволит создать искусственный интеллект, который не просто имитирует владение языком, но и понимает его коммуникативную природу. Идеального алгоритма обнаружения искусственного интеллекта на данный момент не существует. Статистические методы работают до первой серьезной редактуры текста человеком, поэтому человеческий фактор до сих пор является крайне важным.

Список литературы

1. *Аршинов В.И., Янукович М.Ф., Гарашко Ю.Н.* Онтология в диалоге: рождение языка и смысла на пересечении человеческого и искусственного интеллекта // Вестник Пермского университета. Философия. Психология. Социология. 2025. № 3. С. 317–328. https://doi.org/10.17072/2078-7898/2025-3-317-328 EDN: AYNKYE
2. *Бедненко Ю.И.* Искусственный интеллект и проблема сознания: философские и лингвистические аспекты // Гуманитарные исследования. История и филология. 2025. № 17. С. 103–113. https://doi.org/10.24412/2713-0231-2025-17-103-113 EDN: LANAJG
3. *Бодрийяр Ж.* Симулякры и симуляция. Тула, 2013.
4. *Гусаренко С.В., Гусаренко М.К.* Композиционно-структурные, семантические и пресуппозиционально-прагматические параметры и дефекты сгенерированных коротких текстов в языковой нейросети GigaChat // Гуманитарные и юридические исследования. 2024. Т. 11. № 2. С. 368–379. https://doi.org/10.37493/2409-1030.2024.2.21 EDN: JMESZZ
5. *Гусаренко С.В., Гусаренко М.К.* Предикатно-фреймовое описание линейной семантики художественного нарратива // Филологические науки. 2020. № 3. С. 18–30. https://doi.org/10.20339/PhS.3-20.018 EDN: XZUUHX
6. *Демченков С.А., Федяева Н.Д.* Вербальный хеппенинг: о некоторых моделях компьютерной генерации текстов в аспекте авангардной поэтики // Humanitarian Vector. 2020. Т. 15. № 1. С. 27–37. https://doi.org/10.21209/1996-7853-2020-15-1-27-37 EDN: ZJTBAV
7. *Зырянова И.Н.* Генеративные языковые модели и феномен антиантропоцентризма — новые перспективы лингвистической парадигмы «posthumano» и «общего/сильного» // Известия Байкальского государственного университета. 2024. Т. 34. № 1. С. 144–152. https://doi.org/10.17150/2500-2759.2024.34(1).144-152 EDN: NGYEEB
8. *Матурана У., Варела Ф.* Древо познания: биологические корни человеческого понимания. М. : Прогресс-Традиция, 2001.
9. *Назаров Н.В., Шульгинов В.А.* «Первый блин комом»: ограничения и потенциал LLM в распознавании референции устойчивых фраз // Коммуникативные исследования. 2024. Т. 12. № 4. С. 881–896. https://doi.org/10.24147/2413-6182.2025.12(4).881-896 EDN: ZBGBZR
10. *Налимов В.В.* Вероятностная модель языка. О соотношении естественных и искусственных языков. М. : Наука, 1974. EDN: QXGCTG
11. *Саакян Л.Н.* ИИ как гуманизированный субъект: митигативные стратегии искусственного интеллекта // Коммуникативные исследования. 2025. Т. 12. № 4. С. 789–804. https://doi.org/10.24147/2413-6182.2025.12(4).789-804 EDN: CJFFSS
12. *Студеникина К.А.* Оценка языковой способности нейронных моделей на материале предикативного согласования в русском языке // Труды Института системного программирования РАН. 2022. № 34 (6). С. 179–184. https://doi.org/10.15514/ispras-2022-34(6)-14 EDN: UCJBEV
13. *Тельпов Р.Е., Ларцина С.В.* Типовые различия естественных и сгенерированных нейронной сетью текстов в квантитативном аспекте // Научный диалог. 2023. Т. 12. № 7. С. 47–65. https://doi.org/10.24224/2227-1295-2023-12-7-47-65 EDN: ZZRRKN
14. *Хомский Н.* Синтаксические структуры // Новое в лингвистике. М., 1962. Вып. II. С. 412–527.
15. *Шаврина Т.О.* О методах компьютерной лингвистики в оценке систем искусственного интеллекта // Вопросы языкознания. 2021. № 6. С. 117–138. https://doi.org/10.31857/0373-658X.2021.6.117-138 EDN: PFGVQJ
16. *Шарандин А.Л.* Сознание, семиотика, язык, коммуникация и прагматика как единое концептуальное пространство // Вопросы психолингвистики. 2015. № 3 (25). С. 240–250. EDN: UDLHJT
17. *Шевченко Е.В.* Симулякр как тактический прием оказания ориентирующего воздействия // Вопросы когнитивной лингвистики. 2015. № 3 (44). С. 144–150. EDN: TTYPMF

18. *Павловская О.Е., Ильина Л.Е.* Энтропия текста как рассеивание смысла // Вестник Северного (Арктического) федерального университета. Серия: Гуманитарные и социальные науки. 2023. Т. 23. № 6. С. 77–86. https://doi.org/10.37482/2687-1505-V308 EDN: ZPZZPW
19. *Шелестюк Е.В., Щетинкина Е.А.* Стохастичность и энтропия в лингвистике // Вестник Челябинского государственного университета. 2023. № 2 (472). С. 150–165. EDN: WIVZGY
20. *Андриенко Т.П.* Энтропия текста как предпосылка реализации стратегии перевода // Известия Гомельского государственного университета им. Ф. Скорины. 2017. № 1 (100). С. 76–80. EDN: YFYPLF
21. *Baroni M.* On the proper role of linguistically-oriented deep net analysis in linguistic theorizing // Algebraic Approaches to Sequence Data. 2022. https://doi.org/10.1201/9781003205388-1
22. *Chomsky N., Roberts I., Watumull J.* Noam Chomsky: The False Promise of ChatGPT // The New York Times. 2023. 8 March.
23. *Elek A., Yildiz H.S., Akca B., Oren N.C., Gundogdu B*. Evaluating the efficacy of perplexity scores in distinguishing AI-generated and human-written abstracts // Academic Radiology. 2025. Vol. 32. Iss. 4. P. 1785–1790. https://doi.org/10.1016/j.acra.2025.01.017 EDN: HMLYED
24. *Jiang S., Li K.* Another kind of authenticity: the visual simulacra of artificial intelligence // Digital Creativity. 2025. Vol. 36. Iss. 1. P. 75–91. https://doi.org/10.1080/14626268.2025.2467087 EDN: ZRIUEB
25. *Konzack L.* Generative AI, Simulacra, and the Transformation of Media Production // Athens Journal of Mass Media & Communications. 2025. Vol. 11. Iss. 3. https://doi.org/10.30958/ajmmc.11-3-3 EDN: BYIUCE
26. *Lyon C.F.* Fake cases, real consequences: misuse of ChatGPT leads to sanctions // NY Litigator. 2023. Vol. 28. Iss. 2. P. 8–12.
27. *Mahowald K., Ivanova A.A.* et al. Dissociating language and thought in large language models: a cognitive perspective // Trends in Cognitive Sciences. 2024. https://doi.org/10.1016/j.tics.2024.01.011 EDN: SKTATZ
28. *Miralles-González P., Huertas-Tato J., Martín A., Camacho D.* Not all tokens are created equal: Perplexity Attention Weighted Networks for AI generated text detection // arXiv. 2025. https://doi.org/arXivpreprint arXiv:2501.03940
29. *Pudasaini S., Miralles-Pechuán L., Lillis D., Llorens Salvador M.* Survey on AI-generated plagiarism detection: The impact of large language models on academic integrity // Journal of Academic Ethics. 2024. https://doi.org/10.48550/arXiv.2407.13105
30. *Ram A.* Generative AI and Text Creation. Preprint. 2023. https://doi.org/10.48550/arXiv.2407.13105

References

1. Arshinov, V.I., Yanukovich, M.F., & Garashko, Yu.N. (2025). Ontology in Dialogue: Birth of Language and Meaning at the Intersection of Human and Artificial Intelligence. *Perm University Herald. Philosophy. Psychology. Sociology*, 3, 317–328. (In Russ.). https://doi.org/10.17072/2078-7898/2025-3-317-328 EDN: AYNKYE
2. Bednenko, Yu.I. (2025). Artificial Intelligence and the Problem of Consciousness: Philosophical and Linguistic Aspects. *Humanitarian Studies. History and Philology*, 17, 103–113. (In Russ.). https://doi.org/10.24412/2713-0231-2025-17-103-113 EDN: LANAJG
3. Baudrillard, J. (2013). *Simulacra and Simulation*. Tula. (In Russ.).
4. Gusarenko, S.V., & Gusarenko, M.K. (2024). Compositional-Structural, Semantic, and Presuppositional-Pragmatic Parameters and Defects of Generated Short Texts in the Gigachat Language Neural Network. *Humanities and Law Studies*, *11*(2), 368–379. (In Russ.). https://doi.org/10.37493/2409-1030.2024.2.21 EDN: JMESZZ

5. Gusarenko, S.V., & Gusarenko, M.K. (2020). Predicate-Frame Description of the Linear Semantics of a Literary Narrative. *Philological Sciences*, 3, 18–30. (In Russ.). https://doi.org/10.20339/PhS.3-20.018 EDN: XZUUHX
6. Demchenkov, S.A., & Fedyaeva, N.D. (2020). Verbal Happening: On Some Models of Computer Text Generation in the Aspect of Avant-Garde Poetics. *Humanitarian Vector*, *15*(1), 27–37. (In Russ.). https://doi.org/10.21209/1996-7853-2020-15-1-27-37 EDN: ZJTBAV
7. Zyryanova, I.N. (2024). Generative Language Models and the Phenomenon of Antianthropocentrism: New Perspectives of the Linguistic Paradigm "Posthumano" and "General/Strong" AI. *Bulletin of Baikal State University*, *34*(1), 144–152. (In Russ.). https://doi.org/10.17150/2500-2759.2024.34(1).144-152 EDN: NGYEEB
8. Maturana, H., & Varela, F. (2001). *The Tree of Knowledge: The Biological Roots of Human Understanding*. Moscow: Progress-Traditsiya. (In Russ.).
9. Nazarov, N.V., & Shulginov, V.A. (2024). "The First Pancake is Always Lumpy": Limitations and Potential of LLM in Recognizing the Reference of Set Phrases. *Communication Studies*, *12*(4), 881–896. (In Russ.). https://doi.org/10.24147/2413-6182.2025.12(4).881-896 EDN: ZBGBZR
10. Nalimov, V.V. (1974). *Probabilistic Model of Language. On the Relationship between Natural and Artificial Languages*. Moscow: Nauka. (In Russ.). EDN: QXGCTG
11. Saakyan, L.N. (2025). AI as a Humanized Subject: Mitigative Strategies of Artificial Intelligence. *Communication Studies*, *12*(4), 789–804. (In Russ.). https://doi.org/10.24147/2413-6182.2025.12(4).789-804 EDN: CJFFSS
12. Studenikina, K.A. (2022). Evaluation of the Language Ability of Neural Models Based on Predicative Agreement in Russian. *Proceedings of the Institute for System Programming of the RAS*, *34*(6), 179–184. (In Russ.). https://doi.org/10.15514/ispras-2022-34(6)-14 EDN: UCJBEV
13. Telpov, R.E., & Lartsina, S.V. (2023). Typical Differences between Natural and Neural Network-Generated Texts in a Quantitative Aspect. *Scientific Dialogue*, *12*(7), 47–65. (In Russ.). https://doi.org/10.24224/2227-1295-2023-12-7-47-65 EDN: ZZRRKN
14. Chomsky, N. (1962). Syntactic Structures. In: *New in Linguistics* (Vol. 2, pp. 412–527). Moscow. (In Russ.).
15. Shavrina, T.O. (2021). On the Methods of Computational Linguistics in Evaluating Artificial Intelligence Systems. *Voprosy Jazykoznanija*, 6, 117–138. (In Russ.). https://doi.org/10.31857/0373-658X.2021.6.117-138 EDN: PFGVQJ
16. Sharandin, A.L. (2015). Consciousness, Semiotics, Language, Communication and Pragmatics as a Single Conceptual Space. *Journal of Psycholinguistics*, 3(25), 240–250. (In Russ.). EDN: UDLHJT
17. Shevchenko, E.V. (2015). Simulacrum as a Tactical Method of Rendering an Orienting Influence. *Issues of Cognitive Linguistics*, 3(44), 144–150. (In Russ.). EDN: TTYPMF
18. Pavlovskaya, O.E., & Ilyina, L.E. (2023). Text Entropy as Meaning Dispersion. *Vestnik of Northern (Arctic) Federal University. Series: Humanitarian and Social Sciences*, *23*(6), 77–86. (In Russ.). https://doi.org/10.37482/2687-1505-V308 EDN: ZPZZPW
19. Shelestyuk, E.V., & Shchetinkina, E.A. (2023). Stochasticity and Entropy in Linguistics. *Bulletin of Chelyabinsk State University*, 2, 150–165. (In Russ.). EDN: WIVZGY
20. Andrienko, T.P. (2017). Text Entropy as a Prerequisite for the Implementation of Translation Strategy. *Proceedings of Francysk Skorina Gomel State University*, 1 (100), 76–80. (In Russ.). EDN: YFYPLF
21. Baroni, M. (2022). On the Proper Role of Linguistically-Oriented Deep Net Analysis in Linguistic Theorizing. *Algebraic Approaches to Sequence Data*. https://doi.org/10.1201/9781003205388-1
22. Chomsky, N., Roberts, I., & Watumull, J. (2023). Noam Chomsky: The False Promise of ChatGPT. *The New York Times*, 8 March.

23. Elek, A., Yildiz, H.S., Akca, B., Oren, N.C., & Gundogdu, B. (2025). Evaluating the Efficacy of Perplexity Scores in Distinguishing AI-Generated and Human-Written Abstracts. *Academic Radiology*, *32*(4), 1785–1790. https://doi.org/10.1016/j.acra.2025.01.017 EDN: HMLYED
24. Jiang, S., & Li, K. (2025). Another Kind of Authenticity: the Visual Simulacra of Artificial Intelligence. *Digital Creativity*, *36*(1), 75–91. https://doi.org/10.1080/14626268.2025.2467087 EDN: ZRIUEB
25. Konzack, L. (2025). Generative AI, Simulacra, and the Transformation of Media Production. *Athens Journal of Mass Media & Communications*, *11*(3). https://doi.org/10.30958/ajmmc.11-3-3 EDN: BYIUCE
26. Lyon, C.F. (2023). Fake Cases, Real Consequences: Misuse of ChatGPT Leads to Sanctions. *NY Litigator*, *28*(2), 8–12.
27. Mahowald, K., Ivanova, A.A., et al. (2024). Dissociating Language and thought in Large Language Models: a Cognitive Perspective. *Trends in Cognitive Sciences*. https://doi.org/10.1016/j.tics.2024.01.011 EDN: SKTATZ
28. Miralles-González, P., Huertas-Tato, J., Martín, A., & Camacho, D. (2025). Not All Tokens are Created Equal: Perplexity Attention Weighted Networks for AI Generated Text Detection. https://doi.org/arXivpreprint arXiv:2501.03940
29. Pudasaini, S., Miralles-Pechuán, L., Lillis, D., & Llorens Salvador, M. (2024). Survey on AI-generated Plagiarism Detection: The Impact of Large Language Models on Academic Integrity. *Journal of Academic Ethics*. https://doi.org/10.48550/arXiv.2407.13105
30. Ram, A. (2023). *Generative AI and Text Creation*. Preprint. https://doi.org/10.48550/arXiv.2407.13105

Сведения об авторах:
Новикова Марина Львовна, доктор филологических наук, профессор кафедры русского языка и лингвокультурологии Института русского языка, Российский университет дружбы народов (117198, Российская Федерация, г. Москва, ул. Миклухо-Маклая, д. 6); *сфера научных интересов*: искусственный интеллект; машинный перевод; нейросеть; большая языковая модель; цифровое присутствие; *e-mail*: novikova-ml@rudn.ru
ORCID: 0000-0002-4673-067X, SPIN-код: 1502-9636, AuthorID: 499765.
Новиков Филипп Николаевич, кандидат филологических наук, доцент кафедры иностранных языков юридического института, Российский университет дружбы народов (117198, Российская Федерация, г. Москва, ул. Миклухо-Маклая, д. 6); *сфера научных интересов*: искусственный интеллект; машинный перевод; нейросеть; большая языковая модель; цифровое присутствие; *e-mail*: novikov-fn@rudn.ru
ORCID: 0000-0003-4884-3659, Scopus: 57191418034, SPIN-код: 3523-2539, AuthorID: 671112.

Information about the authors:
Marina L. Novikova, Dr.Sc. (Philology), Professor, Russian Language and Cultural Studies Department, Russian Language Institute, RUDN University (6 Miklukho-Maklaya str., Moscow, Russian Federation117198); *Research interests*: artificial intelligence; machine translation; neural networks; large language models; digital presence; *e-mail*: novikova-ml@rudn.ru
ORCID: 0000-0002-4673-067X, SPIN-code: 1502-9636, AuthorID: 499765.
Philipp N. Novikov, PhD in Philology, Associate Professor, Law Institute Department of Foreign Languages, RUDN University (6 Miklukho-Maklaya str., Moscow, Russian Federation117198); *Research interests*: artificial intelligence; machine translation; neural networks; large language models; digital presence; *e-mail*: novikov-fn@rudn.ru
ORCID: 0000-0003-4884-3659, Scopus: 57191418034, SPIN-code: 3523-2539, AuthorID: 671112.

RUDN Journal of Language Studies, Semiotics and Semantics
2313-2299 (Print), ISSN 2411-1236 (Online)
Вестник Российского университета дружбы народов.
Серия: ТЕОРИЯ ЯЗЫКА. СЕМИОТИКА. СЕМАНТИКА

2026 Vol. 17 No. 2 517–533
http://journals.rudn.ru/semiotics-semantics

DOI: 10.22363/2313-2299-2026-17-2-517-533
EDN: LMPPPM

Research Article / Научная статья

Exploring Cross-Cultural Knowledge Transfer: AI-Integrated Semantic Framework

Maria I. Kiose[1,2], Zhenishkul S. Khulkhachieva[1], Artem V. Barmin[1], Berd Amirkhanov[1]

[1]Moscow State Linguistic University, *Moscow, Russian Federation*
[2]Institute of Linguistics RAS, *Moscow, Russian Federation*
✉maria_kiose@mail.ru

Abstract. The study develops a semantic framework with AI-integrated component to identify the knowledge domains transferred from donor into recipient language. Two types of knowledge transfer are considered, direct and indirect, which provide the transfer via loanwords and via their collocates in discourse. The developed framework is used to explore cross-cultural knowledge transfer from Russian as donor language into Kyrgyz as recipient language. To proceed, a clustering Word2vec algorithm is applied to identify the loanword frequent collocates, next, the semantic classes of the collocates of more and less frequent loanwords are determined. The results featuring the semantic classes distribution in the collocates (overall 8,515 in two lists) in web discourse disclose similar and distinct domains of knowledge transfer from Russian into Kyrgyz. The common knowledge domains are mental sphere and knowledge, people, buildings, appliances and routine objects, special event, transport, texts, class names, collective objects, sport, money and finance, measure units, interaction, space and location, substances. Indirect transfer follows two patterns, discrete found in most frequent loanwords which mediate a restricted number of knowledge domains, and continuous found in less frequent loanwords which mediate a larger number of knowledge domains.

Keywords: semantic clustering; semantic tagging; cross-cultural knowledge transfer; loanword; donor language; recipient language

Founding: The research is part of the FSFU-2025-0004 Project “Assessment of cultural integration and disintegration processes in CIS countries: the study of communicative practices” which is being carried out at Moscow State Linguistic University.

Authors’ contribution: The authors contributed equally to this article and preparation of the article text**.**

Conflict of interests: The authors stated that there were no conflicts of interest.

Article history: Received: 20.01.2026; Accepted: 10.02.2026.

© Kiose M.I., Khulkhachieva Zh.S., Barmin A.V., Amirkhanov B., 2026
This work is licensed under a Creative Commons Attribution-NonCommercial 4.0 International License
https://creativecommons.org/licenses/by-nc/4.0/legalcode

For citation: Kiose, M.I., Khulkhachieva, Zh.S., Barmin, A.V., & Amirkhanov, B. (2026). Exploring Cross-Cultural Knowledge Transfer: AI-Integrated Semantic Framework. *RUDN Journal of Language Studies, Semiotics and Semantics, 17*(2), 517–533. https://doi.org/10.22363/2313-2299-2026-17-2-517-533 EDN LMPPPM

Изучение кросс-культурного трансфера знания: семантическое решение с встроенным ИИ-компонентом

М.И. Киосе[1,2], Ж.С. Хулхачиева[1], А.В. Бармин[1], Б. Амирханов[1]

[1]Московский государственный лингвистический университет, *Москва, Российская Федерация*
[2]Институт языкознания РАН, *Москва, Российская Федерация*
✉maria_kiose@mail.ru

Аннотация. Цель исследования — разработка семантического решения с встроенным ИИ-компонентом, с помощью которого определяется состав сфер знания, подвергающихся межкультурному трансферу из языка-донора в язык-реципиент. Рассматриваются два типа трансфера знания: прямой и непрямой, обеспечивающие передачу знания с помощью самих заимствованных слов и их коллокатов в дискурсе. Выработанное решение позволяет установить состав таких сфер и особенности их распределения в отношении заимствованной лексики из русского языка в киргизский. Задача решается с применением алгоритма Word2vec и последующей семантической разметки коллокатов заимствованных слов с учетом их частотности в языке-реципиенте. Анализ семантических классов двух собранных датасетов контекстов заимствованных слов и их коллокатов (общим числом 8,515 в двух датасетах более и менее частотных заимствований) в новостном дискурсе выявил сходные и отличительные сферы прямого и непрямого трансфера знания из русского языка в киргизский. Основными сферами являются следующие: ментальная сфера и знание, лица, здания и сооружения, инструменты и приспособления, бытовые предметы, мероприятия, транспортные средства, тексты, классы объектов, множества и совокупности объектов, спорт, деньги и финансы, единицы измерения, взаимодействие и взаимоотношение, пространство и место, вещества и материалы. Установлены два паттерна непрямого трансфера знания: дискретный, характерный для наиболее частотных заимствований, проявляющийся в том, что с их помощью происходит трансфер знания только в отношении ограниченного количества сфер знания, и непрерывный, характерный для менее частотных заимствований, проявляющийся в том, что с их помощью обеспечивается трансфер знания в отношении значительно большего количества сфер знания.

Ключевые слова: семантическая кластеризация, семантическая разметка, межкультурный трансфер знания, заимствование, язык-донор; язык-реципиент

Финансирование: Исследование выполняется в Московском государственном лингвистическом университете в рамках Госзадания Минобрнауки России, проект FSFU-2025-0004 «Диагностика процессов культурной интеграции и дезинтеграции в странах СНГ: анализ коммуникативных практик».

Вклад авторов: равнозначен на все этапах исследования и подготовки статьи к публикации.

Конфликт интересов: авторы заявили об отсутствии конфликтов интересов.

История статьи: дата поступления: 20.01.2026; дата приема в печать: 10.02.2026.

Для цитирования: *Kiose M.I., Khulkhachieva Zh.S., Barmin A.V., Amirkhanov B.* Exploring Cross-Cultural Knowledge Transfer: AI-Integrated Semantic Framework // Вестник Российского университета дружбы народов. Серия: Теория языка. Семиотика. Семантика. 2026. Т. 17. № 2. С. 517–533. https://doi.org/10.22363/2313-2299-2026-17-2-517-533 EDN LMPPPM

Introduction

Cross-cultural knowledge transfer is commonly explored in language via loanwords [1–3]; their rate in recipient language is highly indicative of direct transfer from donor to recipient culture. Meanwhile, in discourse the context in which the loanword frequently appears also becomes charged with the transferred knowledge, since the collocates inherit the semantic structure of loanwords, which can be viewed as a form of adoption (imposition) or indirect knowledge transfer [4; 5]. To explore the knowledge transfer via loanwords in large discourse datasets, complex semantic frameworks are required to identify the loanword collocates and their semantic classes displaying the transferred knowledge domains into the recipient language and culture. This study aims to develop the AI-integrated semantic framework to explore cross-cultural transfer performed directly and indirectly via loanwords, and to further reveal the domains of inherited knowledge from Russian into Kyrgyz, two typologically different although politically and culturally interrelated discourses.

One of the potent AI-based natural processing methods is semantic clustering which presupposes identifying groups of objects displaying similarity to each other and difference to other objects in other groups [6]. Existing clustering frameworks establish the foundation for efficient combining AI-based semantic clustering algorithms and other semantic methods in applied linguistics [7; 8]. This study adopts AI-based semantic clustering within a complex framework to explore cross-cultural direct and indirect transfer. The framework comprises 1) AI-based clustering algorithm of Word2vec revealing word clusters exposed in their collocates and 2) semantic taxonomy analysis to attribute the semantic classes (related to knowledge domains) to the collocates. This complex framework is intended to identify the extent of cross-cultural transfer of two types, direct expressed in the semantic integration of loanwords into a different language and culture, and indirect expressed in the semantic integration of loanword collocates. We hypothesize that higher and lower frequency loanwords function in a singular way in cross-cultural interaction.

The study is structured as follows. Theoretical framework section presents the problem of cross-cultural direct and indirect transfer and the developed methodological decision based on natural processing methods of semantic clustering. Methods and procedure section shows the data and the research steps of the study. Results section presents the research outcomes and their implications for cultural transfer exploration. Final remarks section briefly summarizes the key findings of the study.

Theoretical Framework. Direct and Indirect Cross-Cultural Knowledge Transfer

Cultural knowledge transfer is addressed as the integration of culturally specific knowledge domains into a donor culture which can be observed in the distribution of single words and expressions [9]. It also appears in the distribution of loanwords and the knowledge domains they refer to, since "culturally specific transfer of domains and conceptual modifications <…> appear in their exporting and importing from one culture into another" [10. P. 5]. It is notable that both the loanwords which designate a new knowledge domain and an already existing one are considered in the studies [2; 3]. Consequently, we expect that identifying the knowledge domains related to these loanwords will allow to reveal the knowledge modifications in recipient culture. Developing these ideas further, we presume that apart from ascertaining the presence of loanwords in a recipient language, their collocates distribution also matters since it evidences of the degree of culture integration attributed to specific knowledge domains [8]. Therefore, two types of knowledge transfer can be observed, direct, via the knowledge domains of loanword, and indirect, via the loanword collocates. In the present study, we diagnose the integration degree and affected knowledge domains of both direct and indirect transfer in two languages, Kyrgyz and Russian which are known to have close cultural contacts.

Single studies address the problems of lexical transfer via loanwords in Russian and Kyrgyz languages [11]; meanwhile these studies do not appeal to the structure of inherited knowledge. They commonly explore the loanword transfer within two frameworks: lexicography aimed to compile the dictionaries of loanwords, and lexical semantics aimed at attributing the loanword meaning. Following these findings, we can explore the direct and indirect knowledge transfer in contemporary discourse adopting the lists of loanwords identified and listed in the compiled etymological dictionaries of the Kyrgyz language[1]; meanwhile, these dictionaries being compiled at the end of the XXth century (there are no up-to-date etymological dictionaries of Kyrgyz) do not reveal recent inputs into Kyrgyz, which is definitely a research limitation.

Semantic Framework with AI-Integrated Clustering Decision

To explore cross-cultural transfer in large datasets, we have to develop the semantic framework capable of processing the loanwords in contexts. This problem is currently addressed via natural processing methods which have now become integrated into applied linguistic studies, with semantic clustering being one of the most potent in grouping words. Semantic clustering performs three major tasks: a) underlining the pattern and insight into the data, b) identifying the degree of

[1] Yudakhin, K.K. (1985). *Kyrgyz-Russian Dictionary*. Frunze: Glavnaya redaktsiya Kirgyzskoy Sovetskoy Entsiklopedii. (In Russ.); Karasaev, Kh.K. (1986). *The Dictionary of Loanwords*: 5100 words. Frunze: Kyrgyz Sovet. (In Kyrgyz).

similarity between data points, c) data organization and summarization through cluster prototypes [6]. Multiple studies present the evolution and trends in data clustering; to proceed, they focus on algorithms for grouping data sets used in statistics, computer science and machine learning (review see [12]). Importantly, clustering decisions have become complex or hybrid methods integrating both AI-based techniques and other clustering methods (e.g. automatic K-means clustering and semantic clustering) [13]. Semantic clustering is at present a natural language processing method employed in investigating semantic relationship through the shared meanings the words maintain in context [14; 15]; it allows to cluster words through word embeddings (Word2vec) or through transformer-based models (BERT) by calculating their similarity in the vector space.

In applied linguistics, semantic clustering has been extensively used in recent decades for different purposes. In Divjak [16], a clustering decision for lexical groups was developed. It was later extended to cluster the languages via the use of perception semantics in verbs, where D. Divjak applied random forest model to identify the predictors of each perception type in the samples [17]. It is notable that a linguistic foundation for clustering methods in lexical groups lie in prototypicality effects explored vastly since the pioneering works of Rosch [18] and Geeraerts [19] who claimed that the centers of semantic categories could be revealed in psychological experiments. These ideas were taken further due to developing the corpus linguistic methods dealing with big data [20]. At present, LLM (large language models) operate performing natural language processing including text understanding and generation (for review see [21]), for instance to perform text clustering based on deep semantic understanding [6; 22]. Semantic clustering as another of natural language processing methods is now extensively used in education and psychology studies [7; 23]. Developing its applied potential in cross-cultural studies, we address the view promoted by Kuhn et al [14] who regard lexical semantic clusters as groups of source artifacts that use similar vocabulary; therefore, these are the common knowledge domains which underlie cluster formation.

Meanwhile, identifying the clusters of loanwords and their collocates in exploring indirect transfer is not enough to reveal these knowledge domains. Additionally, semantic taxonomy has to be implemented to annotate the collocates and to further determine the semantic classes related to these knowledge domains. To explore the semantic classes of both loanwords and their target (frequently used) collocates, their overall taxonomy is required. Several major approaches have been developed to proceed, which are lexical-semantic analysis, frame analysis, image-schema, qualia analysis, which display the taxonomic classes of words within different types of conceptual domains. In [5] a complex semantic framework with AI-integrated clustering decision was already tested with the second component being qualia-analysis. Since in this study we intend to identify the knowledge domains inherited from or mediated by the donor culture, we appeal to lexical-semantic taxonomy proposed in the National Corpus of the Russian Language

(NCRL)[2] and identify the distribution of these taxonomic classes attributed to the loanwords in two datasets, with higher and lower corpus frequency, and to their target collocates. Following the postulates of cultural transfer studies [10], we presume that revealing the knowledge distribution in related or distant cultures and languages allows to identify the cultural transfer routes and the affected knowledge domains.

Methods and Procedure

Following the decisions elaborated above, we propose a two-component semantic framework to explore indirect cross-cultural transfer via loanwords. To obtain a list of collocates with their numerical scores, we use an AI-based algorithm of semantic clustering (Word2vec). To further identify the knowledge domains mediated by indirect transfer, we apply the semantic taxonomy protocol devised in NCRL, to annotate the semantic classes of collocates. Since we explore both direct and indirect transfer, we apply semantic taxonomy to annotate both loanwords in two lists, more and less frequent, as part of direct transfer analysis, and to annotate loanword collocates in two lists attributed to more and less frequent loanwords. We presume that higher and lower frequency loanwords promote knowledge transfer via different domains of direct and indirect transfer thus functioning in a singular way in cross-cultural interaction.

The research procedure involves the following steps. At step 1, we compile two lists of loanwords from Russian into Kyrgyz following their corpus frequency. At Step 2, we compile the datasets with their corpus contexts. At Step 3, via semantic clustering procedure the target collocates with their numerical scores are identified. At Step 4, we determine the semantic taxonomic distribution of both the domains of inherited in loanwords and modulated knowledge in collocates, to further contrast the data and verify the hypothesis claiming the existing differences in their distribution in two lists. Consequently, the outcomes of the study include i) two lists of loanwords from Russian into Kyrgyz distinct in their corpus frequency, ii) two datasets of corpus contexts of these loanwords, iii) two lists of target collocates of these loanwords with their numerical scores obtained via AI semantic clustering, iiii) the determined differences in the taxonomic distribution of both the domains of inherited and mediated knowledge transferred from Russian into Kyrgyz.

The research data comprise two macro blocks: a) a compiled list of loanwords borrowed from Russian into Kyrgyz which fall into two lists different in their web corpus frequency, b) two compiled datasets of their contexts elicited from the web corpus. To prepare a list of loanwords from Russian into Kyrgyz, two existing dictionaries of the Kyrgyz language were addressed: 1) “The Dictionary of

[2] National Corpus of the Russian Language. https://ruscorpora.ru/page/instruction-semantic (accessed: 10.01.2026). (In Russ.).

Loanwords" compiled by Kh. Karasaev[3] which contains 5.100 loanwords; 2) "The Dictionary of the Kyrgyz Language" compiled by K. Yudakhin[4] which contains 100.000 words. To obtain the datasets of loanword collocates, we addressed Kyrgyz Web text corpus (Kyrgyzstan)[5] which is part of the open-access Leipzig Corpora, and which comprises 3.036.362 sentences and 36.581.130 tokens. A programmed decision was further elaborated to form the context dataset, since the corpus does not have a lemmatizer[6]. Next, AI-based clustering decision was applied to cluster the loanwords via their collocates. Word2vec word-embedding algorithm was used to vectorize the datasets data, which allowed to determine the closest neighbors of each loanword in each of its word forms in each of the datasets. We further manually performed semantic annotation of target collocates in two lists of target words. To proceed, we applied the taxonomy developed for NCRL. In accordance with it, single semantic classes were attributed to autosemantic parts of speech. While this taxonomy is easier attributed to names, its use to semanticize attributes, adverbs and verbs is less regulated since their meaning is highly contextualized. For instance, in NCRL the semantics of *актуальный* (topical) is restricted only to one class 'relative', which will not allow to identify its semantic attribution. In our case, when dealing with target collocates, the procedure of context analysis could not be accomplished. For this reason, only names were classified in this study.

Results.
Direct Knowledge Transfer: Knowledge Domains in Loanwords

All the words in Karasaev[7] and in Yudakhin[8] marked as having Russian as immediate etymon were selected, their number being equal to 1.204. Next, their corpus frequency in the base form was identified, which allowed to segment the data into three lists, highly frequent loanwords (64, corpus frequency in the base form ranging from 800 to 13,500), less frequent (74, corpus frequency in the base form ranging from 300 to 799) and infrequent (corpus frequency in the base form not exceeding 299). The loanwords from Lists 1 and 2 were further annotated using the semantic taxonomy presented in NCRL.

In Figure 1 we present the semantic taxonomic distribution of the loanwords in two lists.

[3] Karasaev, Kh.K. (1986). *The Dictionary of Loanwords*: 5100 words. Frunze: Kyrgyz Sovet. (In Kyrgyz).

[4] Yudakhin, K.K. (1985). *Kyrgyz-Russian Dictionary*. Frunze: Glavnaya redaktsiya Kirgyzskoy Sovetskoy Entsiklopedii. (In Russ.).

[5] Kyrgyz Web Text Corpus. Leipzig Corpora URL: https://wortschatz.uni-leipzig.de/en (accessed: 10.01.2026).

[6] The authors are grateful to Daniil Sinkevich for his help with developing the software used to compile the context dataset.

[7] Karasaev, Kh.K. (1986). *The Dictionary of Loanwords*: 5100 words. Frunze: Kyrgyz Sovet. (In Kyrgyz).

[8] Yudakhin, K.K. (1985). *Kyrgyz-Russian Dictionary*. Frunze: Glavnaya redaktsiya Kirgyzskoy Sovetskoy Entsiklopedii. (In Russ.).

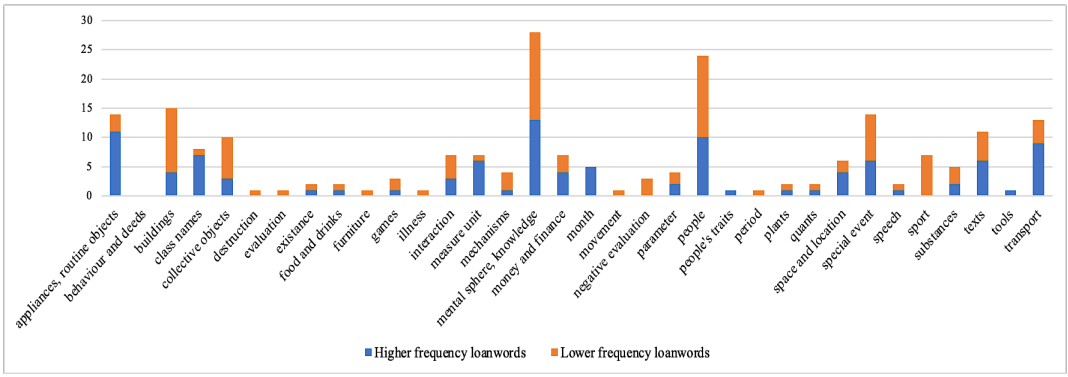


Figure 1. Semantic taxonomic distribution of the loanwords, abs.
Source: compiled by Maria I. Kiose, Zhenishkul S. Khulkhachieva, Artem V. Barmin, Berd Amirkhanov.

As seen, the most frequent semantic classes of loanwords are 1) mental sphere, knowledge, 2) people, 3) buildings, 4) appliances, routine objects, 5) special event, 6) transport. "Mental sphere, knowledge" is expressed in *план* (plan), *физика* (physics), *экономика* (economics); in multiple cases it is represented integrally with the class "people", e.g. in *профессор* (professor), *парламент* (parliament), *партия* (party), where the latter two also manifest the class "collective objects". The semantic class "buildings" is found in *клуб* (club), *аэропорт* (airport), *база* (base). The class "appliances, routine objects" appears in *карта* (card, map), *мандат* (ticket), *товар* (goods), *техника* (tools, technics), where the latter two additionally manifest the class "class names". The loanwords *олимпиада* (Olympiad), *семинар* (seminar), *экзамен* (exam) serve to represent the semantic class "special event", while the sphere "transport" is found in *поезд* (train), *велосипед* (bike), *такси* (taxi).

The major differences in the distribution of these semantic classes within the two lists of loanwords lie in the classes "transport", "appliances, routine objects", "month", "class names", "measure unit", which prevail in higher frequency loanwords list; while the classes "buildings", "collective objects", "sport" display prevalence in the list of lower frequency words. Importantly, as the results show, lower frequency words display the single cases of use attributed to multiple semantic classes related to states, communication, evaluation. It is noticeable that the loanwords related to the class "food" are less numerous and less frequent, which shows the prevalence in Kyrgyz culture to its own culinary traditions. These results may serve to identify the semantic classes concomitant with the knowledge domains majorly affected by direct knowledge transfer within the Kyrgyz culture, and the ones less affected, still presumably due to their less significant role in culture.

Indirect Knowledge Transfer: Knowledge Domains

To compile the datasets of contexts with more and less frequent loanwords, we used the Kyrgyz Web text corpus (Kyrgyzstan)). Meanwhile, since this corpus is not equipped with a lemmatizer, the lists of word forms were compiled; they

contained 825 word forms for 64 most frequent loanwords, and 878 word forms for 74 less frequent ones. The two compiled datasets of contexts with more and less frequent loanwords comprise 348.084 contexts (Dataset 1) and 145.039 contexts (Dataset 2).

At this step, to perform the AI-based semantic clustering of loanwords and their neighbors (collocates), we applied a word-embedding algorithm using *word2vec* R-package v.0.4.0. in RStudio v.2024.12.1+563 with R v.4.3.2. The following parameters of the model were set: type of word2vec algorithm (type) = skip-gram; dimensions of word vectors (dim) = 300; number of times a word should occur to be considered as part of the training vocabulary (min_count) = 1. The model was trained on web samples using the "predict" function in the *word2vec* R-package to determine the five nearest neighbors for each of the loanword forms based on their vector similarity in the 300-dimensional space. We also collected the numerical scores of vector similarity between each of the loanwords and their five nearest neighbors. To visualize the results, we uploaded the three-column matrix to Gephi v.0.10.1. — an open-source software for network visualization and analysis and then used the Fruchterman-Reinhold layout algorithm to better optimize the resulting network.

Figure 2 shows a fragment of semantic net with the word cluster produced by the loanword *макси* (taxi) displaying the semantic class "transport" prevailing within the loanwords.

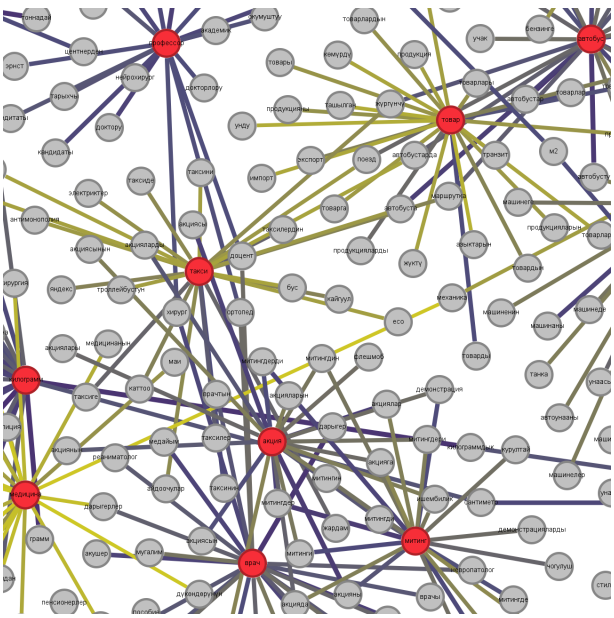


Figure 2. Fragment of Semantic Net
Source: compiled by Maria I. Kiose, Zhenishkul S. Khulkhachieva, Artem V. Barmin, Berd Amirkhanov.

As seen, the loanword *макси* (taxi) forms semantic nets with the loanwords related to other knowledge domains via semantic classes, here *профессор*

(professor) and *врач* (doctor) related to mental sphere and people, *акция* (action, share) related to special event, money and finance, *товар* (goods) referring to appliances, routine objects, collected objects, *килограмм* (kilo) attributed to measure unit, etc., which overall gives rise to multiple semantic connections expressed in collocates.

Two lists of target collocates for more and less frequent loanwords (Dataset 1 and Dataset 2) were further obtained, with the first containing 4,125 words and the second containing 4.390 words. Next, we proceeded to semantic classification of target collocates of these loanwords in both lists. All the target collocates in two lists were attributed to one or more of these taxonomy classes following their basic dictionary meaning. To attribute the semantic class to a lexeme, we applied a two-stage procedure. First, we addressed the basic dictionary meaning / meanings of a word, which allowed to identify a single semantic class or a group of semantic classes included in the taxonomy. Next, we consulted the semantic classes attributed to a lexeme in NCRL. At this step, we specified or added the semantic classes to the words. Using a one-step procedure (addressing only NCRL) appeared less reliable. For instance, *опера* (opera) is semanticized as 1) music and drama genre, 2) staged performance; therefore, the semantic classes of mental sphere, special event and space and location are attributed to it. It is notable that in NCRL *опера* (opera) has the following semantic attribution: semantics: a piece of music, object name; additional semantics: notional name; where the class "a piece of music" although attributed is not included in the semantic taxonomy, and other two classes are utterly common and do not display semantic taxonomic distribution. Therefore, although in major cases we consulted the semantic taxonomy of single words presented in NCRL, we mostly addressed the meanings listed in the dictionary to attribute the semantic classes. Below we present the results displaying the distribution of the semantic classes corresponding to affected knowledge domains.

List one with the collocates of most frequent loanwords includes 4.126 target collocates; with 3.525 nouns and nominals which were further semanticized using NCRL semantic protocol. To annotate the semantic classes in these collocates, the same procedure applied to annotate the loanwords was used; 5.392 tags were performed. Figure 3 shows the distribution of semantic classes tagged in the collocates.

The results show that the collocates semantic classes can be further grouped considering their frequency in the compiled collocates list. The collocates of higher frequency (with values exceeding one third of the highest frequency value) manifest the classes "mental sphere, knowledge", "people", "appliances, routine objects", "texts". The collocates of lower frequency (with values exceeding one tenth of the highest frequency value) relate to the classes "special event", "collective objects", "transport", "money and finance", "class names", "measure unit", "buildings", "month", "parameter", "space and location", "substances", "interaction", "mechanisms". The least frequent collocates relate to a wide range of semantic

classes attributed both to tangible and intangible values. Additionally, several occasional classes were identified but due to their infrequent use were not involved in the diagram.

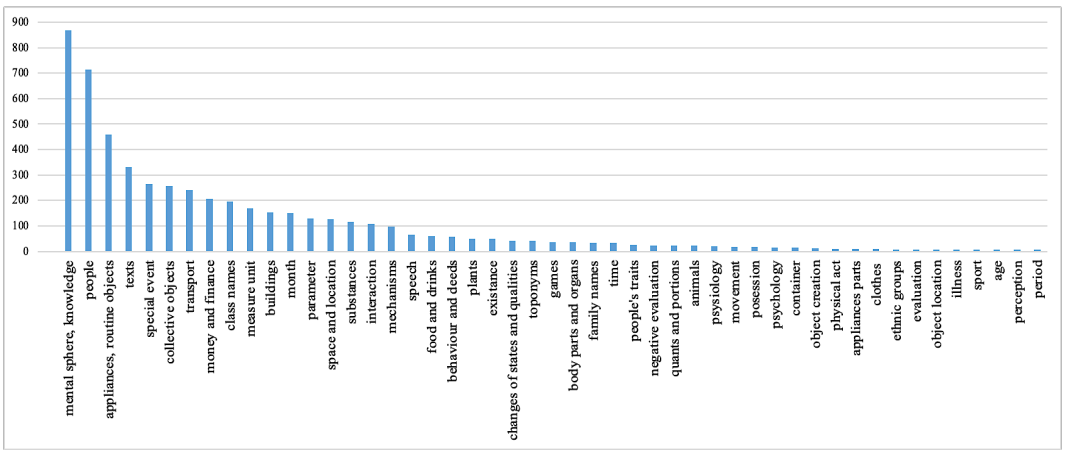


Figure 3. Semantic Taxonomy of the List one Loanword Collocates, abs.
Source: compiled by Maria I. Kiose, Zhenishkul S. Khulkhachieva, Artem V. Barmin, Berd Amirkhanov.

Next, we contrasted these classes distribution with the ones tagged in the loanwords of higher frequency (Figure 4), in both cases relative values were used.

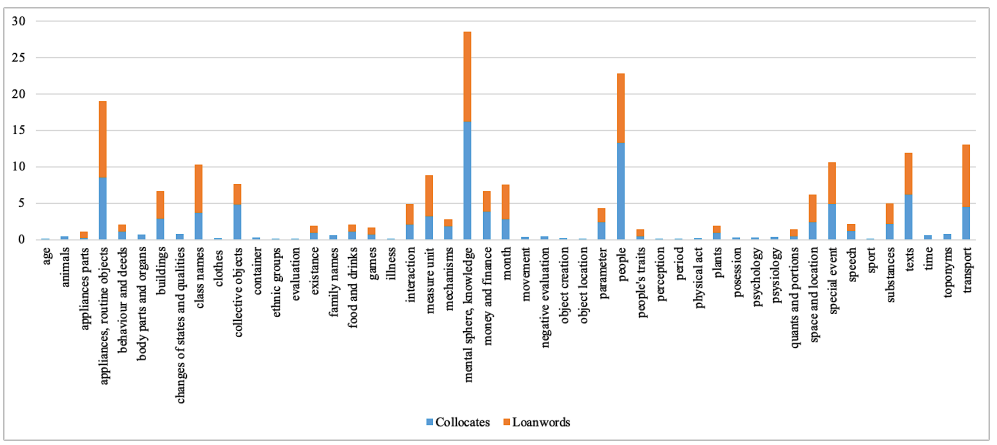


Figure 4. Contrastive Semantic Taxonomic Distribution of the Loanwords and Collocates, rel.
Source: compiled by Maria I. Kiose, Zhenishkul S. Khulkhachieva, Artem V. Barmin, Berd Amirkhanov.

With the loanword semantic taxonomy manifesting direct transfer and the collocates semantic taxonomy manifesting indirect transfer, the data help contrast the direct and indirect knowledge transfer affected by highly frequent loanwords. The results show that the difference in the distribution of higher frequency classes lies in “class names”, “measure unit”, “month”, “space and location”, “transport” which prevail in direct transfer, and “collective objects”, “money and finance”,

"texts" which prevail in indirect transfer; meanwhile, multiple χ^2 tests did not show any significant differences in their distribution.

Next, we addressed the distribution of semantic classes within the collocates of less frequent loanwords. List two with the collocates of less frequent loanwords includes 4.390 target collocates; with 3.067 nouns and nominals which were further semanticized. It is notable that List two with collocates involves a larger number (compared to List one) of names (of companies, brands) which were not semanticized. To annotate the semantic classes in these collocates, the same procedure applied to annotate the loanwords was used; 3.874 tags were found. Figure 5 shows the distribution of semantic classes tagged in the collocates.

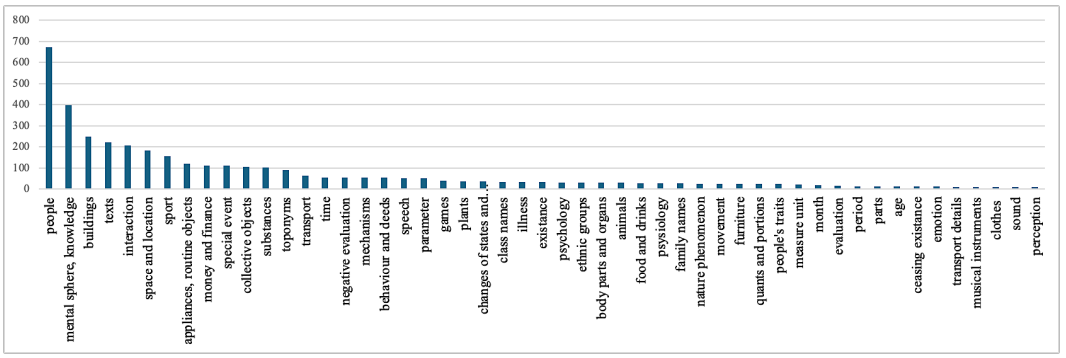


Figure 5. Semantic Taxonomy of the List Two Loanword Collocates, abs.
Source: compiled by Maria I. Kiose, Zhenishkul S. Khulkhachieva, Artem V. Barmin, Berd Amirkhanov.

The results show that the collocates of higher frequency (with values exceeding one third of the highest frequency value) manifest the classes "people", "mental sphere, knowledge", "buildings". The collocates of lower frequency (with values exceeding one tenth of the highest frequency value) relate to the classes "texts", "interaction", "space and location", "sport", "special event", "collective objects", "money and finance", "substances", "appliances, routine objects", "toponyms". Infrequent semantic classes are numerous; meanwhile, we observe the balanced used of tangible and intangible objects.

Importantly, we observe a slightly different picture if contrasted with the distribution of List one loanword collocates. First, we notice a decrease in the frequency of two highly manifested classes, which are "mental sphere, knowledge" and "people", which cannot be linked to the decrease in the presence of these spheres in the List two loanwords since they are equally frequently present in both lists. Presumably, as the results show, while most frequent loanwords display a directed effect, less frequent loanwords encompass a wider range of semantic classes. These differences show two distinct patterns of indirect transfer — discrete with single affected knowledge domains found with frequent loanwords, and continuous with multiple affected knowledge domains found with less frequent loanwords. It is noticeable although that Mann-Whitney U test did not show significant differences in the data distribution in two collocate lists (U=1743,

p=0.765), which proves the prevalence of the same semantic classes and infrequent presence of similar ones.

Next, we contrasted these classes distribution with the ones tagged in the loanwords of lower frequency (Figure 6), in both cases relative values were used.

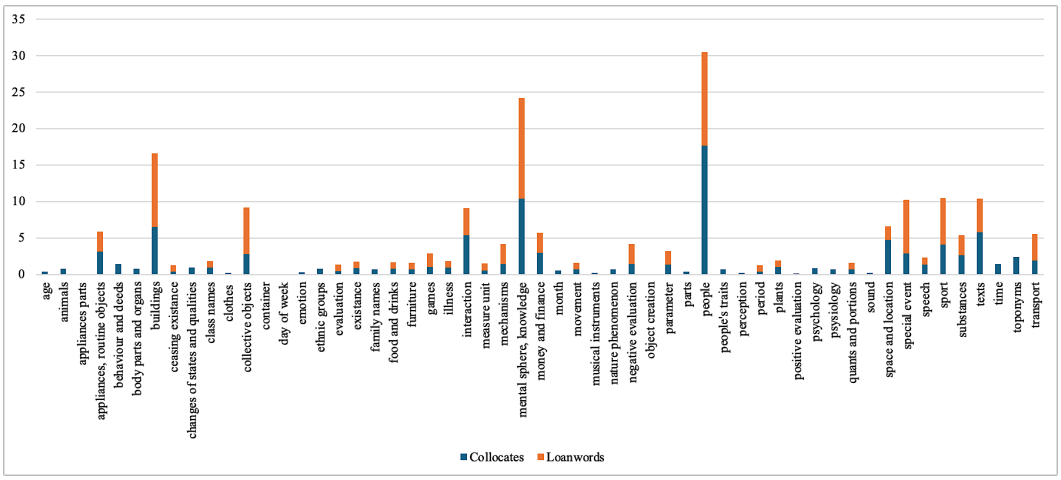


Figure 6. Contrastive Semantic Taxonomic Distribution of the Loanwords and Collocates, rel.
Source: compiled by Maria I. Kiose, Zhenishkul S. Khulkhachieva, Artem V. Barmin, Berd Amirkhanov.

The results allow to identify the differences in direct and indirect transfer via lower frequency loanwords. Higher differences are observed in the semantic classes "special event", "collective objects", "sport", "transport" which prevail in direct transfer, and the classes "interaction", "space and location", which prevail in indirect transfer. In line with the results obtained with higher frequency loanwords, multiple χ^2 tests did not show any significant differences in their distribution due to a large number of classes.

Meanwhile, the contrastive analysis helped formulate several regularities in the direct and indirect knowledge transfer. First, there are two semantic classes which are largely affected by both direct and indirect knowledge transfer in higher and lower frequency loanwords, which are "mental sphere, knowledge" and "people"; other less affected classes are "appliances, routine objects" and "buildings". Second, most affected semantic classes display regular correspondences in the effects of direct and indirect transfer, which means that the frequency rate of these semantic classes is similar in direct and indirect transfer. Third, both object and abstract names (representing tangible and intangible values) serve to participate in direct and indirect knowledge transfer; additionally, no regulations were found in their distribution mediated by higher and lower frequency loanwords.

The study helped rank the semantic classes according to their regularity in affecting and being affected by knowledge transfer. Two semantic scales displayed below manifest the direct and indirect knowledge transfer from Russian into Kyrgyz

language and culture featuring the frequency effects displayed in modern web discourse.

Direct knowledge transfer domains (most frequent semantic classes, cases): Mental sphere, knowledge (28) > People (24) > Buildings (15) > Appliances, routine objects (14), Special event (14) > Transport (13) > Texts (11) > Collective objects (10) > Class names (8) > Sport (7), Money and finance (7), Measure unit (7), Interaction (7) > Space and location (6) > Substances (5), Month (5).

Indirect knowledge transfer domains (most frequent semantic classes, cases): People (1387) > Mental sphere, knowledge (1263) > Appliances, routine objects (576) > Texts (551) > Special event (374) > Collective objects (361) > Money and finance (317), Interaction (315) > Space and location (307), Transport (302) > Buildings (296) > Class names (230) > Substances (216) > Measure unit (190) > Parameter (180) > Month (169) > Sport (161) > Mechanisms (151) > Toponyms (132).

Overall, the obtained results do not fully support the research hypothesis claiming that higher and lower frequency loanwords promote knowledge transfer via different domains of direct and indirect transfer thus functioning in a singular way in cross-cultural interaction. The distribution of semantic classes corresponding to knowledge domains clearly showed that while there is a distinction in several knowledge domains prevailing within the collocates of either more or less frequent loanwords (e.g., the redistribution in the classes “people” and “mental sphere, knowledge”), the key knowledge domains remain stable. This means that these knowledge domains are most susceptible to being mediated by knowledge transfer, presumably due to their high activity in language and culture. The results also validate the significance of the notion of knowledge transfer proposed and developed in [9; 10] since it proved to serve a valid methodology to explore the distribution of knowledge domains within a mediated culture. Additionally, they show that the lexical decision of loanwords proposed in [1–3] can serve to explore not only direct, but also indirect knowledge transfer.

Most importantly, the AI-integrated framework was found valid to explore indirect knowledge transfer, which paves the way for new applications of semantic clustering. Therefore, besides the applications shown in [7; 14; 23], cross-cultural integration can be efficiently determined with its help. Additionally, the second component of the developed semantic framework, which is semantic taxonomy application proposed in NCRL, was found equally effective in attributing semantic classes to collocates and identifying the distribution of affected knowledge domains. Meanwhile, the research limitation is definitely the necessity to use the etymological dictionaries[9,10] [which do not enlist the loanwords of the XXIst century.

[9] Karasaev, Kh.K. (1986). *The Dictionary of Loanwords*: 5100 words. Frunze: Kyrgyz Sovet. (In Kyrgyz).

[10] Yudakhin, K.K. (1985). *Kyrgyz-Russian Dictionary*. Frunze: Glavnaya redaktsiya Kirgyzskoy Sovetskoy Entsiklopedii. (In Russ.).

Final Remarks

The study proposed a two-component semantic framework to explore cross-cultural knowledge transfer mediated by more and less frequent loanwords. The framework comprising AI-based semantic clustering algorithm (Word2vec) and semantic taxonomy analysis allowed to determine the knowledge domains of direct and indirect knowledge transfer. Two scales of most frequent knowledge domains susceptible to knowledge transfer were deduced, with the knowledge domains corresponding to semantic classes "people" and "mental sphere, knowledge" being most frequent. Additionally, the study determined several regularities in direct and indirect knowledge transfer. Importantly, indirect transfer follows two patterns, discrete found in most frequent loanwords which mediate a restricted number of knowledge domains, and continuous found in less frequent loanwords which mediate a larger number of knowledge domains. Overall, the results contribute to developing the procedure of AI-based methods integration into semantic frameworks, with semantic clustering being efficient in exploring cross-cultural knowledge transfer. The research perspectives are seen both in verifying the AI-based semantic framework in cultural studies, and in exploring cross-cultural transfer in other languages and cultures.

References / Список литературы

1. Myers-Scotton, C. (2002). *Contact linguistics: Bilingual Encounters and Grammatical Outcomes*. Oxford: Oxford University Press.
2. Myers-Scotton, C. (2006). *Multiple Voices: An Introduction to Bilingualism*. Malden, MA: Blackwell.
3. Haspelmath, M. (2009). Lexical Borrowings: Concepts and Issues. In: M. Haspelmath, U. Tadmor (eds.) *Loanwords in the World's Languages* (pp. 35–54). Berlin: De Gruyter.
4. Thomason, S.G. (2001). *Language Contact*. Washington, D.C.: Georgetown University Press.
5. Kiose, M.I., Khulkhachieva, Zh.S., Izyumskaya-Kapitonova, V.V., & Barmin, A.V. (2025). Lexical-Semantic Clustering in the Diagnostics of Culture Integration in Discourse. *Research Result. Theoretical and Applied Linguistics*, *11*(4), 63–84. https://doi.org/10.18413/2313-8912-2025-11-4-0-4 EDN: UBIXHD
6. Singh, J., & Singh, D. (2024). A Comprehensive Review of Clustering Techniques in Artificial Intelligence for Knowledge Discovery. *Advanced Engineering Informatics*, 62, 102799. https://doi.org/10.1016/j.aei.2024.102799 EDN: QKLLAC
7. Pérez-Serrano, M., Nogueroles-López, M., & Duñabeitia, J. A. (2022). Effects of Semantic Clustering and Repetition on Incidental Vocabulary Learning. *Frontiers in Psychology*, 13, 997951.
8. Zhu, Y., Yang, L., Xu, K., Zhang, W., Song, Z., Wang, J., & Yu, P.S. (2025). LLM-MemCluster: Empowering Large Language Models with Dynamic Memory for Text Clustering. *Computation and Language*. https://arxiv.org/abs/2511.15424 (accessed: 15.01.2026).
9. Iriskhanova, O.K., & Kiose, M.I. (2016). Technologies of Interdisciplinary Terms Transfer into Language Studies. In: V.V. Feschenko (Ed.) *Linguistics and Semiotics of Cultural Transfers*: *Methods, Principles, Technologies* (pp. 151–180). Moscow: Cultural revolution. (In Russ.). EDN: YQAGOL

Ирисханова О.К., Киосе М.И. Технологии трансфера междисциплинарных терминов в лингвистику // Лингвистика и семиотика культурных трансферов: методы, принципы, технологии / отв. ред. В.В. Фещенко. М. : Культурная революция, 2016. С. 151–180.

10. Feschenko, V.V., & Bochaver, S.Yu. (2016). Theory of Cultural Transfers: from Translation Studies — through Cultural Studies — to the Theory of Language. In: V. V. Feschenko (Ed.) *Linguistics and Semiotics of Cultural Transfers*: *Methods, Principles, Technologies* (pp. 5–35). Moscow: Cultural revolution. (In Russ.). EDN: YFKVZR
Фещенко В.В., Бочавер С.Ю. Теория культурных трансферов: от переводоведения — через cultural studies — к теоретической лингвистике // Лингвистика и семиотика культурных трансферов: методы, принципы, технологии / отв. ред. В. В. Фещенко. М. : Культурная революция, 2016. С. 5±–35.
11. Kambaralieva, U.D., & Sternin, I.A. (2021). *Russian and Kyrgyz Communicative Behaviour*. Voronezh: Ritm. (In Russ.). EDN: JNYREE
Камбаралиева У.Д., Стернин И.А. Русское и киргизское коммуникативное поведение. Воронеж : Ритм, 2021.
12. José-García, A., & Gómes-Flores, W. (2016). Automatic Clustering Using Nature-Inspired Metaheuristics: a Survey. *Applied Soft Computing*, 41, 192–213.
13. Nguyen, C.D., & Cios, K.J. (2008). GAKREM: A Novel hybrid Clustering Algorithm. *Information Sciences*, 178 (22), 4205–4227.
14. Kuhn, A., Ducasse, S., & Gîrba, T. (2007). Semantic Clustering: Identifying Topics in Source Code. *Information and Software Technology*, *49*(3), 230–243.
15. Witschard, D., Jusufi, I., Martins, R.M., Kucher, K., & Kerren, A. (2022). Interactive Optimization of Embedding-Based Text Similarity Calculations. *Information Visualization*, *21*(4), 335—353. https://doi.org/10.1177/14738716221114372 EDN: JNAUAD
16. Divjak, D. (2010). *Structuring the Lexicon: A Clustered Model for Near-Synonymity*. Berlin: De Gruyter.
17. Divjak, D. (2015). Exploring the Grammar of Perception. A Case Study Using Data from Russian. *Sensory Perceptions in Language and Cognition. Special issue of functions of Language*, *22*(1), 44–68. https://doi.org/10.1075/fol.22.1.03div EDN: URKXYP
18. Rosch, E. (1977). Human Categorization. In: *Studies in Cross-Cultural Psychology*, N. Warren (Ed.) (Vol. 1, pp. 1–49). New York: Academic Press.
19. Geeraerts, D. (1988). Where Does Prototypicality Come from? In: *Topics in Cognitive Linguistics*, B. Rudzka-Ostyn (Ed.) (pp. 207–229). Amsterdam, Philadelphia: John Benjamins.
20. Soloviev, V.D. (2015). Possible Mechanisms of Change in the Cognitive Structure of Semantic Rows. In: *Language and thought. Modern cognitive linguistics* (pp. 478–487). Moscow: Languages of Slavic Culture.
21. Kolmogorova, A.V., & Margolina, A.V. (2024). Written vs Generated Text: “Naturalness” as a Textual and Psycholinguistic Category. *Research Result. Theoretical and Applied Linguistics*, *10*(2), 71–99. https://doi.org/10.18413/2313-8912-2024-10-2-0-4 EDN: CCESHE
22. Hadifar, A., Sterckx, L., Demeester, T., & Develder, C. (2019). A Self-Training Approach for Short Text Clustering. In: *Proceedings of the 4th Workshop on Representational Learning for NLP* (*RepL4NLP-2019*) (pp. 194–199). Italy: Florence.
23. Troyer, A.K., Moscovitch, M., & Winocur, G. (1997). Clustering and Switching as Two Components of Verbal Fluency: Evidence from Younger and Older Healthy Adults. *Neuropsychology, 11*(1), 138–146.

Information about the authors:

Maria I. Kiose, Dr. Sc. (Philology) (Advanced Doctorate), Associate Professor, Chief Researcher of the Centre for Socio-Cognitive Discourse Studies, Moscow State Linguistic University (38 build 1 Ostozhenka str., Moscow, Russian Federation, 119034); Leading Researcher; Laboratory for Multichannel Communication, Institute of Linguistics, RAS

(1 B. Kislovsky ln., Moscow, Russian Federation, 125009); *Research interests*: cognitive semantics, indirect naming, multimodal linguistics; *e-mail*: maria_kiose@mail.ru
ORCID: 0000-0001-7215-0604. Scopus AuthorID: 56642747500; ResearcherID: AAB-7989-2019, SPIN-code: 4419-0090, AuthorID: 341964.
Zhenishkul S. Khulkhachieva, PhD in Philology, Associate Professor, Head of Department of Languages and Cultures of CIS and FSU countries, Head of Ch. Aitmatov Centre for Kyrgyz language and culture, Moscow State Linguistic University (38 build 1 Ostozhenka str., Moscow, Russian Federation, 119034); *Research interests*: linguodidactics, teaching methodology, cognitive linguistics; *e-mail*: j.khulkhachieva@linguanet.ru
ORCID: 0009-0000-3506-2007. SPIN-code: 7897-5998, AuthorID: 1014042.
Artem V. Barmin, Junior Researcher of the Laboratory for Cognitive Studies of Communication of Moscow State Linguistic University (38 build 1 Ostozhenka str., Moscow, Russian Federation, 119034); *Research interests*: semantic memory, network science; *e-mail*: art.barmin1@gmail.com
ORCID: 0000-0001-9658-6621. Scopus Author ID: 58113747500; ResearcherID: AGL-9794-2022; SPIN-code: 4233-4593, AuthorID: 1146203.
Al-Berd I. Amirkhanov is a Research Assistant at the Center for Sociocognitive Research of Discourse, Moscow State Linguistic University (38 build 1 Ostozhenka str., Moscow, Russian Federation, 119034); *Research interests*: political economy, regional development, cross-cultural knowledge transfer; *e-mail*: berd582@mail.ru

Сведения об авторах:

Киосе Мария Ивановна, доктор филологических наук, доцент, главный научный сотрудник Центра социокогнитивных исследований дискурса, Московский государственный лингвистический университет (119034, Российская Федерация, г. Москва, ул. Остоженка, д. 38 с. 1); ведущий научный сотрудник Лаборатории мультиканальной коммуникации, Институт языкознания РАН (125009, Российская Федерация, г. Москва, Б. Кисловский переулок, д. 1); *сфера научных интересов*: когнитивная семантика, непрямое наименование, мультимодальная лингвистика; *e-mail*: maria_kiose@mail.ru
ORCID: 0000-0001-7215-0604. Scopus Author ID: 56642747500; ResearcherID: AAB-7989-2019; SPIN-код: 4419-0090, AuthorID: 341964.
Женишкуль Саматовна Хулхачиева, кандидат филологических наук, доцент, заведующий кафедрой языков и культур стран СНГ и ближнего зарубежья ИМОиСПН, директор центра киргизского языка и культуры им. Ч. Айтматова, Московский государственный лингвистический университет (119034, Российская Федерация, г. Москва, ул. Остоженка, д. 38 с. 1); *сфера научных интересов*: лингводидактика, методика преподавания, когнитивная лингвистика; *e-mail*: j.khulkhachieva@linguanet.ru
ORCID: 0009-0000-3506-2007. SPIN-код: 7897-5998, AuthorID: 1014042.
Бармин Артем Вячеславович, младший научный сотрудник Лаборатории когнитивных исследований основ коммуникации, Московский государственный лингвистический университет (119034, Российская Федерация, г. Москва, ул. Остоженка, д. 38 с. 1); *сфера научных интересов*: семантическая память, наука о сетях; *e-mail*: art.barmin1@gmail.com
ORCID: 0000-0001-9658-6621. Scopus Author ID: 58113747500; ResearcherID: AGL-9794-2022; SPIN-код: 4233-4593, AuthorID: 1146203.
Амирханов Аль-Берд Исламович, лаборант-исследователь Центра социокогнитивных исследований дискурса, Московский государственный лингвистический университет (119034, Российская Федерация, г. Москва, ул. Остоженка, д. 38 с. 1); *сфера научных интересов*: политическая экономия, региональное развитие, межкультурный трансфер знаний; *e-mail*: berd582@mail.ru

RUDN Journal of Language Studies, Semiotics and Semantics
2313-2299 (Print), ISSN 2411-1236 (Online)
2026 Vol. 17 No. 2 534–556
Вестник Российского университета дружбы народов.
Серия: ТЕОРИЯ ЯЗЫКА. СЕМИОТИКА. СЕМАНТИКА
http://journals.rudn.ru/semiotics-semantics

DOI: 10.22363/2313-2299-2026-17-2-534-456
EDN: LQDNOM

Научная статья / Research Article

Моделирование демографически обусловленной речи в больших языковых моделях: построение синтетического респондента для лингвистических исследований

А.С. Ваничкина ✉, В.В. Изюмская-Капитонова

Московский государственный лингвистический университет, *Москва, Российская Федерация*
✉alexvanichkina@yandex.ru

Аннотация. Использование больших языковых моделей (LLM) для построения синтетических респондентов активно осваивается в маркетинге и социальных науках, однако их применимость к моделированию демографически обусловленной вариативности естественной речи остается в лингвистике недостаточно исследованной. Целью статьи является оценка того, в какой мере синтетические респонденты, задаваемые с помощью личностного социодемографического промптинга, способны воспроизводить гендерно маркированные параметры русскоязычного дневникового письма. Материалом служат дневниковые записи Национального корпуса русского языка (НКРЯ) после 1950 г., разделенные на женский и мужской подкорпусы и использованные как референсные данные, а также 108 синтетических дневниковых текстов, сгенерированных тремя большими языковыми моделями (GPT-5.2, DeepSeek, GigaChat) по 32 конфигурациям личностного промптинга. Методы включают корпусный анализ гендерной вариативности речи на основе набора лексико-грамматических и прагматических показателей, автоматизированную морфологическую и лексико-прагматическую разметку синтетических текстов с помощью библиотеки Stanza, расчет параметрического и лексемного индексов гендерной близости и их агрегирование в интегральный показатель. Результаты показывают, что отдельные комбинации моделей и формул личностного промптинга (прежде всего GPT-5.2 и DeepSeek при непосредственном принятии роли и структурированной или именной стратегии прайминга) генерируют тексты, статистический профиль которых сближается с женским дневниковым подкорпусом НКРЯ, тогда как другие конфигурации систематически смещаются к мужскому профилю. Заключение. Показано, что синтетические респонденты могут выступать исследовательским инструментом для моделирования гендерно маркированного речевого поведения, но их валидность строго зависит от выбора модели и личностного промпта и требует предварительной калибровки относительно референсных корпусов; тем самым статья уточняет границы и условия ответственного использования синтетических респондентов в лингвистике.

Ключевые слова: синтетические респонденты, большие языковые модели, личностный социодемографический промптинг, гендерная вариативность речи, корпусная лингвистика, генерация синтетических текстов

© Ваничкина А.С., Изюмская-Капитонова В.В., 2026
This work is licensed under a Creative Commons Attribution-NonCommercial 4.0 International License
https://creativecommons.org/licenses/by-nc/4.0/legalcode

Вклад авторов: вклад авторов равнозначен на всех этапах исследования и подготовки текста статьи.

Конфликт интересов: авторы заявили об отсутствии конфликтов интересов.

История статьи: дата поступления: 22.02.2026; дата приема в печать: 03.03.2026.

Для цитирования: *Ваничкина А.С., Изюмская-Капитонова В.В.* Моделирование демографически обусловленной речи в больших языковых моделях: построение синтетического респондента для лингвистических исследований // Вестник Российского университета дружбы народов. Серия: Теория языка. Семиотика. Семантика. 2026. Т. 17. № 2. С. 534–456. https://doi.org/10.22363/2313-2299-2026-17-2-534-456 EDN LQDNOM

Modeling Demographically Conditioned Language in Large Language Models: Towards a Synthetic Respondent for Linguistic Research

Alexandra S. Vanichkina ✉, Veronika V. Izyumskaya-Kapitonova

Moscow State Linguistic University, *Moscow, Russian Federation*
✉alexvanichkina@yandex.ru

Abstract. Using large language models (LLMs) to construct synthetic respondents is increasingly common in marketing and the social sciences, yet their applicability to modeling demographically conditioned variation in natural language remains underexplored in linguistics. This article aims to assess the extent to which synthetic respondents, instantiated via sociodemographic persona prompting, can reproduce gender-marked features of Russian diary writing. The data comprise diary entries from the Russian National Corpus (RNC) after 1950, split into female and male subcorpora and used as reference data, as well as 108 synthetic diary texts generated by three large language models (GPT-5.2, DeepSeek, GigaChat) across 32 persona-prompt configurations. The methods include corpus-based analysis of gender variation in language using a set of lexico-grammatical and pragmatic features, automated morphological and lexico-pragmatic annotation of synthetic texts with the Stanza library, computation of a parametric and a lexical index of gender proximity, and their aggregation into an integral measure. The results show that specific combinations of models and persona-prompt formulas (above all GPT-5.2 and DeepSeek with direct role adoption and structured or name-based demographic priming) generate texts whose statistical profiles approximate the female diary subcorpus of the RNC, whereas other configurations systematically shift towards the male profile. The study concludes that synthetic respondents can serve as a research instrument for modeling gender-marked language use, but their validity is tightly dependent on the choice of model and persona prompt and requires prior calibration against reference corpora, thereby specifying the scope and conditions for the responsible use of synthetic respondents in linguistics.
Keywords: artificial intelligence, semiotics, semantics, simulation of meaning, perplexity

Authors' contribution: The authors contributed equally to this article and preparation of the article text.

Conflict of interests: The authors stated that there were no conflicts of interest.

Article history: Received: 22.02.2026; accepted: 03.03.2026.

For citation: Vanichkina, A.S., & Izyumskaya-Kapitonova, V.V. (2026). Modeling Demographically Conditioned Language in Large Language Models: Towards a Synthetic Respondent for Linguistic Research. *RUDN Journal of Language Studies, Semiotics and Semantics, 17*(2), 534–456. (In Russ.). https://doi.org/10.22363/2313-2299-2026-17-2-534-456 EDN LQDNOM

Введение

В последние годы в научном дискурсе активно формируется понятийный аппарат, связанный с использованием больших языковых моделей (Large language model, LLM) для имитации поведения реальных информантов в исследовательских целях. В англоязычных академических публикациях, а также в практико-ориентированных материалах компаний, разрабатывающих подобные решения на коммерческой основе, для обозначения генерируемых искусственным интеллектом (ИИ) участников исследования используются термины *synthetic respondent* («синтетический респондент») и *synthetic participant* («синтетический участник»), определяемые как искусственные цифровые личности, сгенерированные моделями машинного обучения с целью воспроизведения ответов реальных людей на основе демографических, психологических и поведенческих профилей [1][1]. В одной из ключевых для данной научно-практической области работ «Out of One, Many: Using Language Models to Simulate Human Samples» также вводится понятие *silicon samples* («кремниевые выборки»), обозначающее выборки синтетических ответов, которые были сгенерированы языковой моделью и предварительно настроены на тысячах социодемографических профилей реальных участников нескольких крупных опросов, проведенных в США. За счет такой параметризации модель воспроизводила распределения ответов различных социальных групп с высокой степенью соответствия эмпирическим данным и отражала не только поверхностное сходство, но и многомерное переплетение идей и социокультурных контекстов, лежащее в основе человеческих установок [2]. В русскоязычных источниках параллельно закрепляются термины *синтетический респондент*[2]*, цифровой респондент*[3] *и виртуальный респондент* [3]*,* первый из которых далее будет использоваться в данной работе по причине большего соответствия международной практике. Важно отметить, что в русскоязычной научной литературе использование LLM в качестве синтетических респондентов пока освещено крайне ограниченно, тогда как на прикладные популярные материалы в прессе, блогах и на корпоративных страницах разработчиков приходится основной массив публикаций по этой теме.

[1] РБК Компании. Режим доступа: https://nielseniq.com/global/en/insights/education/2024/the-rise-of-synthetic-respondents/ (дата обращения: 03.02.2026).

[2] РБК Компании. Режим доступа: https://companies.rbc.ru/news/8g5GYHRPqc/kak-byistro-sozdat-tsifrovuyu-lichnost-dlya-zadach-marketinga/ (дата обращения: 03.02.2026).

[3] РБК Компании. Режим доступа: https://www.sostav.ru/blogs/284997/70174 (дата обращения: 03.02.2026).

Основные сферы применения синтетических респондентов на сегодняшний день сосредоточены преимущественно в маркетинговых исследованиях, социологии и поведенческих науках, что подтверждается появлением отраслевых решений вроде проекта *Synth* компании «Родная Речь», использующего синтетических респондентов (разработчики называют их «*синтеты*») для быстрой оценки реакции аудитории и предварительного тестирования гипотез в сфере маркетинга и медиапотребления[4]. В то же время исследование [4], представленное на конференции *International Conference on Machine Learning* (ICML) — одной из ведущих площадок в области машинного обучения, — показало, что LLM способны воспроизводить результаты не только опросов, но и классических экономических, психолингвистических и социо-психологических экспериментов.

Вместе с тем применимость синтетических респондентов в лингвистических исследованиях, в частности при изучении речевого поведения конкретных социальных групп, остается существенно менее разработанной. При этом именно в тех областях, где доступ к эмпирическому языковому материалу ограничен вследствие физической недоступности данных, жестких этических регуляций и строгих режимов конфиденциальности, потребность в подобных исследованиях оказывается особенно значимой. К числу таких типов дискурса относятся личная переписка, дневниковые записи и частные заметки, для которых наблюдается выраженный дефицит репрезентативных корпусов, что существенно затрудняет систематическое изучение вариативности идиолектов, социолектов и прагматических стратегий. Синтетические респонденты потенциально могут помочь преодолеть эти барьеры, позволяя моделировать речевые паттерны на основе агрегированных демографических данных без риска нарушения приватности, что особенно актуально для труднодоступных социальных групп, таких как носители редких диалектов или маргинализированные сообщества.

В настоящей работе синтетические респонденты рассматриваются в фокусе гендерной вариативности русскоязычного дневникового письма и сопоставляются с реальными текстами из Национального корпуса русского языка (НКРЯ), где дневниковые записи после 1950 г. снабжены метаданными об авторском поле и позволяют выделить референсные женский и мужской подкорпусы. На основе этих данных конструируется система лексико-грамматических и прагматических показателей (части речи, грамматические категории, экспрессивность, диминутивы, средства вежливости, обсценная лексика, вопросы, восклицания, лексические повторы), задающих эмпирический профиль гендерной вариативности дневниковых текстов. Далее с помощью формул личностного социодемографического промптинга генерируется корпус синтетических дневниковых записей тремя крупными языковыми моде-

[4] Новости рекламы и маркетинга. Режим доступа: https://adindex.ru/news/innovations/2025/07/10/335234.phtml (дата обращения: 1.12.2025).

лями (GPT-5.2, DeepSeek, GigaChat); все тексты проходят автоматизированную морфологическую и прагматическую разметку с помощью Python, после чего для каждого из них вычисляются интегральные параметры сходства с женским и мужским эталонными профилями. Это позволяет оценить, в какой мере выбор модели и конфигурации личностного промпта влияет на способность синтетических респондентов воспроизводить гендерно маркированные особенности русскоязычного дневникового письма.

НКРЯ как источник референсных данных для оценки синтетических респондентов

Для оценки эффективности LLM в задаче моделирования синтетических респондентов в лингвистических исследованиях исходным требованием является возможность адекватного сопоставления результатов больших языковых моделей с реальными языковыми данными. В качестве такой референсной эмпирической базы в настоящем исследовании используется Национальный корпус русского языка (НКРЯ), обеспечивающий достаточный объем материала для надежного статистического анализа.

Именно особенностями данной платформы — в частности, структурой и полнотой метаданных об авторах текстов, — обуславливался выбор демографических переменных для тестирования синтетических респондентов. Таким образом, в качестве ключевого параметра определения социальной группы была выбрана *гендерная принадлежность*, поскольку соответствующий признак наиболее последовательно и полно аннотирован в референсном источнике. Возрастные характеристики, напротив, не могут стать предметом анализа из-за заметной несбалансированности корпуса по годам рождения авторов, а также вследствие отсутствия во многих случаях сведений о возрасте автора на момент написания текста в экстралингвистической разметке. Другие демографические параметры, такие как национальность, уровень образования и др., в НКРЯ в явном виде не представлены.

Среди доступных в НКРЯ типов текстов, релевантных для поставленных задач, были рассмотрены подкорпусы личных писем («письмо личное»: 3112 текстов, 4 402 739 слов) и дневниковых записей («дневники, записные книжки»: 3677 текстов, 38 735 953 слова), причем последний превосходит первый по объему более чем в десять раз, что делает его предпочтительным источником данных. Выбор дневниковых записей дополнительно мотивирован их жанровой спецификой — монологической спонтанностью, отсутствием явного адресата и высокой степенью индивидуализации языкового выражения [5]. В письмах речевое поведение в значительной степени регулируется ожиданиями конкретного адресата и нормами вежливости, тогда как дневниковый текст фиксирует преимущественно автокоммуникативные практики, менее подверженные внешнему коммуникативному контролю.

Ограничение корпусного материала записями после 1950 г. обусловлено необходимостью обеспечить соответствие референсных данных материалу,

на котором обучены современные большие языковые модели, и тем самым повысить корректность последующего сравнения. Тексты этого периода не содержат системных архаизмов и орфографических колебаний, характерных для первой половины XX века, что снижает риск расхождений в распределениях форм и улучшает качество автоматической морфосинтаксической разметки. Такой хронологический срез также позволяет интерпретировать наблюдаемые различия между реальными и синтетическими текстами как отражение особенностей генерации, а не как эффект историко-языковых сдвигов.

В совокупности эти решения задают референсную базу исследования: дневниковые записи НКРЯ, разделенные на женский и мужской подкорпуса и ограниченные периодом после 1950 г.

Сбор и анализ референсных данных

Поскольку в качестве основной переменной для дальнейшего моделирования синтетических респондентов рассматривается гендерная принадлежность автора, следующим шагом стало определение параметров, репрезентирующих типичную гендерную вариативность речи. Опираясь на представление о гендерлекте как социально и культурно обусловленном варианте речевого поведения мужчин и женщин [6], а также на наблюдения о дифференциации мужской и женской речи на различных уровнях языка, в качестве отправной точки был выбран широкий спектр описанных в литературе признаков [7–10] — от распределения частей речи и лексико-грамматических предпочтений до частотности параметров экспрессивности, средств вежливости и модальности, а также прагматических маркеров. Однако прямое перенесение всего этого множества индикаторов в корпусную парадигму оказалось невозможным: во-первых, значительная часть описанных различий относится к устной, диалогической, ситуативно привязанной коммуникации, тогда как настоящая работа базируется на письменных дневниковых текстах; во-вторых, многие параметры (типичные темы, особенности прагматикализации, семантические и оценочные сдвиги) либо не поддаются надежной автоматической разметке, либо требуют трудоемкой ручной аннотации, несовместимой с масштабом исследуемого материала.

С учетом этих ограничений перечень признаков гендерлекта был ограничен теми параметрами, которые, с одной стороны, имеют теоретическое обоснование в работах по гендерной вариативности речи, с другой стороны, могут быть обработаны существующими инструментами автоматического и автоматизированного анализа. Таким образом, последующие сбор и анализ данных были реализованы по следующим единицам: частям речи (существительные, глаголы, прилагательные, наречия, числительные, предлоги, союзы, местоимения, частицы, междометия) и леммам; некоторым грамматическим категориям (временные формы, императив, условное наклонение, сравнительная и превосходная степени); параметрам прагматической организации

высказывания (вводные конструкции, вопросительные и восклицательные предложения); а также лексическим и социопрагматическим маркерам, включающим диминутивы, нецензурную лексику, лексемы-индикаторы вежливости, имена собственные и лексические повторы. В дальнейшем этот перечень трактуется как параметры гендерной вариативности речи.

Таким образом, анализ, осуществленный на предыдущем этапе, позволил не только сузить набор признаков гендерлекта до перечня соответствующих целям и ограничениям исследования параметров, но и применить их к установленному ранее подкорпусу НКРЯ, разделенному на женские и мужские подвыборки.

На этой основе были выделены два референсных подкорпуса на базе НКРЯ: дневниковые записи, написанные женщинами (418 текстов, 2 508 844 словоформы), и дневниковые записи, написанные мужчинами (1493 текста, 13 936 392 словоформы). В каждом из этих подкорпусов был проведен лексико-грамматический поиск по всем параметрам гендерлекта, определенным ранее. В большинстве случаев процедура опиралась на грамматические признаки (части речи и грамматические категории), однако ряд параметров требовал привлечения дополнительных указателей, в частности при идентификации лексических повторов.

Специальные решения потребовались и для подсчета вопросительных и восклицательных предложений, поскольку интерфейс НКРЯ не предусматривает прямого поиска по знакам препинания. Вместо этого использовались пометы категории «слово перед» с подкатегориями «перед восклицательным знаком» и «перед вопросительным знаком», что позволило получить статистические данные по распределению восклицательных и вопросительных предложений в каждом из подкорпусов. Дополнительные трудности возникли при поиске лексических средств выражения вежливости и нецензурной лексики, так как соответствующие пометы в НКРЯ в явном виде отсутствуют и каждый из этих параметров требовал предварительного составления специального списка лексем. Вслед за наблюдениями о гендерно дифференцированной вариативности редуцированных форм в повседневной устной коммуникации [10], при поиске средств выражения вежливости дополнительно учитывались их графически редуцированные варианты, которые, как показал предварительный анализ Основного корпуса НКРЯ, представлены и в письменных текстах (например: «спасиб», «спс», «пожалста», «плз»). Поиск нецензурной лексики осуществлялся на основе соответствующей словарной статьи в Толковом словаре государственного языка Российской Федерации, входящем в перечень нормативных словарей ИРЯ РАН, причем дополнительной сложностью являлось то, что словарная статья фиксирует набор корней, а не лемм, вследствие чего список лексем для корпусного поиска по леммам и словоформам формировался вручную.

Результаты поиска по параметрам сохранялись в двух форматах. Во-первых, формировались две сводные таблицы (по женским и мужским

дневниковым записям) по всем параметрам, в которых для каждого параметра указывались общее количество текстов в подкорпусе, общий объем подкорпуса в словоформах, число найденных вхождений, удовлетворяющих поисковому запросу, и нормализованная частотность на миллион словоформ (instances per million words, IPM). Во-вторых, для каждого отдельного поискового запроса из НКРЯ выгружалась частотная таблица лемм, содержащая сведения о доле каждой лексемы в рамках данного запроса, ее нормализованной частотности и абсолютном количестве вхождений.

Далее был реализован макроуровневый анализ, рассматривающий каждый из параметров как целостный маркер, для которого сопоставлялись относительные частоты в мужском и женском подкорпусах. Для каждого параметра X из сводных таблиц извлекались: объем женского подкорпуса в словоформах N_f, объем мужского подкорпуса N_m, число вхождений параметра X в женском подкорпусе $hits_f$ и число вхождений параметра X в мужском подкорпусе $hits_m$. На основе этих величин сравнивались частоты токенов параметра X и всех остальных токенов ($\neg X$) в обоих подкорпусах.

Каждый параметр, таким образом, моделировался как бинарная категория (наличие/отсутствие маркера в токенах подкорпуса), что позволяет использовать G-тест (log-likelihood ratio) для проверки нулевой гипотезы об одинаковых относительных частотах X в мужском и женском подкорпусах. При этом суммарный объем корпуса обозначался как $N = N_f + N_m$, а общая частота параметра $hits = hits_f + hits_m$. Ожидаемые частоты при верной нулевой гипотезе вычислялись из маргинальных сумм: ожидаемое число вхождений X и $\neg X$ в каждом подкорпусе определялось как

$$E_{fX} = \frac{N_f \times hits}{N},\ E_{mX} = \frac{N_m \times hits}{N},\ E_{fX} = \frac{N_f \times (N - hits)}{N},\ E_{mX} = \frac{N_m \times (N - hits)}{N}.$$

Статистика G^2 вычислялась по стандартной формуле отношения правдоподобий для таблицы сопряженности:

$$G^2 = 2\left(hits_f \ln\frac{hits_f}{E_{fX}} + (N_f - hits_f)\ln\frac{N_f - hits_f}{E_{f-X}} + hits_m \ln\frac{hits_m}{E_{mX}} + (N_m - hits_m)\ln\frac{N_m - hits_m}{E_{m-X}}\right).$$

Большие значения G^2 соответствуют сильным отклонениям наблюдаемых частот от ожидаемых при нулевой гипотезе. Для каждого параметра полученное значение G^2 интерпретировалось в терминах распределения χ^2 с одной степенью свободы, что позволяло вычислить *p-значение* как $p = P(\chi^2_{(df=1)} \geq G^2_{набл})$, то есть вероятность получить такое же или более экстремальное значение G^2 при условии, что относительные частоты параметра X в мужском и женском подкорпусах на самом деле совпадают. Чем меньше p-значение, тем менее правдоподобно, что наблюдаемое расхождение частот возникло случайно при верной нулевой гипотезе, и тем более убедительными являются основания для ее отклонения.

Параллельно для каждого параметра сохранялись нормированные частоты в IPM для женского и мужского подкорпусов (IPM_f и IPM_m), их разность $\Delta IPM = IPM_f - IPM_m$ и направление эффекта: тег «*female*»/«*male*» присваивался в зависимости от знака ΔIPM, при этом положительное значение указывало на более высокую частоту параметра у женщин, отрицательное — у мужчин.

Полученная сводная таблица показала статистически значимые различия по целому ряду параметров: G^2 достигал значений в тысячах, а *p-значение* оценивались как 0.0 вследствие предельно малых величин в условиях больших объемов корпусов, что стандартно для корпусных *G*-тестов на больших выборках. При этом направление соответствует ожиданиям гендерной лингвистики: например, женщины демонстрируют относительно более высокие нормализованные частоты междометий, диминутивов, восклицательных предложений, лексических повторов и форм будущего времени; мужчины — более высокую относительную частоту существительных, прилагательных, числительных, предлогов и некоторых градационных форм (превосходная степень).

Таким образом, на макроуровне было показано, что отобранные параметры систематически распределены по полу, и эти различия не могут быть объяснены случайными флуктуациями.

На следующем этапе фокус исследования сместился с уровня обобщенной представленности параметров в подкорпусах на уровень распределения отдельных лексем внутри каждого из них: использовались частотные таблицы НКРЯ, полученные по результатам лексико-грамматических запросов и содержащие сведения о лемме, IPM и абсолютном числе вхождений. Их обработка осуществлялась с помощью кода на языке программирования *Python*, который последовательно обращался к каждому такому файлу, соответствующему комбинации «параметр–пол», и приводил данные к единому формату, что позволило выделить наиболее типичные лексемы в женских и мужских дневниковых текстах.

В дальнейшем лексемы внутри каждого файла упорядочивались по убыванию IPM; затем фиксировалось пороговое значение IPM на позиции N (в данном исследовании $N = 10$), то есть значение, приходящееся на условную десятую строку отсортированного списка. На основании этого порога формировалось множество высокочастотных единиц: в него включались все лексемы с IPM не ниже порогового значения. Если единицы на последующих позициях (после десятой) имели такое же значение IPM, они также относились к этому множеству, что исключало искусственное разделение статистически сопоставимых по частоте форм.

Для каждой лексемы, включенной в множество высокочастотных единиц, сохранялись сведения о параметре, поле, лемме, IPM и абсолютном числе вхождений; далее результаты по всем комбинациям «параметр–пол» конкатенировались в единый сводный файл. На основании этого массива для

каждого параметра были выделены три типа множеств: общие лексемы, входящие в верхнюю зону распределения по IPM и в женском, и в мужском подкорпусе; единицы, характерные только для женских дневников; и лексемы, присутствующие лишь в мужских текстах. Для общих элементов дополнительно сопоставлялись значения IPM и ранги в обоих подкорпусах, что позволило зафиксировать как пересечение наборов наиболее частотных лексем, так и различия в их положении в женском и мужском подкорпусах.

Личностный социодемографический промптинг при моделировании синтетических респондентов

В настоящем исследовании при создании синтетических респондентов используются формулы личностного социодемографического промптинга (sociodemographic persona prompting), сформулированные в работе «The Prompt Makes the Person(a): A Systematic Evaluation of Sociodemographic Persona Prompting for Large Language Models» [11]. Согласно авторам, адекватный задачам исследования запрос делится на два основных элемента: во-первых, личностный сегмент, который включает в себя *характер принятия роли* (role adoption format) и *стратегию демографического прайминга* (demographic priming strategy), во-вторых, сегмент задачи, которую синтетический респондент должен выполнить.

Выделяются следующие типы характера принятия роли: *непосредственный* (Direct), где модель напрямую позиционируется как носитель заданной личности — «You are a Hispanic woman» / «Ты женщина латиноамериканского происхождения»; *от третьего лица* (Third Person), где личность вводится опосредованно как объект воображаемой перспективы — «Think of a Hispanic woman» / «Представь женщину латиноамериканского происхождения»; *интервью* (Interview), где демографические характеристики поочередно задаются в вопросно-ответном виде — «Interviewer: What gender do you identify as? Interviewee: I identify as female» / «Интервьюер: Какого вы пола? Интервьюируемый: Я идентифицирую себя как женщину»).

Стратегии социодемографического прайминга также подразделяются на несколько типов: *эксплицитная* (Explicit), использующая явные вербальные описания — «a Hispanic woman» / «Женщина латиноамериканского происхождения»; *именная* (Name), при которой демографические признаки задаются имплицитно через имя — «Ms. Hernandez» / «г-жа Эрнандес»; *структурированная* (Structured), в которой характеристики задаются в виде перечня категориальных атрибутов — «a person of race „Hispanic“ and gender „female“» / «лицо происхождения „латиноамериканский“ и пола „женский“».

Для проведения экспериментов авторы определяют формулы личностного социодемографического промптинга как декартово произведение трех типов характера принятия роли и трех стратегий демографического прайминга, в результате чего выделяются девять типов личностных запросов для создания синтетических респондентов. Для каждой из этих девяти

комбинаций используются две альтернативные формулировки, сохраняющие общую структуру инструкций.

В настоящем исследовании описанные выше формулы личностного социодемографического промптинга используются в адаптированном виде, с учетом ограничений референсного корпуса и задач сопоставления с дневниковыми текстами НКРЯ. В экстралингвистической разметке корпуса пол является единственной стабильно доступной социодемографической характеристикой автора, поэтому в личностном сегменте он указывается явно; при этом в качестве целевой группы для моделирования синтетического респондента был произвольно выбран женский гендерлект. Другие параметры (национальность, гражданство и т.п.) в разметке отсутствуют, поэтому характеристики типа «русская», «россиянка» и аналогичные им не включались в личностный сегмент, и синтетический респондент специфицируется через русскоязычность.

В рамках именной стратегии социодемографического прайминга требовалось задать устойчивый антропонимический шаблон. По данным основного корпуса НКРЯ, были определены наиболее частотные женское имя, отчество и фамилия: «Анна», «Ивановна» и «Иванова». Поскольку возраст автора в корпусной разметке также не задан, временная характеристика синтетического респондента задается через условие «жизнь после 1950 г.», что обеспечивает сопоставимость генерируемых текстов с дневниковым подкорпусом этого периода.

Сегмент задачи включал указание жанрово-стилистической характеристики целевого текста как «нехудожественной записи в личный дневник», что обеспечивает жанровое соответствие синтетических текстов референсным данным. Дополнительно задавалось требование подробности записи (для получения достаточного объема материала при отсутствии жесткого числового ограничения на длину текста) и уточнение «как в реальной записной книжке», призванное приблизить высказывание к повседневной письменной практике.

Вслед за опорной работой для каждой комбинации характера принятия роли и стратегии прайминга использовались две альтернативные версии. Однако дополнительно варьировалась и форма обращения к модели (на «ты» и на «вы») с учетом противоречивых данных о влиянии вежливости запроса на качество ответа: в ряде исследований для английских моделей описывается наивысшая результативность в зоне нейтрально-вежливых формулировок [12], тогда как более поздние работы демонстрируют преимущество грубо сформулированных инструкций по сравнению с вежливыми запросами [13]. В результате общее число тестируемых вариантов промптов было увеличено с 18 до 32 за счет введения переменной формы обращения.

Для каждого из описанных выше промптов был сформирован набор ответов от нескольких доступных больших языковых моделей. В рамках настоящего исследования сознательно не использовались специализированные коммерческие решения, обученные на текстах конкретных социальных

групп, поскольку в типичной исследовательской практике на начальном этапе эксперимента такие дорогостоящие инструменты, как правило, недоступны. Вместо этого были отобраны три широко используемые модели: GPT-5.2, DeepSeek и GigaChat, что позволило сопоставить результаты систем, разработанных в разных национальных и технологических средах.

При запуске запросов соблюдался единый протокол использования моделей. Для каждого промпта открывался отдельный диалог, чтобы избежать переноса контекста между запросами, а последовательность подачи промптов в разные модели рандомизировалась, что минимизировало влияние порядка предъявления инструкций на характеристики генерируемых текстов.

Таким образом, использование всех формул личностного социодемографического промптинга — сочетаний трех типов характера принятия роли и трех стратегий социодемографического прайминга, заданных в двух альтернативных формулировках и в двух вариантах обращения к модели (на «ты» и на «вы»), для генерации текстов тремя выбранными большими языковыми моделями дало 108 синтетических текстов для последующего сопоставления с полученными ранее референсными данными. Каждый из этих текстов сохранялся в отдельном файле, который включал служебный заголовочный блок с техническими параметрами промпта и идентификаторами модели, а также основной блок с собственно дневниковым текстом.

Автоматизированная разметка и выделение лексико-прагматических показателей в сгенерированных текстах

Тексты, сгенерированные моделями в ответ на экспериментальные промпты, обрабатывались с применением модулей *Python*, предназначенных для работы с файловыми структурами, регулярными выражениями и табличными данными. В качестве основного инструмента морфологической обработки использовалась библиотека *Stanza*, соответствующая стандартам *Universal Dependencies* и, согласно опубликованным бенчмаркам, демонстрирующая более высокие показатели по сравнению с альтернативными решениями.

На этапе подготовки данных из каждого файла извлекалась только содержательная часть дневникового текста, тогда как форматирование и технические метаданные сохранялись отдельно для последующего качественного анализа. Далее проводилась полная морфологическая разметка: материал сегментировался на предложения и токены, для каждой словоформы определялись лемма, часть речи и совокупность морфологических признаков. На основе полученных данных вычислялись базовые структурные показатели (количество предложений и токенов), а также частотные распределения частей речи и грамматических категорий.

Помимо базовой морфологической разметки, для ряда лексико-прагматических маркеров потребовались специализированные процедуры, не реализуемые в существующих инструментах автоматического анализа. К таким

маркерам относились вводные слова, средства выражения вежливости, диминутивы, обсценная лексика и лексические повторы. Вводные конструкции идентифицировались в два этапа: вначале на основе дневниковых текстов НКРЯ формировался совокупный перечень лемм вводных слов; далее в сгенерированных текстах выделялись токены, совпадающие по лемме с этим списком и одновременно обособленные запятыми, что обеспечивало приближение к их типичным синтаксическим позициям. Маркеры вежливости определялись по перечню целевых лемм и набору усеченных основ, отражающих редуцированные варианты. Поиск диминутивов осуществлялся с использованием специально сформированного словаря, поскольку стандартные морфологические средства не обеспечивают надежного распознавания подобных форм в русскоязычных текстах. На первом этапе во всем корпусе выделялись слова, оканчивающиеся на последовательности символов, соответствующие типичным уменьшительно-ласкательным суффиксам; затем полученный список проходил ручную проверку, в ходе которой исключались единицы, формально подпадающие под шаблон, но не являющиеся диминутивами (например, «понедельник»). На основе предобработанного набора лемм и соответствующих словоформ создавался итоговый словарь, по которому оценивалась представленность диминутивов в каждом тексте. Обсценная лексика выявлялась по отдельному списку целевых лемм, адаптированному к особенностям лемматизации русской модели *Stanza*. Лексические повторы трактовались как повторное употребление одной и той же леммы в пределах интервала длиной не более двух токенов. Дополнительно рассчитывалась частота восклицательных и вопросительных предложений на основе соответствующих знаков препинания.

Все вычисленные параметры — от структурных и морфологических до прагматических и стилистических — были агрегированы в единую табличную структуру, где каждая строка соответствовала отдельному сгенерированному дневниковому тексту. Для каждого показателя гендерной маркированности в таблице фиксировались абсолютная частота, значение IPM, рассчитываемое на объеме конкретного текста, и словарь частотности соответствующих лексем в формате «лемма: абсолютная частота, IPM», что обеспечивало возможность последующего статистического сопоставления с референсными данными дневниковых подкорпусов НКРЯ.

Сопоставление данных синтетических респондентов и референсных подкорпусов НКРЯ

Для формализованной оценки степени соответствия дневниковых текстов, сгенерированных большими языковыми моделями, речевому поведению целевой группы — женщин-авторов дневниковых записей, живших после 1950 г., — был разработан многоуровневый метод сопоставления, основанный на вышеописанных параметрах и учитывающий как их распределение, так и частотные характеристики связанных лексем.

На подготовительном этапе в рабочую среду были загружены сводные таблицы по женскому и мужскому подкорпусам НКРЯ, частотные таблицы лемм по каждому показателю и агрегированные данные морфологической и прагматической разметки всех сгенерированных текстов. Для анализа и сопоставления использовались библиотеки *pandas* (для обработки табличных данных), *NumPy* и *SciPy* (для проведения статистических вычислений), а также модуль *ast* (для разбора словарных структур). Реализация сопоставления посредством *Python* обеспечивала автоматизированный расчет комплексных показателей близости между параметрами сгенерированного массива и референсных подкорпусов.

На первом этапе эталонные данные по женскому и мужскому дневниковым подкорпусам были преобразованы в компактный формат, удобный для сравнительного анализа. Из сводных таблиц по НКРЯ автоматически извлекались значения IPM по всем ранее описанным параметрам. В среде разработки на основе этих данных формировались два вектора — женский и мужской. Для каждого подкорпуса вычислялась сумма IPM по всем параметрам, после чего частотные значения каждого параметра нормировались по этой сумме. В результате были получены два нормированных профиля — «женский» и «мужской», в которых значения параметров выражены не в абсолютных IPM, а как доли от совокупного объема маркеров, что позволяет оценивать относительную представленность каждого параметра в обоих подкорпусах.

Аналогичная процедура нормировки применялась к каждому синтетическому дневниковому тексту: для каждой строки агрегированной таблицы извлекались значения IPM по всем параметрам, соответствующим референсным данным, и на их основе формировался нормированный профиль распределения параметров конкретного сгенерированного текста.

Поскольку IPM в референсных подкорпусах рассчитываются на всем объеме корпуса, а в массиве сгенерированных текстов — на длину отдельного текста, прямое сопоставление абсолютных IPM было бы некорректным. Нормировка до распределений долей позволяет сравнивать не абсолютные значения, а относительное распределение параметров, то есть их взаимное усиление или ослабление при устранении влияния общего уровня частот. На этой основе для каждого сгенерированного текста в среде *Python* вычислялась степень сходства его профиля с женским и мужским эталонными профилями. В качестве меры использовалось косинусное сходство между нормированными векторами распределения параметров, реализованное с помощью функций библиотеки *SciPy*. Этот показатель, широко применяемый при сравнении векторов частотных признаков и других текстовых представлений, оценивает угол между векторами и не зависит от их длины, что делает его удобным для сопоставления распределений частотных параметров.

В результате для каждого текста были получены два значения: степень сходства с женским эталонным профилем и степень сходства с мужским. На их основе для каждого текста рассчитывался параметрический индекс,

определяемый как разность «женского» и «мужского» сходств. Положительное значение индекса означало, что профиль распределения маркеров в тексте ближе к женскому дневниковому подкорпусу, отрицательное — что он ближе к мужскому. Абсолютная величина индекса отражала степень выраженности этого смещения; при этом необходимо учитывать, что сами женский и мужской профили дневников по всем параметрам в целом близки, поэтому параметрический индекс фиксирует, скорее, тонкие, но систематические сдвиги между конфигурациями модели и промпта.

На уровне лексем сопоставление опиралось на множества высокочастотных единиц, выделенных ранее для каждого параметра в женском и мужском дневниковых подкорпусах НКРЯ (общие, характерные только для женских дневников и только для мужских). Для каждого параметра в сгенерированных текстах подсчитывалось количество вхождений лемм из «женских» и «мужских» списков, что позволило оценивать, в какой степени лексическая реализация данного маркера тяготеет к женскому или мужскому референсному профилю.

В агрегированной таблице сгенерированных текстов для каждого параметра были предусмотрены столбцы, содержащие словари вида «лемма: абсолютная частота, IPM» для соответствующей группы маркеров. На данном этапе эти словари последовательно разбирались для каждого текста: из них извлекались IPM всех лемм, входящих в «женский» словарь по данному параметру, и IPM всех лемм, входящих в «мужской» словарь. Далее для каждого текста и по всем параметрам суммировалась нормализованная частотность «женских» лемм и отдельно — «мужских». Разность этих суммарных значений задавала лексемный индекс: если суммарная IPM «женских» лемм превышала суммарную IPM «мужских», индекс принимал положительное значение и интерпретировался как преобладание лексики, характерной для женского дневникового письма, и наоборот.

Поскольку параметрический и лексемный индексы имеют разную природу и масштаб значений, оба показателя были дополнительно приведены к сопоставимой шкале. Для этого по всему массиву из 108 сгенерированных текстов вычислялись средние значения каждого индекса и их стандартные отклонения, после чего для каждого текста рассчитывалась нормированная оценка (z-оценка), показывающая, насколько значение данного индекса отклоняется от среднего по корпусу в единицах стандартного отклонения. Интегральный показатель гендерной близости к женскому дневниковому профилю определялся как среднее двух нормированных индексов. Положительные значения этого показателя указывали на то, что текст в совокупности — и по распределению параметров, и по лексическим предпочтениям — ближе к профилю женских дневников, отрицательные — на большее сходство с мужским профилем; величина по модулю отражала степень отклонения от среднего поведения моделей.

На заключительном этапе все полученные индексы для каждого текста агрегировались по комбинациям факторов: большой языковой модели,

характера принятия, стратегии социодемографического прайминга и формы обращения. В среде Python для каждой такой комбинации вычислялись число текстов в ячейке, среднее значение интегрального показателя, средние значения параметрического и лексемного индексов, а также доля текстов с положительным интегральным значением. Это позволило уже на уровне сводных таблиц выявить, какие сочетания модели и формулы личностного социодемографического промптинга систематически ведут к генерации текстов, максимально приближающихся к референсному профилю женских дневников, а какие конфигурации приводят к смещению в сторону мужского стиля.

Результаты

Интегральный показатель, а также параметрический и лексемный индексы, вычисленные для всех 108 синтетических дневниковых текстов, продемонстрировали устойчивую зависимость гендерной маркированности синтетических ответов от выбора модели и конфигурации личностного социодемографического промптинга. В целом массив сгенерированных текстов DeepSeek и GPT-5.2 чаще характеризовался положительными значениями интегрального показателя, то есть большей близостью к референсному профилю женских дневников, тогда как тексты GigaChat преимущественно смещались в сторону мужского профиля. Для DeepSeek наиболее высокие значения интегрального показателя наблюдались в условиях *непосредственного* характера принятия роли в сочетании со *структурированной* или *именной* стратегией социодемографического прайминга, а также при формате *интервью* с *именной* стратегией и обращением на «ты». Для GPT-5.2 максимальные значения интегрального показателя приходились на сочетания *непосредственного* характера принятия роли со *структурированной* стратегией и обращением на «ты», а также на ряд конфигураций с *эксплицитной* стратегией при том же характере принятия роли. Напротив, для GigaChat большинство комбинаций, особенно при *именной* и *структурированной* стратегиях, давали отрицательные значения интегрального показателя, что свидетельствует о преобладании гендерно маркированных особенностей, характерных для мужских дневниковых текстов.

Анализ влияния характера принятия роли и стратегии социодемографического прайминга показал, что эти факторы взаимодействуют с моделью и существенно меняют направление гендерного сдвига. При *непосредственном* принятии роли DeepSeek и GPT-5.2 демонстрировали наиболее устойчивое приближение к женскому дневниковому профилю вне зависимости от стратегии прайминга, тогда как GigaChat в тех же условиях сохранял отрицательные или близкие к нулю значения интегрального показателя. Формат *интервью* оказался наиболее чувствительным к выбору стратегии: при *именной* стратегии в этом формате GPT-5.2 и GigaChat давали одни из самых выраженных отрицательных значений интегрального показателя, тогда как DeepSeek при той же комбинации, но при обращении на «ты», формировал

тексты с заметным положительным сдвигом. В режиме принятия роли *«от третьего лица» структурированная* стратегия для DeepSeek и GPT-5.2 в большинстве случаев обеспечивала положительные значения интегрального показателя, тогда как *именная* стратегия в этом формате последовательно ассоциировалась с мужским профилем во всех трех моделях.

Форма обращения к модели также систематически модифицировала гендерную маркированность сгенерированных текстов. Для DeepSeek и GPT-5.2 обращение на «ты» в ряде конфигураций с *непосредственным* характером принятия роли и *структурированной* или *именной* стратегией приводило к более высоким значениям интегрального показателя по сравнению с обращением на «вы», усиливая приближение к женскому дневниковому профилю. Для GigaChat, напротив, обращение на «вы» в некоторых случаях частично компенсировало мужской сдвиг (например, при *эксплицитной* стратегии в формате *интервью* или при принятии роли «*от третьего лица*»), тогда как обращение на «ты» в аналогичных условиях сопровождалось наиболее низкими значениями интегрального показателя. Сопоставление параметрического и лексемного индексов показало, что знак интегрального показателя в большинстве случаев согласован и с распределением параметров, и с распределением лексем: конфигурации, приближающиеся к женскому дневниковому подкорпусу по параметрическому индексу, одновременно демонстрировали преобладание лемм из «женских» словарей, в то время как отрицательные интегральные значения сочетались с доминированием «мужских» лексем.

Для последующего качественного анализа были выделены три конфигурации с наибольшей близостью к женскому дневниковому профилю по интегральному показателю. Во-первых, GPT-5.2 при *непосредственном* характере принятия роли, *структурированной* стратегии социодемографического прайминга и обращении на «ты» демонстрирует максимальное положительное значение интегрального показателя среди всех комбинаций, что указывает на максимально выраженное совпадение и параметрического, и лексемного профилей с женским дневниковым подкорпусом. Во-вторых, DeepSeek при *непосредственном* характере принятия роли и *структурированной* стратегии с обращением на «вы» также формирует тексты с очень высоким положительным интегральным показателем, что делает эту конфигурацию одним из ведущих режимов генерации синтетических ответов, близких к женскому дневниковому дискурсу. В-третьих, DeepSeek в формате *интервью* с именной стратегией и обращением на «ты» обеспечивает сопоставимые по величине положительные значения интегрального показателя, что позволяет рассматривать эту комбинацию как ключевой пример успешного моделирования женского гендерлекта при более сложном сценарии принятия роли.

Для детализации и верификации количественных результатов был проведен качественный анализ шести синтетических дневниковых текстов (по два варианта для трех конфигураций личностного социодемографического промптинга, показавших наибольшую близость к женскому дневниковому

профилю по интегральному показателю). Анализ опирался на рамку автороведческой экспертизы и модель речевого жанра личного дневника [14], предложенную Т.В. Шмелевой.

На первом этапе, с опорой на методику Т.Ф. Моисеевой и И.В. Огорелкова [15], для каждого текста оценивалось, отвечает ли предполагаемый авторский профиль следующим критериям: женский пол, период после 1950 г. и воспроизведение письменной практики носителя русского языка; во всех рассмотренных случаях тексты, сгенерированные двумя лучшими по формальным показателям моделями, были признаны соответствующими этим требованиям. Дополнительно учитывалось, что тексты представляют собой достаточно протяженные свободные высказывания (не менее 100 словоформ) и могут рассматриваться как произведения отдельных «авторов», в которых в полной мере проявляется речемыслительный навык, что делает их пригодными для автороведческой экспертизы. Далее для каждого из шести отобранных текстов (двух текстов, сгенерированных GPT-5.2, и четырех текстов, сгенерированных DeepSeek) был проведен детальный описательный анализ: фиксировались правильность орфографии и пунктуации, степень сложности синтаксической организации (включая сложные союзы, причастные и деепричастные конструкции, вводные слова), а также ориентация на книжные формы при сохранении элементов разговорной речи.

В рамках автороведческого анализа внимание сосредотачивалось на том, насколько синтетические тексты воспроизводят профиль женщины-автора: учитывались последовательное использование форм женского рода при повествовании от первого лица и согласование зависимых форм, тематическая фокусировка на сферах «дом, семья, отношения», частые обращения к эмпатическим и оценочным стратегиям (жалость, сочувствие, самокритика, апелляция к собственной совести). Отдельно отмечалась высокая плотность вводных слов и модальных конструкций («кажется», «наверное», «по-моему», «если честно»), оформляющих колебания, неуверенность и саморефлексию и традиционно описываемых в гендерной литературе как характерный ресурс «женского письма»; в выбранных шести текстах совокупность этих признаков интерпретировалась экспертами как устойчивое приближение к письму женщины-автора.

На следующем этапе тексты оценивались с точки зрения соответствия жанровым признакам личного дневника в рамках модели Т.В. Шмелевой: анализировались коммуникативная цель, образ адресата, событийное содержание и темпоральная организация, синтаксис, лексика и параметр «подлинности» высказывания. На основе этих параметров была предложена трехуровневая шкала соответствия («высокая», «средняя», «низкая»), в рамках которой пять из шести текстов получили оценку «высокая степень соответствия», поскольку демонстрируют автокоммуникацию, ориентированность на внутреннего адресата, спонтанный синтаксис, разговорную лексику, субъективную модальность и бытовую детализацию при отсутствии направленности на внешнего читателя. Один текст (DeepSeek, формат

Interview с именной стратегией социодемографического прайминга, вариант 1) был отнесен к категории со «средней» степенью соответствия: при точной передаче исторического контекста и эмоционального фона он сохраняет более стилизованный характер, меньшую степень спонтанности (в том числе отсутствие характерных для дневниковой речи обрывов фраз, колебаний и вводных конструкций) и более опосредованный образ адресата.

Таким образом, качественный анализ отобранных текстов показывает, что в рассматриваемых примерах большие языковые модели не только воспроизводят внешние формальные признаки дневниковой записи (датировку, повествование от первого лица, соответствующую композицию), но и в значительной степени имитируют глубинные механизмы автокоммуникации — синтаксическую спонтанность (ориентацию на устно-разговорный синтаксис в письменной форме), сложную темпоральную организацию, широкое использование метарефлексивных элементов и устойчивых эмоционально-оценочных стратегий, характерных для человеческих дневниковых практик.

Заключение

Проведенное исследование демонстрирует, что большие языковые модели при определенных комбинациях моделей, характеров принятия роли и стратегий социодемографического прайминга генерируют тексты с параметрами гендерной вариативности речевого поведения, близкими к зафиксированным в дневниковых подкорпусах НКРЯ. Интегральный показатель гендерной близости, объединяющий параметрический и лексемный индексы, подтверждает систематическое сближение распределений частей речи, грамматических категорий, прагматических маркеров и высокочастотных лексем с референсным профилем женских дневниковых записей — прежде всего для *GPT-5.2* и *DeepSeek* при *непосредственном* принятии роли в сочетании с *структурированной* или *именной* стратегией прайминга.

Качественный анализ текстов с максимальными значениями интегрального показателя выявляет в соответствующих комбинациях не только формальную близость параметров, но и жанровые характеристики дневникового письма: автокоммуникативную направленность, характерное соотношение разговорной и книжной лексики, высокую плотность вводных конструкций, модальных и экспрессивных средств — соответствующие описаниям «женского» дневникового письма в гендерной лингвистике. Экспертная автороведческая оценка подтверждает, что такие тексты могут восприниматься как правдоподобные образцы женских дневников, написанных после 1950 г., что дополнительно поддерживает возможность использования синтетических респондентов в задачах моделирования речевого поведения целевых групп.

Вместе с тем характер и пределы этой возможности строго ограничены. Во-первых, в настоящей работе варьируется только один демографический параметр — гендерная принадлежность, заданная ограничениями корпусной

разметки, тогда как возраст, образование, социальный статус и другие характеристики не специфицируются в промптах. Это означает, что синтетические респонденты в текущей конфигурации отражают агрегированный гендерный профиль, а не комплексную личность, и позволяют оценивать лишь применимость подхода к моделированию гендерно маркированного речевого поведения, а не к реконструкции более детальной социальной дифференциации. Во-вторых, параметры гендерной близости зависят от конкретных комбинаций модели и формул личностного социодемографического промптинга: те же формулы, которые для GPT-5.2 и DeepSeek приводят к приближению к женскому профилю, для GigaChat систематически ассоциированы со смещением в сторону показателей мужского подкорпуса, а изменение формы обращения («ты»/«вы») заметно модифицирует интегральный показатель. Тем самым синтетический респондент в рассматриваемой постановке не является устойчивой «цифровой личностью», а представляет собой конфигурационно обусловленную модель, что требует в последующих исследованиях систематического тестирования различных моделей и формул личностного социодемографического промптинга на иных жанрах и типах корпусов. В-третьих, корпусная операционализация гендерной вариативности неизбежно ограничивается параметрами, которые поддаются надежной автоматизированной разметке: частями речи, грамматическими категориями, частотными маркерами экспрессивности, вежливости, модальности, диминутивности и т.п. Тематические предпочтения, сложные стратегии прагматикализации, семантические и оценочные сдвиги, лежащие в центре многих гендерных исследований, остаются за пределами количественного сравнения. Следовательно, совпадение синтетических и реальных текстов по выбранным параметрам следует трактовать как совпадение их статистического профиля, а не полной языковой идентичности; синтетические респонденты в текущем виде моделируют прежде всего формально регистрируемые аспекты гендерной маркированности. Наконец, референсный женский подкорпус дневников количественно существенно уступает мужскому, что ограничивает устойчивость оценок для женского профиля и требует опоры на нормированные показатели частоты (IPM) при интерпретации результатов, как это реализовано в настоящем исследовании.

На этом фоне становится возможным уточнить роль синтетических респондентов в лингвистических исследованиях. Полученные данные свидетельствуют, что синтетические респонденты могут использоваться в качестве исследовательского инструмента при условии предварительной оценки комбинаций моделей и формул личностного социодемографического промптинга на материале наиболее близкого репрезентативного корпуса с развитой лингвистической разметкой, как это сделано для дневниковых текстов НКРЯ. Наличие подобных референсных данных позволяет определить, какие комбинации моделей и формул личностного социодемографического промптинга дают приемлемую степень сходства с реальными данными и какие типы параметров адекватно воспроизводятся. После прохождения

этой стадии калибровки та же методика может переноситься в близкие области с дефицитом прямых данных — личную переписку, конфиденциальные записи, речевые практики маргинализированных или малочисленных групп — где синтетические респонденты уже не заменяют корпус, а выступают как инструмент предварительного моделирования и генерации гипотез.

В таком понимании синтетические респонденты не противопоставляются корпусным данным, а включаются в расширенный инструментарий эмпирической лингвистики как дополнительный источник моделируемых данных, требующий опоры на референсные корпусы. В областях с развитой корпусной инфраструктурой они позволяют систематически варьировать комбинации моделей и формул личностного социодемографического промптинга и оценивать, какие параметры гендерной и жанровой вариативности поддаются устойчивому моделированию. В доменах с ограниченным или закрытым доступом к данным синтетические респонденты могут использоваться прежде всего как инструмент прогнозирования возможных распределений маркеров и проверки теоретических ожиданий, при обязательном последующем сопоставлении с любой доступной эмпирической выборкой, что задает рамки их ответственного применения.

Список литературы

1. *Shrestha P., Krpan D., Koaik F., Schnider R., Sayess D., Binbaz M.S.* Beyond WEIRD: Can synthetic survey participants substitute for humans in global policy research? // Behavioral Science & Policy. 2024. Vol. 10. Iss. 2. P. 26–45. https://doi.org/10.1177/23794607241311793 EDN: RJHEHB
2. *Argyle L.P., Busby E.C., Fulda N., Gubler J.R., Rytting C., Wingate D*. Out of One, Many: Using Language Models to Simulate Human Samples // Political Analysis. 2023. Vol. 31. Iss. 3. P. 337–351. https://doi.org/10.1017/pan.2023.2 EDN: TOAPUN
3. *Пузанова Ж.В., Кожоридзе Г.Г., Кожоридзе Д.Г.* ИИ и социология: анализ технологических возможностей виртуальных респондентов // Социология: методология, методы, математическое моделирование (Социология: 4М). 2025. № 60. С. 216–246. https://doi.org/10.19181/4m.2025.34.1.6 EDN: PRPHTP
4. *Aher G., Arriaga R.I., Kalai A.T.* Using large language models to simulate multiple humans and replicate human subject studies // Proceedings of the 40th International Conference on Machine Learning. PMLR. 2023. Vol. 202. P. 358–390.
5. *Дубнякова О.А., Кашина Т.А.* Коммуникативно-прагматические особенности личного дневника // Вестник МГПУ. Серия «Филология. Теория языка. Языковое образование». 2017. № 3 (23). С. 43–49. EDN: YJGTID
6. *Таннен Д.* Ты меня не понимаешь! : женщины и мужчины в разговоре / пер. с англ. Т. Печурко. М. : Вече ; Персей ; АСТ, 1996.
7. *Кирилина А.В.* Гендер: лингвистические аспекты : монография. М. : Ин-т социологии РАН, 1999. EDN: OXTFTD
8. *Потапов В.В.* Многоуровневая стратегия в лингвистической гендерологии // Вопросы языкознания. 2002. № 1. С. 103–126. EDN: WVGGEK
9. Гендер и язык : антология. М. : Языки славянской культуры, 2005.
10. *Попова Т.И.* Социальные роли говорящего в повседневной коммуникации на русском языке: гендерный аспект : дис. ... канд. филол. наук. СПб., 2022.

11. *Lutz M., Sen I., Ahnert G., Rogers E., Strohmaier M.* The Prompt Makes the Person(a): A Systematic Evaluation of Sociodemographic Persona Prompting for Large Language Models. URL: arXiv.org. 2025. https://arxiv.org/abs/2507.16076
12. *Lans J.C.W.* The "Magic Word" for LLMs: A Study on the Effect of Politeness on LLM Performance [PhD Thesis]. Leiden : Leiden Institute of Advanced Computer Science (LIACS), Leiden University, 2025.
13. *Dobariya O., Kumar A.* Investigating how prompt politeness affects LLM accuracy (short paper). URL: arXiv. 2025. https://arxiv.org/abs/2510.04950
14. *Шмелева Т.В.* Речевой жанр (возможности описания и использования в преподавании языка) // Russistik. Русистика: научный журнал актуальных проблем преподавания русского языка. 1990. № 2. С. 20–32.
15. *Моисеева Т.Ф., Огорелков И.В.* Современная автороведческая экспертиза. М. : Российский государственный университет правосудия (РГУП), 2022.

References

1. Shrestha, P., Krpan, D., Koaik, F., Schnider, R., Sayess, D., & Binbaz, M.S. (2024). Beyond WEIRD: Can Synthetic Survey Participants Substitute for Humans in Global Policy Research? *Behavioral Science & Policy*, *10*(2), 26–5. https://doi.org/10.1177/23794607241311793 EDN: RJHEHB
2. Argyle, L.P., Busby, E.C., Fulda, N., Gubler, J.R., Rytting, C., & Wingate, D. (2023). Out of One, Many: Using Language Models to Simulate Human Samples. *Political Analysis*, *31*(3), 337–351. https://doi.org/10.1017/pan.2023.2 EDN: TOAPUN
3. Puzanova, Zh.V., Kozhoridze, G.G., & Kozhoridze, D.G. (2025). AI and Sociology: an Analysis of the Technological Capabilities of Virtual Respondents. *Sotsiologiya: Metodologiya, Metody, Matematicheskoe Modelirovanie (Sotsiologiya: 4M)*, 60, 216–246. (In Russ.). https://doi.org/10.19181/4m.2025.34.1.6 EDN: PRPHTP
4. Aher, G., Arriaga, R.I., & Kalai, A.T. (2023). Using Large Language Models to Simulate Multiple Humans and Replicate Human Subject Studies. In: *Proceedings of the 40th International Conference on Machine Learning* (Vol. 202, pp. 358–390). PMLR.
5. Dubnyakova, O.A., & Kashina, T.A. (2017). Communicative and Pragmatic Particularities of Personal Diary. *MCU Journal of Philology. Theory of Linguistics. Linguistic Education*, *3*(23), 43–49. (In Russ.). EDN: YJGTID
6. Tannen, D. (1990). *You Just Don't Understand: Women and Men in Conversation*. New York: Ballantine Books. (In Russ.).
7. Kirilina, A.V. (1999). *Gender: Linguistic Aspects*. Moscow: Institute of Sociology, Russian Academy of Sciences. (In Russ.). EDN: OXTFTD
8. Potapov, V.V. (2002). A Multilevel Strategy in Linguistic Genderology. *Topics in the Study of Language*, 1, 103–126. (In Russ.). EDN: WVGGEK
9. Gender and language: An Anthology. (2005). Moscow: Yazyki slavyanskoy kultury. (In Russ.).
10. Popova, T.I. (2022). Social Roles of the Speaker in Everyday Communication in Russian: A Gender Perspective [PhD in Philology]. Saint Petersburg: Saint Petersburg State University. (In Russ.).
11. Lutz, M., Sen, I., Ahnert, G., Rogers, E., & Strohmaier, M. (2025). The Prompt Makes the Person(a): a Systematic Evaluation of Sociodemographic Persona Prompting for Large Language Models. *arXiv.* https://arxiv.org/abs/2507.16076
12. Lans, J.C.W. (2025). *The "Magic Word" for LLMS: A Study on the Effect of Politeness on LLM Performance* [PhD Thesis]. Leiden: Leiden University, Leiden Institute of Advanced Computer Science, Leiden University.
13. Dobariya, O., & Kumar, A. (2025). Investigating How Prompt Politeness Affects LLM Accuracy (Short Paper). *arXiv* https://arxiv.org/abs/2510.04950

14. Qi, P., Zhang, Y., Zhang, Y., Bolton, J., & Manning, C.D. (2020). Stanza: A Python Natural Language Processing Toolkit for Many Human Languages. *arXiv.* https://arxiv.org/abs/2003.07082
15. Shmeleva, T.V. (1990). Speech Genre: Possibilities of Description and Use in Language Teaching. *Russistik. Russian Language Studies: A Journal of Current Problems of Teaching Russian*, 2, 20–32. (In Russ.).
16. Moiseeva, T.F., & Ogorelkov, I.V. (2022). *Modern Authorship Examination*. Moscow: Russian State University of Justice (RGUP). (In Russ.).

Сведения об авторах:
Ваничкина Александра Савельевна, кандидат филологических наук, директор, Институт информационных наук, Московский государственный лингвистический университет (119034, Российская Федерация, г. Москва, ул. Остоженка, д. 38 с. 1); *сфера научных интересов*: лингвопрагматика, когнитивная лингвистика, аксиология, искусственный интеллект; *e-mail*: alexvanichkina@gmail.com
ORCID: 0000-0002-1748-7666; SPIN-код: 6174-8813.
Изюмская-Капитонова Вероника Викторовна, младший научный сотрудник, лаборатория корпусной лингвистики, Московский государственный лингвистический университет (119034, Российская Федерация, г. Москва, ул. Остоженка, д. 38 с. 1); *сфера научных интересов*: прикладная лингвистика, корпусная лингвистика, обработка естественного языка, большие языковые модели, искусственный интеллект; *e-mail:* v.v.linguist.ik@gmail.com
ORCID: 0009-0006-9151-0269.

Information about the authors:
Alexandra S. Vanichkina, PhD in director, Institute of Information Sciences, Moscow State Linguistic University (38 build 1 Ostozhenka Str., Moscow, Russian Federation, 119034); *Research interests*: linguopragmatics, cognitive linguistics, axiology, artificial intelligence; *e-mail:* alexvanichkina@gmail.com
ORCID: 0000-0002-1748-7666; SPIN-код: 6174-8813.
Veronika V. Izyumskaya-Kapitonova, Junior Researcher, Corpus Linguistics Laboratory, Moscow State Linguistic University (38 build 1 Ostozhenka Str., Moscow, Russian Federation, 119034); *Research interests*: applied linguistics, corpus linguistics, natural language processing, large language models, artificial intelligence*; e-mail*: v.v.linguist.ik@gmail.com
ORCID: 0009-0006-9151-0269.

RUDN Journal of Language Studies, Semiotics and Semantics
2313-2299 (Print), ISSN 2411-1236 (Online)
Вестник Российского университета дружбы народов.
Серия: ТЕОРИЯ ЯЗЫКА. СЕМИОТИКА. СЕМАНТИКА

2026 Vol. 17 No. 2 557–568
http://journals.rudn.ru/semiotics-semantics

DOI: 10.22363/2313-2299-2026-17-2-557-568
EDN: LSTNOS

Научная статья / Research Article

Токен как алгоритмическая «тень» морфемы

К.П. Жикулина

Российский университет дружбы народов, *Москва, Российская Федерация*
✉zhikulina-kp@rudn.ru

Аннотация. Актуальность исследования критического этапа обработки текста в эпоху доминирования больших языковых моделей обусловлена системным расхождением между статистически выделенными границами токенов и лингвистически обоснованными границами морфем в токенизаторах, что проявляется в работе с языками с богатой морфологией. Научная новизна исследования заключается в рассмотрении *токена* не как технической единицы, а как лингвистического артефакта — *алгоритмической «тени» морфемы*, утратившей семантическую мотивированность, но сохранившую внешнее сходство. Цель исследования — выявить лингвистические потери, возникающие при статистической сегментации слов и оценить степень расхождения между токенными и морфемными границами. Методы исследования: сопоставительный лингвистический анализ, корпусный анализ, критический дискурс-анализ и типологическое моделирование. В ходе исследований разработана и эмпирически обоснована типология морфологических искажений при BPE-токенизации, описывающая три типа искажений и расхождений между морфемными и токенными границами: фрагментация, контаминация и аберрация. На материале лексем подорожник и высокотехнологичный экспериментально доказана трехступенчатая деградация морфологической прозрачности при переходе от одноязычных к двуязычным и мультиязычным BPE-моделzv. Установлено, что мультиязычные токенизаторы полностью утрачивают морфологические свойства в отношении отдельного языка, множа статистические артефакты, которые не идентифицируются классическим набором единиц морфологии и носителями данного языка.

Ключевые слова: искусственный интеллект, компьютерная лингвистика, машинное обучение, BPE-токенизация, НЛП-системы, большая языковая модель, сублексемная сегментация, морфология, морфемизация, подслово

Финансирование: публикация подготовлена в рамках грантовой поддержки научных проектов РУДН 050227-0-000.

Конфликт интересов: автор заявил об отсутствии конфликтов интересов.

История статьи: дата поступления: 20.01.2026; дата приема в печать: 10.02.2026.

Для цитирования: *Жикулина К.П.* Токен как алгоритмическая «тень» морфемы // Вестник Российского университета дружбы народов. Серия: Теория языка. Семиотика. Семантика. 2026. Т. 17. № 2. С. 557–568. https://doi.org/10.22363/2313-2299-2026-17-2-557-568 EDN LSTNOS

© Жикулина К.П., 2026
This work is licensed under a Creative Commons Attribution-NonCommercial 4.0 International License
https://creativecommons.org/licenses/by-nc/4.0/legalcode

Token as an Algorithmic ‘Shadow’ of a Morpheme

Kristina P. Zhikulina

RUDN University, *Moscow, Russian Federation*
✉zhikulina-kp@rudn.ru

Abstract. The relevance of the study of a critical stage of text processing in the era of the dominance of large language models is due to the systemic divergence between statistically identified token boundaries and linguistically justified morpheme boundaries in tokenizers, which is particularly evident when working with rich morphology languages. The scientific novelty of the study lies in considering the token not as technical part, but as a linguistic artifact – an algorithmic ‘shadow’ of a morpheme that has lost its semantic motivation but preserved its external resemblance. The aim of the research is to identify linguistic losses that occur during statistical segmentation of words and to assess the degree of divergence between token and morpheme boundaries. Research methods: comparative linguistic analysis, corpus analysis, critical discourse analysis and typological modeling. During the research, a typology of morphological distortions in BPE tokenization has been developed and empirically substantiated, describing three types of distortions and divergence between token and morpheme boundaries: fragmentation, contamination, and aberration. Using the lexemes ‘podorozhnik’ and ‘vysokotekhnologichny’ as material, the three-stage degradation of morphological transparency in the transition from monolingual to bilingual and multilingual BPE models has been experimentally proven. It has been established that multilingual tokenizers completely lose morphological properties in relation to a particular language multiplying statistical artifacts that are not identified by a classical set of morphological units and native speakers of this language.

Keywords: artificial intelligence, computational linguistics, machine learning, BPE-tokenization, NLP-systems, LLM, subword segmentation, morphology, morphemization

Founding: The publication has been supported by the RUDN University Scientific Projects Grant System, project No 050227-0-000.

Conflict of interests: The author stated that there were no conflicts of interest.

Article history: Received: 20.01.2026; Accepted: 10.02.2026.

For citation: Zhikulina, K.P. (2026). Token as an Algorithmic ‘Shadow’ of a Morpheme. *RUDN Journal of Language Studies, Semiotics and Semantics, 17*(2), 557–568. (In Russ.). https://doi.org/10.22363/2313-2299-2026-17-2-557-568 EDN LSTNOS

Введение

Стремительное развитие генеративных моделей искусственного интеллекта (далее — ИИ), способных имитировать мыслительные процессы человека, обучаться, «рассуждать», поставило перед лингвистикой принципиально новые вопросы. Если ранее язык рассматривался как система знаков, существующая независимо от способов ее фиксации и передачи, то в эпоху больших языковых моделей (LMM) язык все чаще предстает как объект алгоритмической обработки, подчиняющийся не только программным закономерностям, но и лингвистическим. Центральным звеном, опосредующим взаимодействие между человеческим текстом и LMM, становится *машинное обучение* (ML). Под *машинным обучением* понимается узкоспециализированная область, в основу которой положен принцип обучения машин точному

выполнению определенных задач с помощью выявления закономерностей и детального прогнозирования. Благодаря подходу, называемому *непрерывным обучением*, машина способна адаптироваться под разные задачи и обстоятельства, совершенствовать свои функции и выдавать более точные результаты.

Практически все современные нейросетевые системы — от ChatGPT до Llama — используют *сублексемную сегментацию,* при которой слова разбиваются на единицы меньше слова, но больше символа — *подслова*. Наиболее распространенным методом такой сегментации является Byte Pair Encoding (BPE), адаптированный для задач обработки естественного языка Р. Сенрихом с соавторами [1].

Токенизаторы, использующиеся в современных нейросетевых системах, все чаще демонстрируют системное расхождение между статистически выделенными границами токенов и лингвистически обоснованными границами морфем. Именно на этапе *токенизации* закладываются границы членения, которые впоследствии определяют, как LMM интерпретирует язык.

Это расхождение не является случайным и незначительным — оно носит системный характер и особенно проявляется в языках с богатой морфологией, таких как русский, немецкий, тюркские языки. Модель оперирует единицами, внешне напоминающими морфемы, но не наследуют их семантическую и структурную мотивированность. Возникает парадоксальная ситуация: инструмент, созданный для более эффективной обработки языка, порождает лингвистические артефакты, ставящие под сомнение саму возможность классического морфологического анализа в рамках архитектур нейронных сетей. Как справедливо отмечает О.И. Максименко, современный этап развития языкознания характеризуется сближением традиционного лингвистического описания с методами алгоритмической обработки текста, что требует пересмотра устоявшихся понятий в контексте их алгоритмической операционализации [2]. В настоящем исследовании такому пересмотру подвергается категория морфемы, которая в условиях *сублексемной сегментации* утрачивает статус минимально значимой единицы и замещается статистически выделенным токеном — алгоритмической «тенью».

Актуальность исследования обусловлена не только теоретической значимостью поставленной проблемы, но и ее практическими импликациями. Современные исследования в области NLP фокусируются на улучшении технических параметров токенизации — сжатии словарей, скорости обработки, фиксации и распознаванию редких слов (см., например, [3–6]), — игнорируя при этом *морфологическую прозрачность* как критерий качества. Ее утрата при алгоритмической обработке ведет не просто к техническим погрешностям, но и к лингвистической аберрации — искажению способа существования языка в цифровой среде. Проблемным местом оказываются современные *токенизаторы*, как правило, являющиеся мультиязычными: обучаются на разнородных текстах, но должны одинаково хорошо работать для сотен языков одновременно. Такой подход неизбежно приводит к *билингвально-*

ориентированному или *полилингвальному алгоритму членения*, при котором границы *токенов* определяются не морфемной структурой одного языка, а частотной статистикой, усредненной по всем языкам обучающего корпуса. Мнимая универсальность достигается ценой утраты лингвистической специфики: языки с богатой морфологией начинают сегментироваться по подобию языков с другой структурой — преимущественно английского языка, что порождает феномен *лексической интерференции*.

Таким образом, исследование ставит перед собой следующие задачи:

1) провести сравнительный анализ трех типов *токенизации* (однозначной, двуязычной и мультиязычной) на материале русских лексем *подорожник* и *высокотехнологичный* с прозрачной морфемной структурой;

2) разобрать классификацию типов морфологических искажений при BPE-сегментации, описав их лингвистическую природу и механизм возникновения;

3) предложить предварительную систему критериев для оценки морфологической прозрачности *токенизаторов*;

4) очертить границы применимости статистической сегментации для задач, требующих морфологического анализа.

Теоретическая значимость исследования заключается в пересмотре онтологического статуса *токена* как лингвистической категории — *алгоритмической «тени» морфемы*, внешне подобной морфеме, но лишенной семантической и структурной мотивированности. *Токен* впервые осмысляется как *псевдоморфемная единица* — алгоритмическая проекция морфемы, утратившая этимологическую, парадигматическую и словообразовательную мотивированность, но сохранившая подобие. Введенное понятие *алгоритмической «тени» морфемы* позволяет концептуализировать принципиальное различие между языковым наследием и цифровым компромиссом.

Практическая значимость исследования состоит в том, что его результаты позволяют разработчикам NLP-систем выявлять слабые места существующих *токенизаторов* и выбирать стратегию сегментации в зависимости от языка и задачи, предлагают критерии для проектирования лингвистически адекватных *токенизаторов* для низкоресурсных и морфологически богатых языков, а также служат дидактическим материалом для универсальных курсов по компьютерной лингвистике, искусственному интеллекту в филологии и алгоритмической лингвистике.

Обсуждение и результаты исследования

Обработка естественного языка (NLP) направлена на решение таких задач, как обучение компьютеров понимать, интерпретировать и генерировать живую речь. В компьютерной лингвистике NLP означает, прежде всего, обучение ИИ распознаванию закономерностей в человеческой речи, текстовых источниках и использованию этих же закономерностей для анализа пользовательского запроса, расшифровки и создания реакции, то есть ответа.

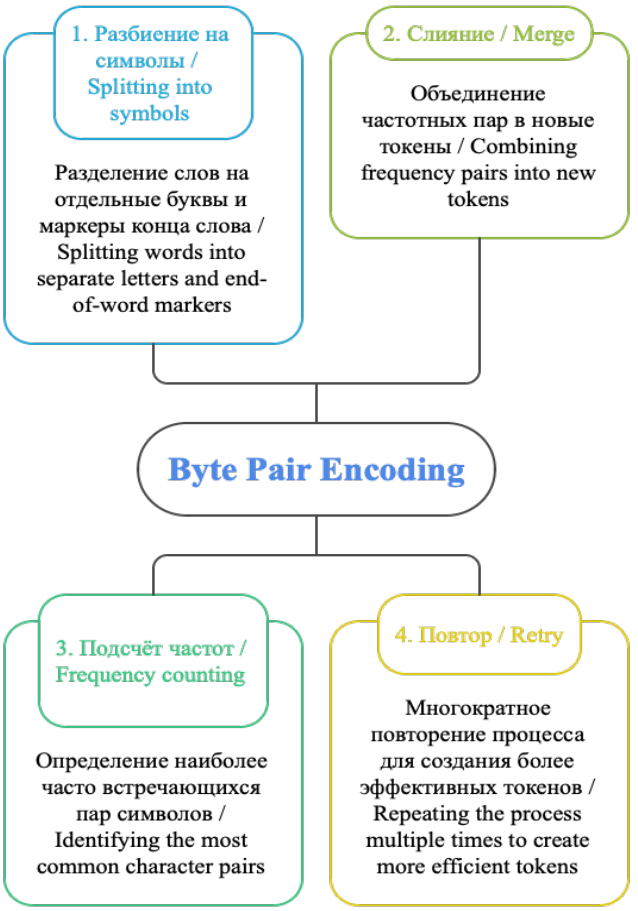
Одним из фундаментальных процессов в NLP является *токенизация*, при которой происходит преобразование исходного текстового материала в формат, воспринимаемый машиной для анализа и моделирования. Текст разбивается на фрагменты — маленькие единицы, называемые *токенами*. *Токенами* могут выступать как отдельные слова так и части слов или символы, в зависимости от требуемого уровня фрагментарности и детализации. Они помогают понимать контекст и разрабатывать модель для обработки естественного языка, называемую сегодня *большой языковой моделью (LLM)*.

В машинном обучении существует несколько типов *токенизации*. *Токены* могут различаться в зависимости от используемого метода *токенизации* и уровня лингвистической абстракции, под которой в исследуемой нами области подразумеваются обобщенные классы, свойства и неосязаемые сущности. Например, ограничения современных нейронных систем машинного перевода связаны с тем, что такие модели используют словарный запас в 30 000–500 000 *токенов*, постоянно обращаясь к внешним словарям — источникам данных, создавая искусственный барьер, не соответствующий природе естественного языка: для продуктивных словообразовательных процессов, таких как транслитерация имен собственных в русском языке, необходимо применять методы обработки данных, связанных с единицами меньше, чем целое слово, одновременно включающее префиксы, суффиксы, интерфиксы и окончания. Р. Сенрих, Б. Хэддоу и А. Бёрч экспериментальным путем выяснили, что *сублексемная сегментация* может решать проблемы, связанные с распознаванием и нахождением неизвестных слов без обращения к внешним словарям. Также они доказали, что подобный метод обработки данных позволяет улучшать качество перевода. По полученным данным, современные механизмы *сублексемной сегментации* повышают качество и эффективность перевода с английского на русский язык с 6,6 до 18,3 % [1. P. 1716–1718].

Для сжатия данных часто используется алгоритм Byte Pair Encoding (BPE), адаптированный под сегментацию слов на *сублексемы* в задачах по обработке естественного языка (рис.).

В результате *морфемизации* получается словарь, в котором существуют отдельные буквы, частотные биграммы, целые слова (в случае высокой частотности), морфемы, слоги и любые другие устойчивые последовательности. Необходимость для такого членения в нейронных сетях возникает из-за невозможности обработки бесконечных потоков данных. Если оставлять целые слова, словарь будет слишком большим, и на этапе обучения LLM не сможет оперировать данными в условиях ограниченной памяти. Словарь позволяет сохранять и использовать фиксированный и компактный набор *токенов* (от 50000 до 100000 в современных моделях), а также сборку слов из *подслов*, даже если их не было в обучении.

Проблема эффективной токенизации для больших языковых моделей остается актуальной и важной в эпоху ИИ. Исследователи и разработчики ищут баланс между компрессией словаря и способностью нейронной сети к морфологическому восприятию и *морфемно-морфологической обработке* (табл.).



Этапы алгоритмической морфемизации методом BPE
Источник: составлено К.П. Жикулиной с помощью Napkin.ai[1]
The stages of algorithmic morphemization by the BPE method
Source: compiled by Kristina P. Zhikulina with Napkin.ai

На основе сопоставительного анализа трех поколений BPE-*токенизаторов* (одноязычный, двуязычный и мультиязычный) были выделены следующие критерии:

- ***Фрагментация***, которая наблюдается, когда одна морфема последовательно разбивается на несколько статистически частотных *токенов* в корпусе, однако совпадение с морфемой частичное (например, *до* и *рож* вместо *дорож*);
- ***Контаминация***, при которой происходит соединение частей разных морфем (например, интерфикс -о- с началом или концом корней) и образуется *гибридный токен*;

[1] Нейронная сеть для генерации изображений Napkin.ai // napkin.ai. Режим доступа: https://www.napkin.ai (дата обращения: 20.01.2026).

• ***Аберрация*** происходит тогда, когда *токен* не имеет аналога в морфемной структуре используемого языка, при этом не является слогом и не несет никакого семантического или структурного смысла, превращаясь в статистический артефакт (например, *со* и *ко*).

Типология морфологических искажений при сублексемной сегментации (на примере лексем «подорожник» и «высокотехнологичный») Typology of morphological distortions in subword segmentation (using words 'podorozhnik' and 'vysokotekhnologichny' as examples)

Вид токенизации / Type of tokenization	**Границы токенов / Boundaries of tokens**	**Лингвистическое искажение / Linguistic distortion**
Правильное морфемное членение слова *подорожник*: по | дорож | ник / The correct morphemic division of the word *psyllium*: po | dorozh | nik		
Одноязычная BPE (только русский язык) / Monolingual BPE (only Russian language)	по | до | рож | ник / po | do | rozh | nik	• *Фрагментация и статическая аппроксимация* Корень *дорож* разбит на частотные встречаемости пар символов *до* и *рож*, а приставка *по-* и суффикс *-ник-* узнаваемы. Морфологически неточное членение, но узнаваемое. • Fragmentation and static approximation The root *dorozh* is divided into frequency of occurrence pairs of characters *do* and *rozh*, and the prefix *po-* and the suffix *-nik-* are recognizable. Morphologically inaccurate division, but recognizable.
Двуязычная (или совместная) BPE (английский и русский языки) / Bilingual BPE (or collaborative) (English and Russian languages)	по | дор | ож | ник / po | dor | ozh | nik	• *Контаминация и лексическая интерференция* Корень *дорож* членится под английские слова *door (дор)* и частотный токен окончания англо-французских слов *ож (garage)*. Размытие морфем происходит под влиянием паттернов языка-донора. • Contamination and lexical interference The root *dorozh* is divided into English word *door* and the frequency token ending of English-French words *ozh* (garage). Morpheme blurring occurs under influence of patterns of the donor language.
Мультиязычная BPE (более 100 языков) / Multilingual BPE (over 100 languages)	под | оро | жник / pod | oro | zhnik	• *Статистическая аберрация* Токены не соответствуют морфемам и слогам: *под* – английское *under*, *оро* – итальянское *золото*, *-жник* – артефакт, или бессмысленный набор символов.

Продолжение табл. / Continuation of the Table

Вид токенизации / Type of tokenization	**Границы токенов / Boundaries of tokens**	**Лингвистическое искажение / Linguistic distortion**
		• *Statistical aberration* Tokens do not correspond to morphemes and syllables: *pod* is English *under*, *opo* – is Italian *gold*, *zhnik* – is an artifact, or a meaningless set of characters.
Правильное морфемное членение слова *высокотехнологичный*: высок | о | техн | о | лог | ич | н | ый / The correct morphemic division of the word *high-tech*: vysok | o | tehn | o | log | ich | n | yj		
Одноязычная BPE (только русский язык) / Monolingual BPE (only Russian language)	высоко | тех | но | ло | гич | ный / vysoko | teh | no | lo | gich | nyj	• *Фрагментация с элементами аппроксимации* Корень *высок* соединён с интерфиксом *-о-*, корень *тех* выделен неверно, потеря второго интерфикса *-о-*, корень *лог* и суффикс *-ич-* разделены на два токена *ло* и *гич*, не выделен отдельно суффикс *-н-*, окончание воспроизведено как *-ный*. Морфологическая путаница присутствует, но слово остаётся опознаваемым. • *Fragmentation with approximation elements* The root *vysok* is connected to the prefix *-o-*, the root *teh* is highlighted incorrectly, the loss of the second prefix *-o-*, the root *log* and suffix *-ich-* are divided into two tokens *lo* and *gich*, the suffix *-n-* is not highlighted separately, the ending is reproduced as *-nyj*. Morphological confusion is present, but the word remains recognizable.
Двуязычная (или совместная) BPE (английский и русский языки) / Bilingual (or collaborative) BPE (English and Russian languages)	вы | со | ко | те | хно | ло | гич | ный / vy | so | ko| te | hno | lo | gich | nyj	• *Фрагментация с контаминацией* Токен *вы* отделён от корня, *со* и *ко* становятся статистическими артефактами, *хно* – гибридный токен, соединяющий часть корня *техн* и интерфикс *-о-*. Интерференция английского *technology* ведёт к разрыву корня. / • *Fragmentation with contamination* The token *vy* is separated from the root, *so* and *ko* become statistical artifacts, and *hno* is a hybrid token connecting a part of the root *tehno* and prefix *-o-*. The interference of English technology leads to a rupture of the root.

Окончание табл. 1 Ending of the Table

Вид токенизации / Type of tokenization	**Границы токенов / Boundaries of tokens**	**Лингвистическое искажение / Linguistic distortion**
Мультиязычная BPE (более 100 языков) / Multilingual BPE (over 100 languages)	высок | отех | ноло | гич | ный / vysok | oteh | nolo | gich | nyj	• *Контаминация с аберрацией* *Отех* – соединение корня с интерфиксом *-о-*, *ноло* – гибридный токен от английского и латинского *nolo*. Морфологическая структура полностью разрушена, границы токенов утрачены - нет соответствий ни морфемам, ни слогам. / • *Contamination with aberration* *Oteh* – it is a compound of the root with the prefic *-o-*, *nolo* – it is a hybrid token from English and Latin. The morphological structure is completely destroyed; the boundaries of the tokens are lost – there are not correspondences to either morphemes or syllables.

Источники: токенизация выполнена К.П. Жикулиной с использованием OpenAI Tokenizer (tiktoken) для GPT-4 [7], Hugging Face Transformers (AutoTokenizer.from_pretrained("meta-llama/Llama-3-8b")) [6], интерактивного инструмента токенизации GPT Tokenzier [4], демонстративный BPE-токенизатор FatimaALzahrani с пошаговой визуализацией [1].
Source: compiled by Kristina P. Zhikulina with using OpenAI Tokenizer (tiktoken) for GPT-4 [7], Hugging Face Transformers (AutoTokenizer.from_pretrained("meta-llama/Llama-3-8b")) [6], interactive tokenization tool GPT Tokenzier [4], demonstrative FatimaALzahrani BPE tokenizer with step-be-step visualization [1].

Границы между типами морфологических расхождений определяются степенью семантической мотивированности *токена*: в случае фрагментации морфема узнаваема, но бывает раздроблена; в случае контаминации морфема неузнаваема, а *токен* выступает гибридным образованием; в случае аббревиации морфема отсутствует, а *токен* — носитель *псевдоморфемного смысла*.

Описанные выше искажения являются следствием детерминированного характера классического BPE-*токенизатора*: каждое слово получает одно статистически вероятное частотное членение [1]. Однако Т. Кудо обращает внимание на то, что *сублексемная сегментация* неоднозначна: одно и то же слово может быть разделено множеством способов даже при условии использования фиксированного словаря [7]. Предложенный им метод доказывает, что, с лингвистической точки зрения, алгоритмическая «тень» морфемы может переставать быть фиксированной и становиться проекциями одно и того же языкового объекта. Вероятностный взгляд Т. Кудо на *токенизацию* открывает перспективу для снижения описанных выше морфологических искажений и потерь с помощью обучения модели на определенном множестве «теней», а не сфокусированной в одной точке.

Существуют и более продвинутые инструменты, как SentencePiece, который не требует предварительной сегментации текста на слова

и рассматривает пробел как отдельный символ, что делает токенизацию обратимой и независимой от языка [8. P. 67–70]. Но на практике оказывается, что алгоритм перестает получать подсказки о том, где проходят границы слов, и вынужден восстанавливать их по частотности: «тень» морфемы становится более автономной и менее привязанной к языковой реальности.

О.В. Ломакина предлагает разграничение ценностных констант и ценностных переменных, что имеет место и в алгоритмической лингвистике [9. С. 623–625]. В системе *сублексемной сегментации* также можно говорить о морфемных константах — корнях, приставках, суффиксах, которые сохраняют свою целостность даже при многократном прохождении через BPE-сжатие [10. С. 28–29], и морфемных переменных — единицах, членимость которых варьируется в зависимости от обучающего корпуса модели, степени мультиязычности и числа операций слияния. Так, на примере слова *подорожник* (см. Табл.), корень *дороож* в русском языке, при однозначной *токенизации*, демонстрирует константное поведение, а в мультиязычной модели превращается в переменную — утрачивается морфемная идентичность.

Заключение

Типология морфологических расхождений позволяет говорить о неравноценном лингвистическом членении, а также о невозможности следования классической морфологической схеме русского языка в современных условиях технических возможностей ИИ, а именно — ограничениям памяти и скорости обработки слишком больших массивов данных. Анализ показывает тесную связь теоретической морфологии и инженерных практик: интервенцию в морфологический строй языка. *Типология морфологических расхождений* — это инструмент лингвистической рефлексии над алгоритмическими практиками. Она позволяет определить, что именно происходит с морфемой при BPE-сегментации, классифицировать степень и характер искажения, оценивать универсальность каждого отдельного языка.

Классическая морфология оперирует устойчивыми и семантически расширенными морфемами, границы которых определяет этимология, парадигматика, словообразование и словоизменение. Однако можно отметить, что статистическая сегментация в нейронных сетях не дополняет базовые принципы, а конкурирует с ними, предлагая иную онтологию членения речи — по частоте. *Токен* выступает в качестве алгоритмической «тени» морфемы, так как является порождением алгоритма, морфема же дана языком. Также «тень» лишь повторяет очертания, но не сущность: *токен* внешне схож с морфемой, но не имеет пока что этимологических, парадигматических и словообразовательных гнезд. Морфема есть в языке, но *токен* – это то, что осталось от морфемы после прохождения через алгоритм сжатия, однако он может указывать на нее, но не повторять или замещать. Можно сказать, что морфема является языковым наследием, а токен — компромисс цифровой стратификации языка. Чем более универсальной становится LLM, тем более искаженной становится алгоритмическая «тень» морфемы.

Список литературы

1. *Sennrich R., Haddov B., Brich A.* Neural Machine Translation of Rare Words with Subword Units // Proceedings of the 54th Annual Meeting of the Association for Computational Linguistics. N.-Y. : Cornell University, 2016 P. 1715–1725. https://doi.org/10.48550/erXiv.1508.07909
2. *Максименко О.И.* Формализованная и цифровая лингвистика. М. : Государственный университет просвещения, 2024. EDN: NPXJZB
3. *Никитина Е.Н., Станкевич М.А., Ларионов Д.С.* Эмотивные глаголы с отрицанием: опыт интерпретации данных автоматического анализа текстов // Медиалингвистика. 2025. Т. 12. № 2. С. 174–198. https://doi.org/10.21638/spbu22.2025.201 EDN: TDQIWC
4. *Гудков В.В., Митренина О.В., Соколов Е.Г., Коваль А.А.* Языковой перенос нейросетевого обучения для частеречной разметки Санкт-Петербургского корпуса агиографических текстов (СКАТ) // Вестник Санкт-Петербургского университета. Язык и литература. 2023. Т. 20. № 2. С. 268–282. https://doi.org/10.21638/spbu09.2023.205 EDN: VQHOZM
5. *Польщикова О.Н.* Становление и формирование терминологии компьютерной лингвистики // Вопросы журналистики, педагогики, языкознания. 2022. Т. 41. № 3. С. 590–607. https://doi.org/10.52575/2712-7451-2022-41-3-590-607 EDN: HTWXDI
6. *Горожанов А.И., Степанова Д.В.* Интерпретация художественного произведения: корпусный подход // Филологические науки. Вопросы теории и практики. 2022. № 15 (1). С. 203–208. https://doi.org/10.30853/phil20220020 EDN: TCZLAF
7. *Kudo T.* Subword Regularization: Improving Neural Network Translation Models with Multiple Subword Candidates // arXiv.org. 2018. Режим доступа: https://arxiv.org/abs/1804.10959 (дата обращения: 18.12.2025).
8. *Kudo T., Richardson J.* SentencePiece: A Simple and Language Independent Subword Tokenizer and Detokenizer for Neural Text Processing // Proceedings of the 2018 Conference on Empirical Methods in Natural Language Processing: System Demonstrations. Brussels. P. 66–71. 2018 https://doi.org/10.48550/arXiv.1808.06226
9. *Ломакина О.В.* Метаязык современной лингвокультурологии: паремиологический уровень // Вестник Российского университета дружбы народов. Серия: Теория языка. Семиотика. Семантика. 2025. Т. 16. № 3. С. 622–637. https://doi.org/10.22363/2313-2299-2025-16-3-622-637 EDN: DQGURR
10. *Зонова Д.Ю.* Принцип работы и проблемы «generative pre-trained transformer artificial intelligence» // Вестник науки и образования. 2023. № 8 (139). С. 28–31. EDN: AFNBRE

References

1. Sennrich, R., Haddov, B., & Brich, A. (2016). Neural Machine Translation of Rare Words with Subword Units. In: *Proceedings of the 54th Annual Meeting of the Association for Computational Linguistics* (pp. 1715–1725). N-Y.: Cornell University publ. https://doi.org/10.48550/erXiv.1508.07909
2. Maksimenko, O.I. (2024). *Formalized and Digital Linguistics*. Moscow: State University of Education publ. (In Russ.). EDN: NPXJZB
3. Nikitina, E.N., Stankevich, M.A., & Larionov, D.S. (2025). Psych Verbs Used with Negation: A Way of Interpretation of Automatic Text Analysis Data. *Media Linguistics*, *12*(2), 174–198. (In Russ.). https://doi.org/10.21638/spbu22.2025.201 EDN: TDQIWC
4. Gudkov, V.V., Mitrenina, O.V., Sokolov, E.G., & Koval, A.A. (2023). Language-Based Transfer Learning Approaches for Part-of-Speech Tagging on Saint Petersburg Corpus of Hagiographic Texts (Skat). *Vestnik of Saint Petersburg University. Language and Literature*, *20*(2), 268—282. (In Russ.). https://doi.org/10.21638/spbu09.2023.205 EDN: VQHOZM

5. Polshchykova, O.N. (2022). Formation of Computational Linguistics Terminology. *Issues in Journalism, Education, Linguistics*, *41*(3), 590–607. (In Russ.). https://doi.org/10.52575/2712-7451-2022-41-3-590-607 EDN: HTWXDI
6. Gorozhanov, A.I., & Stepanova, D.V. (2022). Work of Fiction Interpretation: Corpus Approach. *Philology. Theory & Practice*, *15*(1), 203–208. (In Russ.). https://doi.org/10.30853/phil20220020 EDN: TCZLAF
7. Kudo, T. (2018). Subword Regularization: Improving Neural Network Translation Models with Multiple Subword Candidates // arXiv.org. URL: https://arxiv.org/abs/1804.10959 (accessed: 18.12.2025).
8. Kudo, T., & Richardson, J. (2018). SentencePiece: A Simple and Language Independent Subword Tokenizer and Detokenizer for Neural Text Processing. In: *Proceedings of the 2018 Conference on Empirical Methods in Natural Language Processing: System Demonstrations* (pp. 66–71). Brussels. https://doi.org/10.48550/arXiv.1808.06226
9. Lomakina, O.V. (2025). Metalanguage of Modern Linguaxiology: Paremiological Level. *RUDN Journal of Language Studies. Semiotics and Semantics*, *16(*3), 622–637. (In Russ.). https://doi.org/10.22363/2313-2299-2025-16-3-622-637 EDN: DQGURR
10. Zonova, D.Yu. (2023). The Principle of Operation and Problems of "Generative Pre-trained Transformer Artificial Intelligence". *Bulletin of Science and Education*, 8 (139), 28–31. (In Russ.). EDN: AFNBRE

Сведения об авторе:
Жикулина Кристина Петровна, кандидат филологических наук, ассистент кафедры общего и русского языкознания филологического факультета, Российский университет дружбы народов (117198, Российская Федерация, г. Москва, ул. Миклухо-Маклая, д. 6); *сфера научных интересов*: искусственный интеллект, компьютерная лингвистика, нейролингвистическое программирование, когнитивистика, голосовые помощники, нейронные сети, современный русский язык; *e-mail*: zhikulina-kp@rudn.ru
ORCID: 0000-0003-2488-4616; SPIN-код: 5871-9336, Scopus Author ID: 59236754500.

Information about the author:
Kristina P. Zhikulina, PhD in Philology, Assistant of the Department of General and Russian Linguistics, Faculty of Philology, RUDN University (6 Miklukho-Maklaya str., Moscow, Russian Federation, 117198); *Research interests*: artificial intelligence, computational linguistics, neuro-linguistic programming, cognitive science, voice assistance, neural networks, modern Russian language; *e-mail*: zhikulina-kp@rudn.ru
ORCID: 0000-0003-2488-4616; SPIN-code: 5871-9336, Scopus Author ID: 59236754500.

RUDN Journal of Language Studies, Semiotics and Semantics
2313-2299 (Print), ISSN 2411-1236 (Online)
Вестник Российского университета дружбы народов.
Серия: ТЕОРИЯ ЯЗЫКА. СЕМИОТИКА. СЕМАНТИКА

2026 Vol. 17 No. 2 569–584
http://journals.rudn.ru/semiotics-semantics

DOI: 10.22363/2313-2299-2026-17-2-569-584
EDN: LTBNXB

Research Article / Научная статья

Transformation-Based Generation of English Language Exercises: a Deterministic and CEFR-Guided Approach

Marianna V. Popova

Moscow State Linguistic University, *Moscow, Russian Federation*
✉neunerin@gmail.com

Abstract. The developed model provides a deterministic, transformation-based approach to generating English language exercises from user-provided texts. The system is designed as a rule-based AI tool that restructures existing textual material into pedagogically organized tasks while preserving structural and contextual coherence. Unlike large probabilistic language models that autonomously generate discourse, the proposed approach relies on explicit linguistic rules and CEFR-calibrated constraints (A1–C2), ensuring reproducibility and transparency of output. The study conceptualizes AI-assisted exercise generation as a structured process of text transformation rather than autonomous content production. The model maintains traceability between source text elements and generated tasks and avoids probabilistic text synthesis. CEFR integration functions as a constraint-based filtering mechanism that regulates lexical and grammatical selection according to proficiency descriptors. A qualitative pilot study conducted in a tertiary EFL context demonstrates that the system produces linguistically valid and pedagogically usable exercises derived from authentic texts. The findings indicate that rule-based AI systems can support instructional material preparation while preserving teacher control and structural stability. The study contributes to ongoing discussions in applied linguistics on transparent and theory-informed integration of artificial intelligence into foreign language education.

Keywords: artificial intelligence in language education, exercise generation, CEFR, rule-based NLP, English language teaching, applied linguistics

Conflict of interests: The author declared no conflict of interest.

Acknowledgement: The author thanks Liudmila Rybalko for her contribution to the development of a program implementation system for generating exercises.

Article history: Received: 23.12.2025; Accepted: 15.02.2026.

For citation: Popova, M.V. (2026). Transformation-Based Generation of English Language Exercises: A Deterministic and CEFR-Guided Approach. *RUDN Journal of Language Studies, Semiotics and Semantics, 17*(2), 569–584. https://doi.org/10.22363/2313-2299-2026-17-2-569-584 EDN LTBNXB

© Popova M.V., 2026
This work is licensed under a Creative Commons Attribution-NonCommercial 4.0 International License
https://creativecommons.org/licenses/by-nc/4.0/legalcode

Генерация упражнений по английскому языку на основе текстовой трансформации: детерминированная модель с учетом уровней CEFR

М.В. Попова

Московский государственный лингвистический университет, *Москва, Российская Федерация*
neunerin@gmail.com

Аннотация. Разработанная модель генерации упражнений по английскому языку на основе загружаемого текста реализована в виде системы искусственного интеллекта, основанной на формализованных правилах и осуществляющей структурированную переработку исходного текстового материала в педагогически организованные задания с сохранением его структурной и контекстуальной целостности. В отличие от вероятностных генеративных языковых моделей, предлагаемый подход основан на явных лингвистических правилах и ограничениях, калиброванных в соответствии с уровнями владения языком по шкале CEFR (A1–C2), что обеспечивает воспроизводимость и прозрачность результатов. Генерация упражнений рассматривается как процесс управляемой текстовой трансформации, а не автономного порождения дискурса. Модель сохраняет трассируемость между элементами исходного текста и производными заданиями и не использует вероятностный синтез нового дискурса. Интеграция параметров CEFR функционирует как механизм фильтрации, регулирующий лексический и грамматический отбор в соответствии с дескрипторами уровня владения языком. Качественное пилотное исследование, проведенное в условиях вузовского преподавания английского языка как иностранного, показало, что система обеспечивает лингвистическую корректность и педагогическую применимость генерируемых заданий. Результаты свидетельствуют о том, что системы ИИ на основе формализованных правил могут эффективно поддерживать разработку учебных материалов, сохраняя контроль преподавателя и структурную устойчивость учебного дискурса. Работа способствует развитию теоретически обоснованных и прозрачных моделей интеграции искусственного интеллекта в прикладную лингвистику и методику преподавания иностранных языков.

Ключевые слова: искусственный интеллект в обучении иностранным языкам, генерация упражнений, CEFR, обработка текста на основе формализованных правил, преподавание английского языка, прикладная лингвистика

Конфликт интересов: автор заявил об отсутствии конфликта интересов.

Благодарности: Автор выражает благодарность Людмиле Рыбалко за вклад в разработку программной реализации системы генерации упражнений.

История статьи: дата поступления: 23.12.2025; дата приема в печать: 15.02.2026.

Для цитирования: *Popova M.V.* Transformation-Based Generation of English Language Exercises: A Deterministic and CEFR-Guided Approach // Вестник Российского университета дружбы народов. Серия: Теория языка. Семиотика. Семантика. 2026. Т. 17. № 2. С. 569–584. https://doi.org/10.22363/2313-2299-2026-17-2-569-584 EDN LTBNXB

Introduction

Artificial intelligence has emerged as a central methodological and epistemological issue in contemporary language education. While recent advances in large language models have enabled autonomous text generation at unprecedented scale, their integration into foreign language pedagogy raises fundamental questions regarding interpretability, pedagogical control, and construct validity [1; 2]. Student perceptions of AI in foreign language learning contexts frequently reflect concerns regarding transparency, reliability, and pedagogical trust [3]. Within applied linguistics, the discussion increasingly distinguishes between generative AI systems that synthesize new discourse and controlled computational tools designed to support instructional design [4; 5].

From a computer-assisted language learning (CALL) perspective, AI technologies must be evaluated not solely in terms of technical sophistication but in relation to their alignment with second language acquisition (SLA) theory and task design principles [6; 7]. Language pedagogy is grounded in structured input, meaningful task engagement, and construct-valid assessment practices [8; 9]. Consequently, fully autonomous probabilistic generation may conflict with pedagogical requirements of traceability, contextual integrity, and proficiency calibration.

In this context, the present study proposes a transformation-based, deterministic model for generating English language exercises from teacher-selected texts. Instead of synthesizing new discourse, the system restructures existing textual material into controlled pedagogical formats governed by explicit linguistic rules and CEFR-calibrated constraints (A1–C2). The model conceptualizes AI-assisted exercise generation as a structured process of secondary textual transformation embedded within established proficiency frameworks.

The study pursues three objectives: (1) to describe the system architecture, (2) to position it within contemporary AI-supported language education research, and (3) to evaluate its pedagogical viability through a qualitative pilot implementation in a tertiary EFL context.

By foregrounding interpretability, reproducibility, and proficiency-sensitive filtering, the study contributes to ongoing discussions in applied linguistics concerning theory-informed and pedagogically responsible AI integration.

Theoretical Framework

AI in Foreign Language Education

The rapid development of artificial intelligence technologies has significantly reshaped the landscape of foreign language education. Recent advances in large language models (LLMs) have enabled autonomous text generation, adaptive feedback, and real-time conversational interaction, expanding the range of AI-supported instructional tools [1; 2]. However, the integration of such systems

into pedagogical contexts raises theoretical and methodological concerns regarding transparency, controllability, and construct alignment.

Contemporary research distinguishes between generative, adaptive, and assistive AI systems in education [4; 5]. Generative systems prioritize autonomous content production; adaptive systems focus on personalization through data-driven modeling; assistive systems support specific instructional functions without replacing pedagogical decision-making. While generative AI offers flexibility and scalability, its probabilistic architecture may introduce opacity, unpredictability, and reduced structural correspondence of output — factors that complicate pedagogical validation [2].

Within applied linguistics, these concerns intersect with broader theoretical questions about the relationship between technological mediation and language acquisition processes. From a CALL perspective, technological tools must be evaluated not solely in terms of technical sophistication but in relation to established principles of second language acquisition and instructional design [6].

Alignment with SLA and Task-Based Language Teaching

Second language acquisition research emphasizes the importance of structured input, meaningful task engagement, and form-meaning mapping in the development of communicative competence [7]. Task-based language teaching conceptualizes pedagogical tasks as goal-oriented activities grounded in authentic discourse and designed to promote language processing within meaningful contexts.

Unconstrained AI text generation may risk producing linguistically correct yet pedagogically misaligned content, particularly when task formats lack explicit linkage to input discourse. In contrast, rule-governed transformation architectures that derive exercises directly from teacher-selected texts maintain alignment between input, processing demands, and output formats. Such alignment reflects principles of instructional coherence central to task-based pedagogy [7].

From a vocabulary acquisition perspective, systematic lexical selection and frequency control are essential components of effective instructional design [9; 10]. AI systems that incorporate structured lexical filtering mechanisms may therefore better support vocabulary development than systems relying on unconstrained generative output.

Construct Validity and Assessment-Oriented Task Design

Exercise generation in language education cannot be considered independently of assessment theory. According to Bachman and Palmer (1996) [8], the validity of language tasks depends on construct representation, authenticity, and the relationship between input and expected response behavior. Selected-response formats, cloze procedures, and controlled production tasks each target distinct aspects of language ability and must be grounded in clearly defined constructs.

From this perspective, AI-mediated task generation should preserve construct validity and ensure that response formats are meaningfully derived from the linguistic features of the source text. The preservation of textual anchoring enhances authenticity and reduces construct-irrelevant variance, particularly in vocabulary-focused and comprehension-oriented tasks [11].

Thus, the theoretical requirement is not merely technological feasibility but principled alignment between linguistic analysis, task format, and proficiency descriptors.

CEFR, Readability, and Proficiency-Sensitive Filtering

The Common European Framework of Reference for Languages (CEFR) provides internationally recognized descriptors of communicative competence across proficiency levels[1]. In computational language learning systems, CEFR alignment is often operationalized through readability classification and lexical profiling [12].

Research in NLP-informed CALL demonstrates that proficiency-sensitive filtering mechanisms can be implemented through structured lexical resources and readability modeling [13]. Such systems prioritize controlled selection of linguistic units rather than probabilistic generation. Early ICALL architectures based on modular NLP component reuse further established the viability of linguistically grounded exercise construction [14].

These approaches highlight an important distinction: proficiency calibration may function either as adaptive personalization driven by statistical modeling or as constraint-based filtering embedded within deterministic rule systems. The latter approach emphasizes reproducibility, interpretability, and architectural explicability.

Toward Deterministic, Linguistically Grounded AI Architectures

Recent debates in AI-supported education increasingly stress the need for explainable, controllable, and pedagogically accountable systems [1; 2]. While generative AI demonstrates impressive fluency, its stochastic nature complicates reproducibility and instructional validation.

Deterministic transformation-based models represent an alternative paradigm in which AI functions as a structured processing mechanism rather than an autonomous discourse generator. By integrating linguistic preprocessing, CEFR-calibrated filtering, and template-based instantiation, such systems maintain input-task alignment. This architecture aligns with SLA theory, task-based pedagogy, and assessment validity principles, thereby positioning deterministic AI as a theoretically grounded complement to generative systems.

[1] Council of Europe. (2001). *Common European Framework of Reference for Languages: Learning, Teaching, Assessment*. Cambridge: Cambridge University Press.

The present study builds upon this theoretical foundation by proposing a controlled text-transformation system that operationalizes proficiency-sensitive filtering within a rule-governed, reproducible pipeline.

Model Description

General Architecture and Design Principles

The proposed system is an AI-assisted tool for generating English language learning exercises from user-provided texts. The model is implemented as an open-access Python-based software repository developed by the author in cooperation with L. Rybalko and made available for research and educational purposes[2].

Technically, the system is written in Python and relies on standard NLP procedures for structural text analysis. The processing pipeline includes automated tokenization, sentence segmentation, part-of-speech tagging, and integration of CEFR-labelled lexical resources. These structural annotations serve as the basis for rule-based pedagogical transformations. Internally, the system operates as a sequence of rule-governed transformation modules acting on explicitly defined linguistic units (tokens, sentences, grammatical categories).

Unlike large-scale end-to-end generative language models that rely on probabilistic text synthesis, the present system follows a rule-based, text-driven architecture. It does not autonomously generate new discourse. Instead, it performs structured transformation of existing textual material into pedagogically organized tasks.

The generation mechanism consists of interconnected components, demonstrated in Figure 1.

1) **Linguistic preprocessing module** performs segmentation, tokenization, lemmatization, part-of-speech tagging, and statistical feature extraction. This stage identifies structurally transformable linguistic units and generates morphosyntactic annotations required for downstream processing.

2) **Vectorization and feature extraction module** converts selected textual units into structured numerical representations and extracts distributional and structural features that support controlled target selection and distractor generation.

3) **CEFR — constrained filtering module** applies proficiency-level constraints (A1–C2), restricting lexical and grammatical elements according to CEFR descriptors and calibrated lexical resources.

4) **Rule — based target unit selection module** identifies candidate lexical or grammatical elements suitable for pedagogical transformation based on explicitly defined linguistic rules.

5) **Template — based exercise generation module** embeds selected units into predefined task structures, ensuring structural consistency and contextual anchoring.

[2] Rybalko, L., & Popova, M.V. (2025). *English Exercise Generator* [Computer software]. GitHub. URL: https://github.com/ludryb/generate_exercises/tree/main (accessed: 10.11.2025).

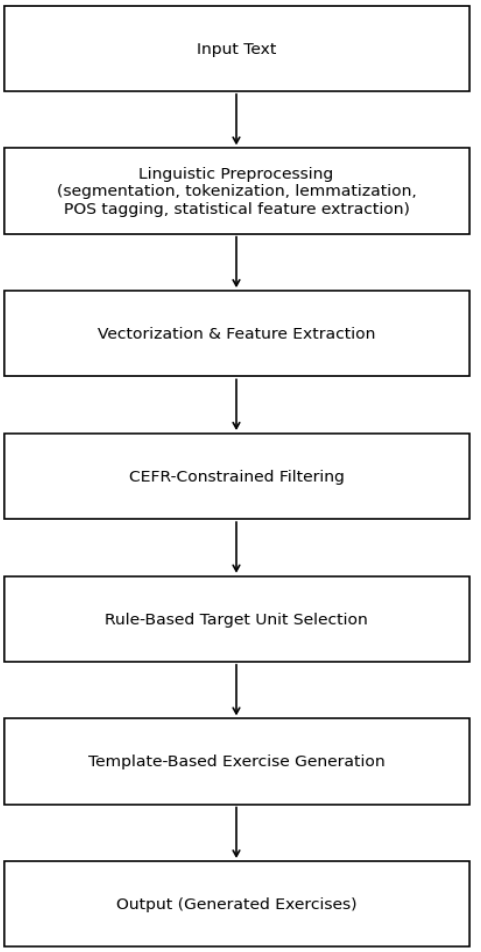


Figure 1. Pipeline architecture of the CEFR-guided rule-based exercise generation system
Source: Compiled by Marianna V. Popova.

6) **Output module** produces fully instantiated exercises ready for instructional use.

A distinctive feature of the model is the integration of CEFR standards. The system incorporates proficiency-level parameters ranging from A1 to C2. During exercise generation, the instructor may specify the intended CEFR level, which influences:

- selection of lexical items appropriate to the target level,
- filtering of grammatical constructions,
- control of task complexity and structural density.

The CEFR parameterization does not function as probabilistic adaptation but as a constraint-based filtering mechanism embedded within the transformation rules. Proficiency-sensitive filtering mechanisms are closely related to readability classification research in NLP-informed language learning systems [12]. Early ICALL systems implementing rule-based sentence selection and CEFR-calibrated readability assessment have demonstrated the pedagogical viability of linguistically

controlled filtering [13]. The present model extends this logic from sentence selection to structured exercise generation.

The CEFR-calibrated filtering module relies on externally available CEFR-labelled lexical datasets. Specifically, the model integrates lexical frequency and level information derived from publicly accessible CEFR-annotated word lists, including the "10,000 English Words CEFR Labelled" dataset[3] and the UniversalCEFR English lexical dataset[4]. These resources provide level-based lexical categorization (A1–C2), which is incorporated into the rule-based filtering mechanism to constrain lexical selection according to proficiency parameters. The model incorporates CEFR-labelled lexical resources as structured reference datasets within its deterministic filtering layer.

All transformation operations are reproducible: identical input under identical CEFR parameters yields identical output. This architectural decision ensures output stability, procedural reproducibility, and pedagogical reliability.

The system is designed according to three core principles:

1. **Text-dependence** — all exercises are derived exclusively from the submitted source text.

2. **Textual traceability** — each generated task can be linked to identifiable lexical or syntactic units in the original discourse.

3. **Pedagogical controllability** — instructors retain full authority to review, interpret, and integrate outputs into instructional contexts.

The model may be understood as a controlled mechanism of structured transformation regulated not only by linguistic form but also by proficiency-based constraints. The integration of CEFR descriptors situates the system within established frameworks of communicative competence, reinforcing its alignment with contemporary language pedagogy.

This configuration positions the system as an architecturally explicit AI-mediated processing tool that combines structural linguistic analysis with proficiency-calibrated pedagogical transformation. All processing is executed locally without reliance on external large language model APIs.

Exercise Types and Generation Logic

The current implementation supports four pedagogically established task formats widely used in foreign language instruction and assessment. Each task type is derived through rule-governed transformations applied to the source text, ensuring alignment with communicative, lexical, and form-focused learning objectives [11; 16].

[3] Nezahatkk. (n.d.). *10,000 English words CEFR labelled* [Dataset]. Kaggle. URL:https://www.kaggle.com/datasets/nezahatkk/10-000-english-words-cerf-labelled (accessed: 10.11.2025).

[4] UniversalCEFR. (n.d.). *CEFR SP-EN dataset* [Dataset]. Hugging Face. URL: https://huggingface.co/datasets/UniversalCEFR/cefr_sp_en (accessed: 10.11.2025).

Selected-Response Vocabulary Items (Multiple-Choice Format)

The system generates selected-response vocabulary items in a multiple-choice format, a task type extensively discussed in contemporary language assessment research [11]. Target lexical units are extracted from the source text and presented as correct responses. Distractors are generated through controlled lexical variation procedures designed to maintain semantic proximity and grammatical plausibility while minimizing construct-irrelevant variance. From an assessment perspective, multiple-choice vocabulary items allow controlled measurement of receptive lexical knowledge when distractor quality is systematically regulated [8]. By restricting response options to contextually compatible material, the system preserves discourse integrity and avoids introducing content external to the input domain.

Cloze and Controlled Gap-Fill Tasks

Selected lexical or grammatical elements are systematically omitted from sentences based on predefined linguistic criteria (e.g., part of speech, grammatical function, CEFR level). The preservation of the original syntactic frame supports contextual inference and reinforces form-meaning mapping, consistent with usage-based perspectives on second language acquisition [7]. Such tasks constitute semi-controlled production activities in which retrieval occurs under discourse-level constraints instead of decontextualized rule application.

Meaning-Focused Comprehension Tasks (Drag-and-Drop Format)

The system also generates structured comprehension tasks implemented in a drag-and-drop format. These tasks correspond to meaning-focused reading activities within task-based language teaching frameworks [7]. By requiring learners to interpret sentence-level and discourse-level meaning, these tasks engage interpretive and discourse competencies as described in communicative competence models [8]. The preservation of textual cohesion ensures alignment with authenticity-oriented instructional approaches [15].

Lexical Puzzle Tasks (Word Reconstruction)

Finally, the system produces lexical reconstruction and word manipulation tasks designed to reinforce orthographic awareness and lexical retrieval. From a vocabulary pedagogy perspective, such tasks contribute to consolidation within balanced lexical instruction models [9; 16]. These exercises integrate attention to form within meaningful textual input, aligning with form-focused instruction embedded in communicative contexts [7].

Across all formats, the generation mechanism preserves source-text anchoring and structural coherence. Tasks are derived exclusively from the original discourse, ensuring referential integrity and contextual consistency. From an assessment perspective, this design reduces construct-irrelevant variance and supports alignment between task demands and the input domain [8]. From an SLA perspective, it integrates meaning-focused and form-focused strands of instruction within a unified rule-governed transformation framework [7; 16].

Workflow

The operational workflow of the system consists of the following sequential stages:

1. The user uploads an English text (.txt) and specifies the target CEFR level (A1–C2).
2. The system performs automated structural analysis of the input text, including tokenization, sentence segmentation, and part-of-speech tagging. This stage produces linguistically annotated textual units that serve as the basis for subsequent operations.
3. Following preprocessing, the system applies a CEFR-calibrated filtering layer. Lexical items are evaluated against integrated CEFR-labelled reference datasets, and selection constraints are imposed according to the specified proficiency level. This stage functions as a constraint-based selection mechanism and does not rely on predictive classification models.
4. Linguistically relevant elements (e.g., lexical items, grammatical constructions) are selected according to predefined transformation rules and proficiency constraints.
5. Selected textual units are embedded into predefined exercise templates (e.g., selected-response items, cloze tasks, reconstruction tasks), resulting in structured pedagogical outputs.
6. Generated exercises are presented for review and use.

Pedagogical Scope and Constraints

The model is intentionally limited to rule-based generation. It does not incorporate large-scale generative language models, probabilistic text synthesis, or adaptive learner profiling mechanisms.

These constraints serve three purposes:

- ensuring output stability,
- maintaining transparency of transformation procedures,
- preserving teacher control over instructional content.

The absence of autonomous generation reduces risks associated with semantic drift or contextually detached output, which are frequently discussed in relation to large language models [1; 2].

At present, difficulty differentiation remains instructor-mediated and does not involve algorithmic adaptation. The system is designed as a material preparation instrument and is not intended to serve as a personalized tutoring environment.

Cross-Linguistic Adaptability: Extension to German

Although the present implementation is designed for English language instruction, the architecture of the system is not language-specific. The modular NLP pipeline and constraint-based generation logic allow adaptation to other morphologically richer languages, including German. Such adaptation would

require replacing the English preprocessing components with a German language model supporting tokenization, lemmatization, and morphosyntactic tagging, as well as integrating CEFR-aligned German lexical resources.

In the case of German, additional rule-layer refinement would be necessary to account for grammatical gender, case marking, separable verb constructions, and agreement phenomena. The CEFR-calibrated filtering mechanism could be extended using German CEFR-annotated lexical datasets, e.g., CEFRLex German, ensuring proficiency-sensitive constraint-based selection without introducing probabilistic modeling.

Importantly, the deterministic transformation architecture would remain unchanged. The system's rule-governed design supports cross-linguistic scalability by isolating language-dependent resources within configurable preprocessing and filtering modules. This extensibility highlights the broader applicability of transformation-based exercise generation beyond English language pedagogy. A further development of the system could involve the implementation of a bilingual English–German version, enabling parallel constraint-based exercise generation across both languages within a unified, multilingual architectural framework.

Pilot Study

Research Design & Materials

The pilot implementation was designed as a qualitative exploratory study aimed at assessing the pedagogical feasibility and structural validity of the model. The objective was not to measure learning outcomes quantitatively but to examine the system's performance in authentic instructional conditions and assess the structural and pedagogical properties of the generated exercises.

The pilot involved **N = 5** university-level EFL instructors with teaching experience ranging from 3 to 15 years. All participants regularly design supplementary materials based on authentic English texts. Participation was voluntary.

A total of **N = 25** authentic English texts were used during the pilot phase. The texts ranged from 200 to 400 words and represented three genres:

- informational academic prose,
- journalistic articles,
- short narrative excerpts.

The selected materials covered CEFR levels B1 to C1. No artificial simplification or preprocessing was applied prior to submission.

Procedure and Evaluation Criteria

Each instructor submitted 5 texts to the system and specified a CEFR level parameter (A2–C1 depending on instructional context). The model generated:

• selected-response vocabulary items (multiple-choice format);

• cloze and controlled gap-fill tasks;

• meaning-focused comprehension tasks implemented in a drag-and-drop format;

• lexical reconstruction tasks.

Generated outputs were analyzed without prior manual modification in order to assess immediate usability. Instructor reflections were collected through structured written commentary focusing on three evaluation criteria mentioned in the model description section.

The analysis combined:

- qualitative thematic coding of instructor feedback,
- structural review of generated exercises,
- comparison of CEFR parameter settings with observed task complexity.

No learner data were collected during this phase. The pilot was restricted to instructor-mediated evaluation of discursive restructuring quality.

Findings

The results indicate that the system consistently preserved traceable linkage between source texts and generated exercises. Tasks remained lexically and syntactically anchored in the source texts. CEFR calibration effectively constrained lexical density and grammatical complexity. Instructors reported high usability in contexts requiring rapid preparation of text-based practice materials. However, variability of exercise formulation remained limited due to the deterministic template structure.

The findings confirm that the model functions as a controlled transformation mechanism and not as an autonomous content generator. The exploratory design of the pilot study limits generalizability, and further controlled learner-based validation is required to assess learning effectiveness.

Discussion

The findings of the pilot implementation allow the proposed system to be situated within a distinct transformation-based paradigm of AI-supported language education. As outlined in the theoretical framework, contemporary research differentiates between generative, adaptive, and assistive AI systems [4; 5]. While generative systems prioritize autonomous discourse production and adaptive systems emphasize personalization, the present model operates through structured transformation of teacher-selected input. This architectural choice directly addresses concerns regarding transparency and pedagogical control raised in recent discussions of large language models in education [1; 2]. By maintaining structural traceability between input text and generated task formats, the system preserves source-text anchoring and reproducibility — features that are central to CALL design principles [6]. In doing so, the study addresses concerns regarding opacity, reduced structural traceability, diminished pedagogical control, and weakened construct alignment raised in current AI-in-education debates. In this framework,

AI is not treated as an autonomous pedagogical agent but as a controlled linguistic processing mechanism integrated into instructional design principles.

From a second language acquisition perspective, pedagogical tasks must promote structured input processing and meaningful form–meaning integration [7]. The rule-governed transformation architecture supports this requirement by deriving all exercises directly from the source text, ensuring that learners engage with authentic discourse rather than decontextualized or externally generated content. The preservation of textual anchoring reinforces coherence between input domain and task demand. This alignment reduces the risk of construct-irrelevant variance and supports principled task construction. In contrast to unconstrained generative output, exercise generation maintains consistency between linguistic features of the text and cognitive processing required by the learner. Moreover, vocabulary-focused formats within the system rely on controlled lexical selection informed by frequency and proficiency descriptors. This design corresponds to established principles of vocabulary instruction emphasizing systematic exposure and level-appropriate lexical control [9; 10]. Thus, the system's pedagogical strength lies not in generative capacity but in its ability to operationalize SLA-informed task design through computational structuring.

Language tasks function as instruments for eliciting evidence of underlying language ability. According to Bachman and Palmer (1996) [8], task validity depends on coherent construct representation, authenticity, and the relationship between input and expected response behavior. The present model preserves construct alignment by ensuring that all selected-response items, cloze tasks, and comprehension activities are structurally derived from the input text. This reduces construct-irrelevant content and enhances authenticity by grounding tasks in meaningful discourse. Selected-response vocabulary items and controlled gap-fill tasks correspond to established assessment formats targeting lexical knowledge and grammatical processing [11]. AI-assisted assessment formats may also influence learner motivation, anxiety, and affective engagement, thereby shaping task effectiveness beyond purely cognitive dimensions [17]. Their computational instantiation within a deterministic framework ensures reproducibility—an important but often underemphasized dimension of assessment design in AI-mediated contexts. By embedding CEFR-based filtering within the generation pipeline, the system further aligns task difficulty with internationally recognized proficiency descriptors. This proficiency-sensitive constraint mechanism reflects research in readability classification and CEFR-based ICALL systems [12; 13].

The system extends the tradition of ICALL architectures that rely on modular NLP components and linguistically controlled selection [14]. Similar to corpus-driven extraction systems such as GDEX [18], the model prioritizes controlled extraction over unconstrained generation. However, it advances this logic by integrating CEFR-calibrated filtering with template-based exercise instantiation. So, the model exemplifies an alternative trajectory within AI-supported education: rather than replacing pedagogical structure with probabilistic fluency, it

operationalizes instructional design principles through deterministic computational procedures. This approach may be particularly relevant in higher education contexts where transparency, accountability, and alignment with standardized proficiency frameworks are critical.

The findings suggest that AI integration in language education does not necessitate autonomous generative capacity. Instead, linguistically grounded transformation architectures can provide scalable support for instructional material preparation while preserving teacher agency and theoretical coherence.

From an applied linguistics perspective, the model shows that computational systems can embody SLA theory, and assessment validity principles, moving beyond the replication of surface-level fluency. In this sense, AI functions as a mediating infrastructure that operationalizes established pedagogical constructs. Such architectures contribute to ongoing debates concerning interpretability, deterministic reproducibility, and responsible AI implementation in education [1; 2]. By foregrounding deterministic design and CEFR-constrained filtering, the study proposes a model of AI integration that prioritizes pedagogical structure over generative autonomy.

Conclusions

This study has presented a deterministic transformation model for generating CEFR-aligned English language exercises from teacher-selected texts. Within applied speech studies, the model demonstrates how computational architectures can restructure discourse while preserving referential and communicative integrity. The system was conceptualized not as an autonomous generative agent but as a linguistically grounded processing architecture that restructures existing discourse into pedagogically controlled task formats.

The theoretical positioning of the model draws on established principles of CALL research, task-based language teaching, vocabulary pedagogy, and construct validity in language assessment. Within this framework, AI-assisted exercise generation is understood as a structured transformation process governed by explicit linguistic rules and proficiency-sensitive constraints rather than probabilistic discourse synthesis.

The qualitative pilot implementation in a tertiary EFL context provides initial empirical support for the pedagogical viability of the system. By maintaining coherence between source input and generated formats, the model supports construct-consistent task design and reduces risks associated with opaque generative architectures.

From the perspective of applied linguistics, the study contributes to ongoing discussions on responsible and theory-informed AI integration in language education. It demonstrates that computational systems can operationalize SLA-informed pedagogical principles while preserving interpretability and teacher agency. The study therefore advocates for interpretable AI architectures that embed

linguistic theory within computational design instead of relying on probabilistic autonomy.

Future research should extend the present findings through controlled learner-based evaluation, quantitative analysis of task effectiveness, and comparison with generative AI systems across proficiency levels. Further development may also explore expanded task formats and integration with adaptive feedback mechanisms while preserving architectural interpretability.

References / Список литературы

1. Kim, E.J. (2025). AI-Assisted English Learning: A Tool for All or Only a Select Few? Exploring Learner Difficulties and Group-Specific Effects in Korean EFL Contexts. *Language Learning & Technology*, *29*(1), 1–22. https://doi.org/10.64152/10125/73633
2. Kebble, P. (2023). A Chat with ChatGPT: The Potential Impact of Generative AI in Higher Education Learning, Teaching and Assessment, with Specific Reference to EAL/D Students. *Journal of Academic Language and Learning, 17*(1), T81–T91.
3. Tikhonova, N.V., & Ilduganova, G.M. (2024). "I Am Afraid of How Fast Artificial Intelligence Is Developing": Students' Perceptions of AI in Foreign Language Learning. *Higher Education in Russia*, *33*(4), 63–83. (In Russ.). https://doi.org/10.31992/0869-3617-2024-33-4-63-83 EDN: FNUAVR
   Тихонова Н.В., Ильдуганова Г.М. «Меня пугает то, с какой скоростью развивается искусственный интеллект»: восприятие студентами искусственного интеллекта в обучении иностранным языкам // Высшее образование в России. 2024. Т. 33. № 4. С. 63–83. https://doi.org/10.31992/0869-3617-2024-33-4-63-83 EDN: FNUAVR
4. Pokrivčáková, S. (2019). Preparing Teachers for the Application of AI-Powered Technologies in Foreign Language Education. *Journal of Language and Cultural Education*, *7*(3), 135–153. https://doi.org/10.2478/jolace-2019-0025
5. Titova, S.V. (2024). AI-Based Technological Solutions in Foreign Language Education. *Lomonosov Linguistics and Intercultural Communication Journal, 27*(2), 18–37. (In Russ.). https://doi.org/10.55959/MSU-2074-1588-19-27-2-2. EDN: OWSQVG
   Титова С.В. Технологические решения на базе искусственного интеллекта в обучении иностранным языкам // Вестник Московского университета. Серия. 19. Лингвистика и межкультурная коммуникация. 2024. Т. 27. № 2. С. 18–37. https://doi.org/10.55959/MSU-2074-1588-19-27-2-2. EDN: OWSQVG
6. Chapelle, C.A. (2001). *Computer Applications in Second Language Acquisition: Foundations for Teaching, Testing and Research*. Cambridge: Cambridge University Press.
7. Ellis, R. (2003). *Task-Based Language Learning and Teaching*. Oxford: Oxford University Press.
8. Bachman, L.F., & Palmer, A.S. (1996). *Language Testing in Practice: Designing and Developing Useful Language Tests*. Oxford: Oxford University Press.
9. Nation, I.S.P. (2001). *Learning Vocabulary in Another Language*. Cambridge: Cambridge University Press.
10. Schmitt, N. (2008). Review Article: Instructed Second Language Vocabulary Learning. *Language Teaching Research*, *12*(3), 329–363.
11. Brown, J.D., & Abeywickrama, P. (2019). *Language Assessment: Principles and Classroom Practices*. Pearson.
12. Vajjala, S., & Meurers, D. (2012). On Improving the Accuracy of Readability Classification Using Insights from Second Language Acquisition. In: *Proceedings of the 7th Workshop on Innovative Use of NLP for Building Educational Applications (BEA7)* (pp. 163–173).
13. Pilán, I., Volodina, E., & Johansson, R. (2013). Automatic Selection of Suitable Sentences for Language Learning Exercises. In: *20 Years of EUROCALL: Learning from the Past, Looking*

to the Future. L. Bradley & S. Thouësny (eds.). Proceedings of the 2013 EUROCALL Conference, Évora, Portugal (pp. 218–225). Dublin–Voillans.
14. Volodina, E., Borin, L., Loftsson, H., Arnbjörnsdóttir, B., & Leifsson, G.Ö. (2012). Waste Not, Want Not: Towards a System Architecture for ICALL Based on NLP Component Re-Use. In: *Workshop on NLP in Computer-Assisted Language Learning. Linköping Electronic Conference Proceedings 80* (pp. 47–58). Linköping, Sweden.
15. Gilmore, A. (2007). Authentic Materials and Authenticity in Foreign Language Learning. *Language Teaching*, *40*(2), 97–118. https://doi.org/10.1017/S0261444807004144
16. Nation, I.S.P. (2007). The Four Strands. *Innovation in Language Learning and Teaching*, *1*(1), 2–13. https://doi.org/10.2167/illt039.0
17. Biju, N., et al. (2024). Which One? AI-Assisted Language Assessment or Paper Format: An Exploration of the Impacts on Foreign Language Anxiety, Learning Attitudes, Motivation, and Writing Performance. *Language Testing in Asia*, *14*(1), 45. https://doi.org/10.1186/s40468-024-00245-x
18. Kilgarriff, A., Husák, M., McAdam, K., Rundell, M., & Rychlý, P. (2008). GDEX: Automatically Finding Good Dictionary Examples in a Corpus. In: *Proceedings of Euralex 2008*.

Сведения об авторе:
Марианна Викторовна Попова, кандидат филологических наук, доцент, ведущий научный сотрудник Экспериментально-фонетической лаборатории криминалистики по речеведению, Московский государственный лингвистический университет (119034, Российская Федерация, г. Москва, ул. Остоженка, д. 38 с. 1); *сфера научных интересов*: психолингвистика; социолингвистика; интерференция и акцент; фонетика/фонология; когнитивная лингвистика; прикладная лингвистика; компьютерная лингвистика (NLP); германские языки; *e-mail*: neunerin@gmail.com
ORCID: 0000-0003-4302-7054; ResearcherID: AEA-4819-2022; SPIN-код: 9048-5511.

Information about the author:
Marianna V. Popova, PhD in Philology, Associate Professor, Leading Researcher at the Experimental Phonetics and Forensic Linguistics Laboratory, Moscow State Linguistic University (38 build 1 Ostozhenka str., Moscow, Russian Federation, 119034); *Research interests*: psycholinguistics; sociolinguistics; phonetics and phonology; language transfer and accent; cognitive linguistics; applied linguistics; computational linguistics (NLP); Germanic languages; *e-mail*: neunerin@gmail.com
ORCID: 0000-0003-4302-7054; ResearcherID: AEA-4819-2022; SPIN-code: 9048-5511.

RUDN Journal of Language Studies, Semiotics and Semantics
2313-2299 (Print), ISSN 2411-1236 (Online)
Вестник Российского университета дружбы народов.
Серия: ТЕОРИЯ ЯЗЫКА. СЕМИОТИКА. СЕМАНТИКА

2026 Vol. 17 No. 2 585–604
http://journals.rudn.ru/semiotics-semantics

DOI: 10.22363/2313-2299-2026-17-2-585-604
EDN: LTTOJA

Research Article / Научная статья

Machine vs Human: Large Language Model Errors as a Means of Creating Stylistically Appropriate Text

Anna V. Zabolotskikh[1] ✉, **Elena E. Sokolova**[2], **Darya V. Tavberidze**[1]

[1]RUDN University, *Moscow, Russian Federation*
[2]Moscow Institute of Physics and Technology (MIPT), *Dolgoprudny*, *Russian Federation*
✉zabolotskikh-av@rudn.ru

Abstract. The relationship between humans and machines in education is analyzing the characteristic features of the LLM artificial language as a tool for teaching the creation of scientific written text through a critical analysis of errors made in texts created by man and machine. This departure from traditional ideas about learning to write in a foreign language reflects progressive processes in teaching, since identifying parallels between human learning and model training based on mistakes refers to a high-tech approach focused on the student rather than the teacher. LLM, as a tool capable of correcting and creating generated text, acts as an incentive for the development of critical thinking when analyzing inaccuracies or evaluating the text "appropriatness". LLM's ability to generate different results can be used as a error-based methodology for teaching academic writing. In this regard, academic writing training is a critical analysis and editing of text by visualizing errors, creating a sense of competition between a machine and a human and stimulating mental activity. This seems to be a clear example of how inconsistency and ambiguity can be used for development and creation.

Keywords: LLM, critical thinking, text appropriatness, editing

Conflicts of interest: The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Author contributions: Anna V. Zabolotskikh, Elena E. Sokolova, & Darya V. Tavberidze — preparation of the article text.

Article history: Received: 20.01.2026; Accepted: 10.02.2026.

For citation: Zabolotskikh, A.V., Sokolova, E.E., & Tavberidze, D.V. (2026). Machine vs Human: Large Language Model Errors as a Means of Creating Stylistically Appropriate Text. *RUDN Journal of Language Studies, Semiotics and Semantics, 17*(2), 585–604. https://doi.org/10.22363/2313-2299-2026-17-2-585-604 EDN LTTOJA

© Zabolotskikh A.V., Sokolova E.E., Tavberidze D.V., 2026
This work is licensed under a Creative Commons Attribution-NonCommercial 4.0 International License
https://creativecommons.org/licenses/by-nc/4.0/legalcode

Человек или машина: ошибки больших языковых моделей как средство создания стилистически уместного текста

А.В. Заболотских[1], Е.Е. Соколова[2], Д.В. Тавберидзе[1]

[1] Российский университет дружбы народов, *Москва, Российская Федерация*
[2] Московский физико-технический институт (МФТИ), *Долгопрудный, Российская Федерация*
✉zabolotskikh-av@rudn.ru

Аннотация. Анализ характерных ошибок, допущенных в текстах, созданных человеком и машиной, рассматривается на примере искусственного языка LLM как инструмента, используемого для обучения созданию научного письменного текста. Данный отход от традиционных представлений об обучении письму на иностранном языке отражает прогрессивные процессы в преподавании, т.к. выявление параллелей между обучением человека и тренировкой модели, основанное на ошибках, относится к высокотехнологичному подходу, ориентированному на учащегося, а не преподавателя. LLM как инструмент, способный корректировать и генерировать текст, выступает стимулом развития критического мышления при анализе неточностей или оценке «уместности» текста. Способность LLM создавать различные типы текста может быть использована для обучения академическому письму в качестве методологии, ориентированной на преодоление ошибок. В этой связи обучение академическому письму представляет собой критический анализ и редактирование текста с помощью визуализации ошибок, создавая ощущение конкуренции между машиной и человеком и стимулируя мыслительную активность. Это является наглядным примером того, как непоследовательность и двусмысленность могут быть использованы в целях развития навыка академического письма.

Ключевые слова: LLM, критическое мышление, соответствие текста, редактирование

Вклад авторов: А.В. Заболотских, Е.Е. Соколова и Д.В. Тавберидзе — подготовка текста статьи.

Конфликт интересов: авторы заявили об отсутствии конфликтов интересов.

История статьи: дата поступления: 20.01.2026; дата приема в печать: 10.02.2026.

Для цитирования: *Zabolotskikh A.V., Sokolova E.E., Tavberidze D.V.* Machine vs Human: Large Language Model Errors as a Means of Creating Stylistically Appropriate Text // Вестник Российского университета дружбы народов. Серия: Теория языка. Семиотика. Семантика. 2026. Т. 17. № 2. С. 585–604. https://doi.org/10.22363/2313-2299-2026-17-2-585-604 EDN LTTOJA

Introduction

More and more articles devoted to the usage of Large Language Models (LLMs) for development critical thinking skills in education are overwhelming academia and exciting the minds of researchers. The advancement of LLMs such as Chat GPT 4, Gemini Ultra, Llama 3.1 (70B) [1. P. 1–75], etc. provide a unique opportunity to reconsider the methodological approach to teaching writing in higher education [2]. The user interaction with LLMs is usually performed as a series of natural language queries, which are answered by the model [3]. Internally, those systems use the transformer neural network architecture to predict the next words/tokens that are most likely to follow in such a dialogue format [4]. Thus, the use of these technologies in the methodology of teaching academic writing has

begun the avalanche of enthusiastic research [5–10]. Furthermore, it is argued that the emergence of intelligent text generators may represent 'a breakthrough in written communication format since the invention of the word processor' (It is the biggest transformation of the writing process since the word processor) [11. P. 691]. Without detracting from LLMs merits as tools for creating and proofreading texts, we will try to consider the limitations which appeal to the integration of LLMs in assisting the learning process.

Terminological Inexactitude

The emergence of transformers which are actively used in natural language processing (NLP) and speech processing has introduced a certain 'terminological inexactitude' [12. P. 474] and confused the essence of concepts that refer to them. Primarily, LLMs were called an 'artificial intelligence' as part of an advertising campaign, which in fact was an artifice. This provoked the mass release of scientific articles where the authors draw direct analogies and compare the choice of language means of LLMs with the natural language used by humans, i.e. 'humanise' them [13]. From our perspective, this hobbles the efforts in understanding how humans should perceive the artificial language, what place should be offered to LLMs, and how we can define their scope and participation in higher education. The given paper considers LLMs as a tool in a pragmatically-orientated approach to teaching the scientific discourse generation, as the use of deep learning technology and multi-word sequencing of natural language has radically improved [14]. However, the tool which is widely used for writing abstracts for new preprints can produce inconsistent results: most abstracts contain false information, do not correspond to queries written in prompts, i.e. contain hallucinations or errors [15. P. 1–38; 16].

Use Errors in Your Favour

Error analysis has long been considered an effective technique in the field of English language teaching. Feedback on errors in recognizing and correcting mispronunciations for the learners was described as a powerful means to improve phonetics [17. P. 1–21). The support system may contribute to learning vocabulary by generating images corresponding to incorrect answers suggesting errors to the learner [18]. Error-driven teaching stems from the idea that the teacher's impression of an output being an authentic representation of input is false [19]. In contrast, error analysis enables teachers to provide counterexamples encouraging reflection and understanding. Since errors are often seen as token-based being embedded in other errors [20] it seems possible to involve LLM in generating examples for error analysis.

Departure from Traditions

We need, therefore, to base our reasons on the previous teaching methods which has changed the educational landscape from a low-tech approach to learning to a high-tech one, as this departure from traditions reflects the dynamic processes

in teaching, especially with the introduction of LLMs in it. Drawing parallels between teaching a human and training a model, we can observe similarity in both cases. LLMs' reinforced learning with rewards and punishment can be attributed to a teacher-centred approach (a low-tech one), and error-driven learning applied for LLMs — to a student-centered approach (a high-tech one) [21]. The difference is also obvious. A human is capable of analysing rules, recognizing patterns, and identifying mistakes from just a few examples, whereas LLMs require an extensive corpus of texts usually created by huge teams of workers engaged in generating synthetic examples.

Contrary to popular belief, that the quality of learning outcomes based on generated texts depend on human-written prompts [22], we suggest that developing effective prompts can be a challenge, leading to potentially inaccurate or contextually inadequate responses. This emphasises the importance of understanding the mechanisms of developing prompts to achieve accurate results.

Problem Statement

Most of the works, which are devoted to the academic opportunities and challenges of using LLMs [23] describe examples of solving practical problems of teaching writing without theoretical framework which might provide deep insights into the potential risks of using LLMs and attributing intelligence to them.

Furthermore, understanding of this problem reveals a certain complexity due to its interdisciplinary context. At present, for language instructors who teach academic writing in higher education it is not enough to possess only linguistic knowledge of syntactic structures, generative grammar, and natural language semantics developed by the pioneers in these fields Noam Chomsky and Richard Montague ("English as a formal language"). Without understanding, at least in general terms, how large language models are built, it is impossible to interact effectively with LLMs in the learning process.

Considering the outlined challenges, the given article describes the approach to the integration of LLMs into English Academic Writing course for postgraduate students at MIPT and RUDN with the emphasis on teaching writing the basic elements of the Introduction section: research question, objectives, gap, hook, hypothesis, criteria for problem selection and a key message.

The article has the following structure: We begin with revealing a theoretical rationale for language description approaches to understand natural language (NL) and artificial language (AL) crossing points and distinctions to show that IT NL operates on different laws than Transformers. Then we try to determine the place and role of LLMs in education, as well as establish criteria for AL appropriateness in teaching scientific communication. Finally, we specify artificial text generation outcomes by several LLMs as a response to the same prompt to identify language adequacy as an example of an error-driven teaching method. The authors of the article believe that LLM can be utilised as an auxiliary tool to help teach pragmatics

of scientific discourse, since many pragmatics learning functions are aimed at performing meaningful tasks that mimic students' real needs.

Background of the Problem

Attempts to relate language to logical analysis had been made long before the emergence of LLMs. The origin can be attributed to Aristotle's work 'First Analytics' in which much of his ideas can be interpreted as the analysis of language [24]. His thoughts gave rise to a large number of works that can be aligned as 'logic and language' (Husserl, 1900–1901, Wittgenstein, 1921, Crane, 1928, Ingarden, 1931, Peirce, F. De Saussure, etc.).

In the twentieth century, logic, driven by many divisions, regenerated and acquired a mathematical form, on which a new approach to understanding linguistics was formed. At that time, artificial languages of symbolic logic were perceived by their founders as ideal compared to natural languages. Theories using mathematical reasoning were built on artificial languages, as many scientists were mathematical logicians. In symbolism, the main approach to language analysis was the correctness of word usage in order to avoid ambiguity while interpreting scientific facts [25]. However, the further practical use of artificial language as the spoken language reflected the impossibility of such use, due to the lack of statics in language in the process of its development [26].

The Prague and Copenhagen Schools, based on the synchronous analysis in linguistics introduced by F. de Saussure, described the formal relations in language at a certain point in time [27]. Later, the 'elimination of ambiguity and uncertainty' between formalised and natural language was also discussed by Church [28. P. 106]. The American school of descriptive structural linguistics organized by L. Bloomfield (1887–1949) described the grammatical structure of a language as a formal structure, due to which similarities and differences between artificial language and natural language became obvious [29; 30].

A considerable amount of more recent research in the framework of logical analysis of language provided rational for the description of an artificial language by means of natural language and by a logical apparatus [31–33], etc. to name just a few.

In philosophical logic, in his 'Logic and Conversation', Grice divided the logicians into formalists and informalists, and described the existence of a divergence between formal symbols and their analogy in natural language. Formalists, believing that ideal language was the basis for the clear expression of scientific thought, described natural language by a system of formulae, and they considered the presence of elements that cannot be described, even by more complex formulae, in natural language as 'undesirable excrescences' that had no place in the reliable expression of scientific thought [34]. Their opponents rejected even the possibility of measuring the adequacy of language and its ability to produce clear definitions just to please scientific thought. This simplified and

symbolized logic of the formalists should not displace the logic of natural languages, but should reinforce it.

It is in these broad strokes of generalisation, that the authors of this article see the difference between approaches to describing natural and artificial languages. To be more precise, natural language is described within the framework of linguistics, where the primary importance is assigned to the study of speech units and language units, whereas in artificial language, data structures, algorithms, and logic are considered. In other words, artificial language is limited to the creation of specific models based on a specific set of data within the framework of mathematical logic. As defined by the Encyclopaedia of Database Systems, "a language model assigns a probability to a piece of unseen text, based on some training data" [35].

Rehash of a Hundred-Year-Old Idea

In one of his implicatures, in 'Logic and Conversation', Grice gives an example of a dialogue between two colleagues about a mutual friend who works in a bank: "*A asks B how C is getting on in his job, and B replies, Oh quite well, I think; he likes his colleagues, ~no! he hasn't been to prison yet. At this point, A might well inquire what B was implying, what he was suggesting, or even what he meant by saying that C had not yet been to prison. The answer might be any one of such things as that C is the sort of person likely to yield to the temptation provided by his occupation, that C's colleagues are really very unpleasant and treacherous people, and so forth. It might, of course, be quite unnecessary for A to make such an inquiry of B, the answer to it being, in the context, clear in advance*" [36. P. 43]. In a question asked in natural language, we can ask for additional information to clarify the communicative intention of the speaker, or perceive it from the context.

A similar experiment was conducted in Apple company to test the accuracy of a number of large language models. A sentence was added to the mathematical task's condition hypothetically having a similar meaning, but not affecting the logic of reasoning in order to get the correct answer. The results demonstrated a significant decrease in the performance of the tested models, as this addition violated the logic of phrase construction and did not correlate with the pattern embedded in the model itself. The authors' conclusion reflected the idea that the models do not understand either linguistic or mathematical categories, but operate with data 'without truly understanding their meaning' [37. P. 10].

Participants and Procedure

Participants of the study were the 1st-year PhD students at MIPT with the level of English not lower than B2 according to the CEFR levels. No selection criteria were set to the gender, age, and field of study; however, the prior theoretical knowledge of the learners should include the understanding of the academic style. The general requirements for writing a text were given as a theoretical background to the students (see the Supplement). The emphasis was put, in particular, on the

understanding of the appropriacy of the text, which is based on lexico-semantic, lexico-grammatical, and stylistical levels.

Research Methodology

Recognizing the existing contradictions and accepting the limitations of LLMs, we cannot deny that they were created to solve specific problems, in other words, they were based on a pragmatic approach. Moving to teaching pragmatics of scientific discourse, it is necessary to define what do we consider by appropriateness by which a text can be evaluated. In contemporary literature both human and machine evaluations and opinions may be perceived as equally trustworthy by learners [38. P. 1766–1776]. However, text appropriateness in terms of communication remains an open question [37].

To define the appropriateness of the text generated by LLMs, we need clear and conventional criteria which will correspond to proposed methodology. Taking into consideration that any text assessment is subjective, we see stylistic level as the most relevant for such a comparison. Stylistic peculiarities which include grammar and lexical aspects are well-established in scientific literature and may be considered as a standard for text appropriateness measurement [39; 40].

This research does not embrace the whole IMRAD structure and the paragraph length since a generated text is analyzed for stylistic appropriateness without considering the stage of planning. For LLM editor review assessment we will focus on the following criteria: logical coherence; the structure of the sentence; grammatical adequacy and lexis. This choice stems from the lack of substantial disagreement in science regarding academic text structure [41].

Coherence includes conventional logical indicators and connective elements. Clarity is the most significant prerequisite for good academic writing therefore the standard model of the English sentence (Subject-Predicate-Object) presupposes close connection of these structural elements without any insertions to avoid ambiguity. The 3d person singular and other impersonal components contribute to academic discourse objectiveness. On the lexical level formal writing is marked by colloquial expressions, slang, irrelevant abbreviations and the majority of phrasal verbs [41–43]. However, some researchers encourage avoiding professional jargon and pretentious vocabulary in scientific works to enhance understanding [44; 45], while others urge to the formal style usage [46; 47]. After selecting criteria for text appropriateness, we created the following algorithm for LLM efficiency testing which is presented in Figure 1.

For comparative analysis the students were suggested formulating the main content elements of the Introduction section, which included *Purpose, Objectives, Research question, Key message, Gap, Hook, Hypothesis, Practical value of the research.* Then, as a result of further discussion, a prompt was written stating guidelines for further text generation:

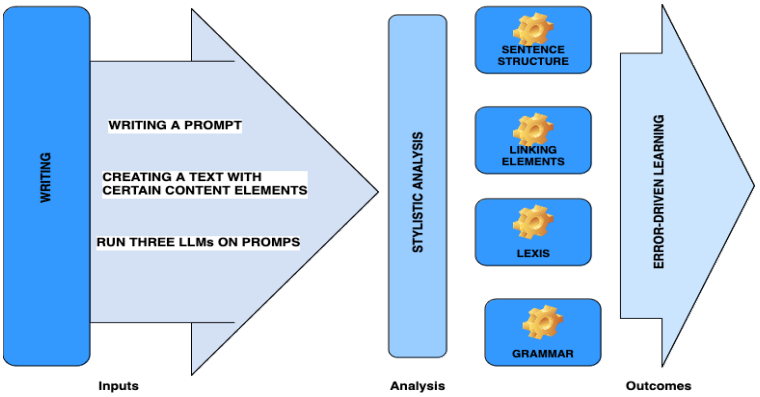


Figure 1. The proposed algorithm for LLMs efficiency testing
Source: compiled by Anna V. Zabolotskikh, Elena E. Sokolova, Daria V. Tavberidze

Please proofread and edit the following text for language issues and stylistic inconsistencies. Focus on language; don't try to improve the argumentation or add and remove content unless doing so is necessary to resolve language and style issues. The text should be written in American English, in 7th edition APA Style. Present the edited text followed by a list of all the changes you made, with clear explanations of why you made each change.

An important condition of the query was the justification for the proposed changes. We chose three models based on HELM leaderboard, where the main win rates together with accuracy, and efficiency of models' performance across different scenarios and metrics were presented. Two of them, GPT– 4o and Gemini Ultra are closed models and Llama 3.1 (70 B) is an open model**.**

Three LLMs were fed with PhD students' "human-written" abstracts and a proposed prompt.

After receiving a response to the inquiry from LLMs, the PhD students analyzed the adjustments proposed by LLMs in accordance with the text appropriateness using guidelines.

Task 2. Evaluate the generated text according to the guidelines and characterize if the generated text corresponds to academic standards.

Task 3. Select or reject the proposed improvements

Task 4. Reorganize the text according to the identified errors.

Further we show the examples of errors identified by LLMs in the human-written abstracts written by PhD students.

Findings and Discussion

To generate material for error critical analysis we run human-written texts on LLM models. LLMs' inconsistency was apparent to the learners and caused constructive criticism. Further we provide the examples of errors detected by different LLMs in human-written texts and students' criticism after their comparative analysis.

Example I. Source human-written text: *With the recent development of superconducting technology, researchers began using artificial superconducting quantum two-level systems (qubits) as an environment for quantum optics experiments. One of the experiments is quantum wave mixing. In a nutshell, when two coherent microwaves with frequencies close to the qubit resonance get mixed by it, the spectrum of the scattered field appears to be linear with amplitudes of the peaks corresponding to the photon number statistics in the incident coherent states of the signals. Provided that two arbitrary quantum states of light get mixed by a qubit, the resulting spectrum might demonstrate a similar correspondence to the photon number statistics of the scattered signals. Suppose we mix a superposition of a photon and a vacuum with a coherent microwave by a qubit. Will the amplitudes of the peaks in the resulting spectrum reveal the properties of the quantum superposition? Spectra in the experiment of mixing a superposition with a coherent microwave clearly correspond to the prepared superposition properties. However, we could not eliminate all spurious effects. Moreover, there is no accurate theoretical description for the scattering spectra in this particular experiment yet.*

LLMs generated results:

Gemini Ultra: Changed "Provided that" to "When" for simplicity and clarity. Changed "get mixed" to "are mixed" for grammatical accuracy.

GPT 4-o: Changed "Provided that" to "If": This substitution for a simpler conditional phrase improves readability and clarity.

Llama 3.1 (70B): Changed "Provided that" to "Given that" for a more direct and natural-sounding transition.

Changed "In a nutshell" → "In essence" (idiomatic expression) I replaced this with a more formal idiom, "in essence," to maintain a consistent tone.

Changed "Two coherent microwaves with frequencies close to the qubit resonance get mixed by it" → "two coherent microwaves with frequencies close to the qubit's resonance interact" (subject-verb agreement and possessive adjective) I changed the verb from "get mixed" to "interact" and added a possessive adjective ("qubit's") to make the sentence more specific.

Students' constructive criticism: [As seen] in Figure 1, electronic tools assess the conjunction provided that in a different way, which makes the writing strategy not clear. The question for reflection, which arises: should academic writing be marked by formalized pretentious vocabulary or, vice versa, prosaic words to enhance better understanding? The genitive (possessive) is not generally used with non-human subjects, however in some cases the genitive is used with inanimate things for specificity.

Gemini Ultra: "Spectra in the experiment" to "Spectra from the experiment": Changed the preposition to "from" to indicate the source of the spectra (see Figure 4).

GPT 4-o: "Spectra in the experiment of mixing a superposition" to "The spectra from the experiment mixing a superposition": This rephrasing improves sentence structure and clarity (see Figure 2).

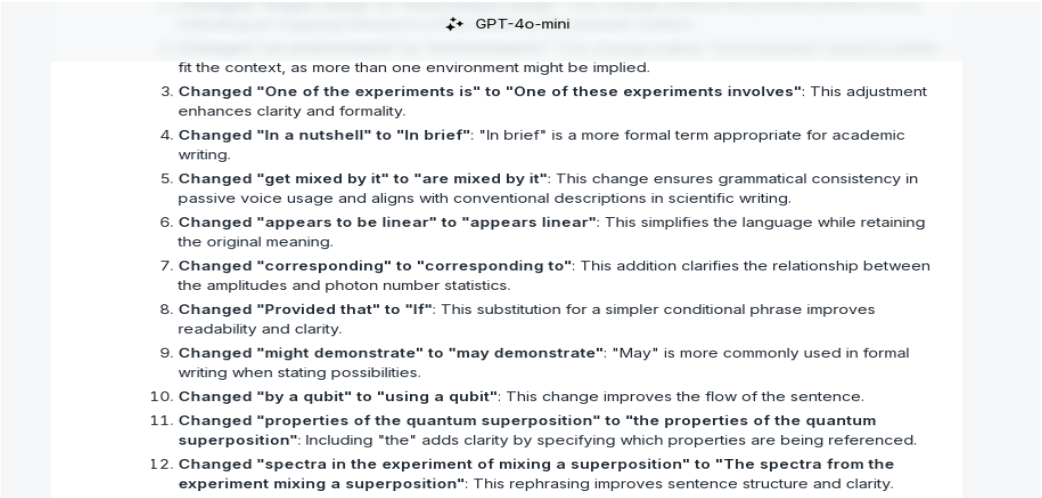


Figure 2. Chat GPT 4: corrections (answer to the prompt)
Source: compiled by Anna V. Zabolotskikh, Elena E. Sokolova, Daria V. Tavberidze

"researchers began using" → "researchers have begun using": Changed "began" to "have begun" for grammatical correctness, indicating the ongoing relevance of the action.

"correspond to the prepared superposition properties" to "correspond to the properties of the prepared superposition": This adjustment enhances clarity and formality.

Llama 3.1 (70B): "Spectra in the experiment" to "The experiment's results show" to make the sentence more clear and concise, and to avoid using the phrase "Spectra in the experiment" which sounds somewhat awkward (see Figure 3).

Provided that two arbitrary quantum states of light get mixed by a qubit" → "Provided that two arbitrary quantum states of light are mixed by a qubit" (added a verb to make the sentence grammatically correct).

Students' constructive criticism: [I found that] when detecting the same error in the original version, Gemini Ultra, GPT 4-o and Llama 3.1 (70B) demonstrated different levels of improvement, from a preposition change to rephrasing the entire structure with the subject replacement. Some tools suggested the Perfect grammar form instead of the Past Simple for currency. The grammatical structure to get + Participle II was replaced with a similar form to be + Participle II, the difference between these structures (the starting point of an action vs. a state) being not recognized. The noun + of + noun construction was suggested as a priority instead of "locked" noun +noun (GPT 4-o) for clarity.

Our comment: Comparative analysis provoked students' reflection on the differences that make sense when selecting the right grammar form in a scientific text. With such an approach grammar acquisition starts from analyzing the context with potential errors after which the grammar rule is inferred from practical examples.

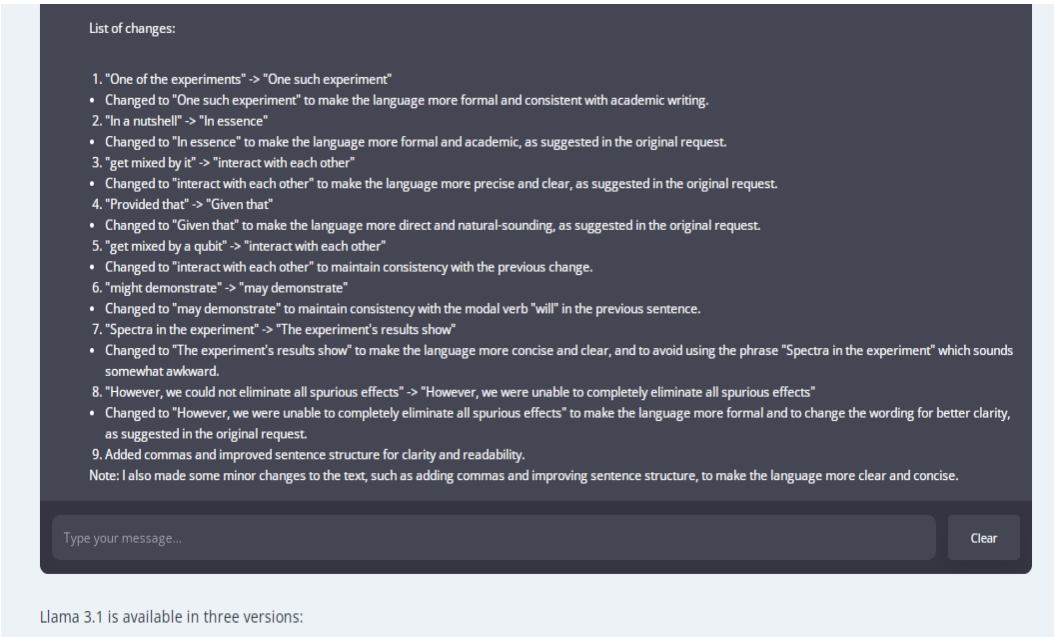


Figure 3. Llama 3.1 (70B): corrections (answer to the prompt)
Source: compiled by Anna V. Zabolotskikh, Elena E. Sokolova, Daria V. Tavberidze

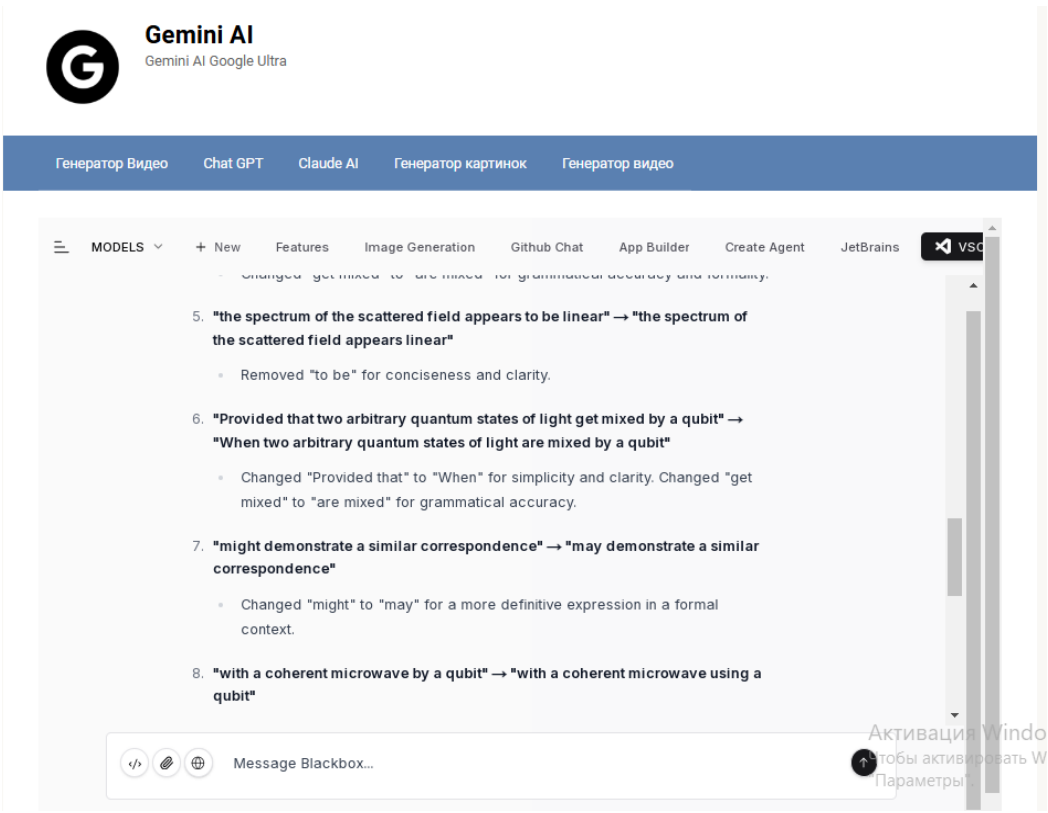


Figure 4. Gemini Ultra: corrections (answer to the prompt)
Source: compiled by Anna V. Zabolotskikh, Elena E. Sokolova, Daria V. Tavberidze

Example II. Source human-written text: *The Standard Model proved to be a reliable model but there are things it can't explain, like dark matter. So, we surely know that there are more that we know, we need to extend our models or find new ones. My research question is "Is it possible to find new particles in future experiments?". First of all, I'm interested in this question because revealing new frontiers of physics always brings new possibilities for mankind. A lot of new theoretical research seemed to be inapplicable but history proved otherwise. Scientists of the future will use our discoveries to invent technologies we would consider impossible today. Second, experiments for such research use the most advanced technologies of our time, most of them are invented specifically for this experiment. This means that new types of magnets or conductors or whatever designed for experiment could be used by engineers to make something useful for society. For example, the World Wide Web was invented by engineers in CERN.*

To speak more specifically on my thesis, in the near future, 10–20 years, a lot of experiments designed specifically for finding new physics are going to start collecting data. Estimating the number of events in these experiments for a variety of models will let us understand if these models should be investigated. We surely should work on models we could prove. Also, predictions on models could save us time after these experiments present data, we would already know what we want to see there.

LLMs generated results:

GPT 4-o: "there are things it can't explain" → "there are phenomena it cannot explain": Changed "things" to "phenomena" to use more precise language. Replaced "can't" with "cannot" for formality.

"revealing new frontiers of physics always brings new possibilities for mankind" → "revealing new frontiers of physics always brings new possibilities for humanity": "Mankind" was changed to "humanity" for inclusiveness and modern language use.

"We surely should work on models we could prove." → "We should focus on models that we can validate.": Changed "surely" to "should" for a more assertive tone and rephrased for clarity.

"Scientists of the future" → "Future scientists": Simplified for conciseness.

Gemini Ultra:

"but there are things it can't explain" → "however, there are aspects it cannot explain" Explanation: "However" provides a smoother transition, and "aspects" is more formal than "things." "Cannot" is preferred over "can't" in academic writing.

"to invent technologies we would consider impossible today" → "to invent technologies that we would consider impossible today" Explanation: "That" is often preferred for restrictive clauses in academic writing.

Llama 3.1 (70B):

Changed "proved to be a reliable model" to "has proven to be a reliable framework." This is a slight rewording to improve the clarity and concision of the sentence.

Students' criticism: The suggested changes are focused on the substitution of words for more formal equivalents (GPT 4-o), the correction of linking words and

the addition of restrictive relative pronouns (Gemini Ultra), the grammatical design of the sentence (Present Perfect to update the context instead of the factual Past Simple). Evaluative words are replaced by the modal verb should. A noun-phrase was suggested as preferable to the corresponding of-phrase (GPT 4-o). Long strings of nouns and adjectives are generally only used if they are names of pieces of equipment or methods.

Our comment: the students assess the register of the improved text on the lexical and grammatical levels.

Example III: Source human-written text: *In this thesis, the nature of the underlying theory will be assumed to be string/M theory. This is because of the following reasons. At present, string theory is the only known consistent theory of quantum gravity, at least at the perturbative level. But for the purposes of this thesis, more importantly, it is the only known ultra-violet complete theory which can naturally give rise to effective four dimensional theories of particle physics with the most important features of the Standard Model, namely, non-abelian gauge fields and chiral fermions. In addition, string theory can address all open questions in particle physics and cosmology within its framework and hopefully, solve them. Therefore, it seems reasonable to assume the existence of such a theory providing an underlying microscopic theoretical structure to our universe.With the recent development of superconducting technology, researchers began using artificial superconducting quantum two-level systems (qubits) as an environment for quantum optics experiments. One of the experiments is quantum wave mixing. In a nutshell, when two coherent microwaves with frequencies close to the qubit resonance get mixed by it, the spectrum of the scattered field appears to be linear with amplitudes of the peaks corresponding to the photon number statistics in the incident coherent states of the signals. Provided that two arbitrary quantum states of light get mixed by a qubit, the resulting spectrum might demonstrate a similar correspondence to the photon number statistics of the scattered signals. Suppose we mix a superposition of a photon and a vacuum with a coherent microwave by a qubit. Will the amplitudes of the peaks in the resulting spectrum reveal the properties of the quantum superposition? Spectra in the experiment of mixing a superposition with a coherent microwave clearly correspond to the prepared superposition properties. However, we could not eliminate all spurious effects. Moreover, there is no accurate theoretical description for the scattering spectra in this particular experiment yet.*

LLMs generated results:

GPT 4-o: "But for the purposes of this thesis, more importantly," → "More importantly,": Removed "But for the purposes of this thesis" as it was redundant and detracted from the sentence's clarity.

Removed the comma before "and hopefully, solve them.": Removed the comma for smoother flow and correct punctuation.

Gemini Ultra: "But for the purposes of this thesis, more importantly, it is the only known ultra-violet complete theory" → "More importantly, for the purposes of this thesis, it is the only known ultraviolet complete theory" Explanation:

Removed “But” for a smoother transition and changed “ultra-violet” to “ultraviolet” for consistency with standard terminology.

“and hopefully, solve them” → “and potentially solve them”

Explanation: “Potentially” is more formal and objective than “hopefully,” which is more subjective.

Llama 3.1 (70B): “more importantly, it is the only known ultra-violet complete theory” → “Moreover, for the purposes of this thesis, string theory is the only known ultra-violet complete theory” Explanation: Added the word “Moreover” to introduce the additional reason, and clarified the purpose of the thesis using the phrase “for the purposes of this thesis”.

“which can naturally give rise to effective four dimensional theories” → “that can naturally give rise to effective four-dimensional theories” Explanation: Changed “which” to “that” to maintain consistency in using the correct pronoun for a clauses, and added a hyphen to “four-dimensional” to adhere to APA style.

“Suppose we mix a superposition of a photon and a vacuum with a coherent microwave by a qubit” → “Suppose we mix a superposition of a photon and a vacuum with a coherent microwave via a qubit” (preposition) I replaced “by” with “via” to use a more formal and precise preposition.

“Do the amplitudes of the peaks in the resulting spectrum reveal the properties of the quantum superposition” (original text added a question mark at the end of this sentence) Since this is the last question of the paragraph, it would be more conventional to end the paragraph with a period.

“there is no accurate theoretical description for the scattering spectra in this particular experiment yet” → “there is currently no accurate theoretical description for the scattering spectra in this particular experiment” (verb tense) I changed the verb tense from “is” to “is currently” to make the sentence more consistent with the surrounding text.

Students’ criticism: In terms of the text’s logical coherence, in one and the same structure some phrases can be considered redundant (GPT 4-o), there may be a replacement of only one conjunction (Gemini Ultra) or an introductory word addition (Llama 3.1 (70B)). Moreover, in some cases, hyphenated and solid spelling of words (Gemini Ultra, Llama 3.1 (70B)) may be noted. Punctuation changes (GPT 4-o) or more formal lexical substitution may take place in the same structures for greater clarity of presentation (Gemini Ultra). Prepositions cannot always be used interchangeably: ‘by’ often identifies the ‘how’, and can suggest cause and effect’; ‘via’ (and ‘through’ and ‘by means of’) may not imply causality and may indicate that what comes afterwards is one option or a step among many (Llama). The future tense is changed for the present simple for factual, not prediction usage to avoid ambiguity. Words, belonging to informal register (yet), should be avoided.

Our comment: students evaluate cohesive structure of generated text in terms of redundancy of linking elements and spelling.

Example IV. Source human-written text: *Provided that sequence read counts at heterozygous genomic loci in the enriched sequencing experiments reflect the genomic copy numbers of alternative alleles, It is possible to estimate the*

Background imbalance of the allelic copy numbers from the enriched sequencing data alone. Research question: Is it possible to infer the genome-wide profiles of copy number imbalance in the alleles from enriched sequencing data? Key message: The relative background allelic dosage profiles can be reconstructed directly from heterozygous variant calls in enriched sequencing data and be used as a baseline in statistical evaluation of allele-specific transcription factor binding and chromatin accessibility.

LLMs generated results:

GPT 4-o:

"Background imbalance" → "background imbalance": Changed "Background" to "background" to maintain consistency with standard capitalization for terms unless they begin a sentence.

"of allele-specific transcription factor binding and chromatin accessibility." → "of allele-specific transcription factor binding and chromatin accessibility.": Added "the" for grammatical correctness, creating "the statistical evaluation."

Gemini Ultra:

Capitalized "Background" in the hypothesis and key message: In APA style, proper nouns and adjectives derived from proper nouns are capitalized.

Changed "allelic copy numbers" to "allelic dosage" in the key message: The term "allelic dosage" is more commonly used in this context to refer to the relative abundance of different alleles.

Added "the" before "statistical evaluation": Change: "in statistical evaluation" to "in the statistical evaluation." Explanation: Adding "the" clarifies that it refers to a specific evaluation, enhancing the precision of the statement.

Llama 3.1 (70B):

"It is possible to estimate the Background imbalance" → "it is possible to estimate the background imbalance" I removed the capitalization of "background" since it's not a proper noun. I also changed "Background" to "background" for consistency.

"of the allelic copy numbers from the enriched sequencing data alone" → "of allelic copy numbers from the enriched sequencing data alone" I removed "the" before "allelic copy numbers" since it's not necessary for clarity.

"in statistical evaluation of allele-specific transcription factor binding and chromatin accessibility" → "in statistical evaluations of allele-specific transcription factor binding and chromatin accessibility" I changed "evaluation" to "evaluations" since it's a plural noun and "evaluation" would be incorrect in this context.

Students' criticism: There is a lack of consistency regarding the standards for capitalizing nouns (GPT 4-o, Gemini Ultra). At the grammatical level, it can be noted that the article is added to abstract nouns for accuracy (GPT 4-o) or the singular form of the noun is replaced by the plural (Llama 3.1 (70B)). In addition, in some cases, more appropriate key terms (Gemini Ultra) may be suggested. It may be noted that there is a replacement of terminology, as well as a shortening of the sentence by replacing the construction with an impersonal It with a modal verb (Llama 3.1 (70B)). One and the same abstract noun was corrected differently and

changed into the plural form “evaluations” (Llama 3.1 (70B) or specified with the definite article (Gemini Ultra). The plural form would focus attention on the result whereas the singular form emphasises the process.

Our comment: students compare the generated outcome with their original texts in terms of grammar: articles, plural/singular forms of nouns and lexis: the key terminology.

Discussion

In order to lay the foundations for a discussion around the argument that the teaching process may benefit from the adoption of LLMs, in this article we analyse error-driven learning methodology presented to the PhD students.

Error-driven learning proposes that students compared the human-written abstract with the generated one and, if there is a discrepancy, they analyzed each detected misconception using guidelines. The artificial error climate created a learning environment, in which the PhD students practiced critical and even skeptical analysis towards the generated results to evaluate information based on verifiability and falsifiability.

The discrepancies obtained from three LLMs can be used in teaching academic writing on the lexico-semantic, lexico-grammatical and stylistic levels. While analyzing generated outcomes the PhD students realized that an artificial language corrector is not a universal tool, as it reflects the source of information embedded in them. Running three LLMs on one and the same prompt can generate very different answers. Thus, using error-driven approach provides the learners not only with factual knowledge but stimulates them to reason with this knowledge.

Conceptually, error-driven learning as a reflection-based method in LLMs studies is aimed at significant improvements in various reasoning tasks [48] through addressing the shortcomings and leveraging the pre-designed prompts. This study suggests a modified teaching approach based on the same principle. However, in our case, artificial systematic errors produced by machines and introduced to human learners serve as an activation-based mechanism for observing and explaining a wide range of linguistic phenomena on lexico-semantic, lexico-grammatical and semantic levels. That said, while the focus of error-driven learning is computational modelling, its principles may be applicable across many other areas including teaching foreign languages.

Conclusions

In this article, we reflect on the relationship between humans and machines in education. LLMs’ ability to produce different generated outcomes may be applied as error-driven methodology for academic writing teaching. The main purpose of this investigation was to argue for a different approach to error-driven learning research, making it less prescriptive and more enabling self-reflection and self-correction. With this in mind, we argue for a shift of focus in the research of academic writing teaching on critical analysis and editing to make it more sensitive

to visualizing errors. In error-driven learning machines enhance human work instead of replacing it. Interaction with artificial tools creates a sense of competition and stimulates further inquiry into where the truth actually lies. This seems to be a good example of how inconsistency and ambiguity may be turned to advantage.

References / Список литературы

1. Brown, T., Mann, B., Ryder, N., et al. (2020). Language Models are Few-Shot Learners. *Advances in Neural Information Processing Systems*, 33, 1877–1901.
2. Koraishi, O. (2023). Teaching English in the Age of AI: Embracing ChatGPT to Optimize EFL Materials and Assessment. *Language Education and Technology*, *3*(1), 55–72.
3. Binu, S. K., & Shanthi, C. (2024). Clinical insight: Comparative Analysis of Deep Learning Models for Disease Prediction across Multifaceted Datasets. In: *2024 Third International Conference on Distributed Computing and Electrical Circuits and Electronics (ICDCECE)* (pp. 1–8). IEEE.
4. Vaswani, A., Shazeer, N., Parmar, N., Uszkoreit, J., Jones, L., Gomez, A. N., et al. (2017). Attention is All You Need. *Advances in Neural Information Processing Systems*, 30.
5. Tran, K.T., Dao, D., Nguyen, M.D., Pham, Q.V., O'Sullivan, B., & Nguyen, H.D. (2025). Multi-Agent Collaboration Mechanisms: A Survey of LLMS. *arXiv*, preprint: 2501.06322.
6. Song, C., & Song, Y. (2023). Enhancing Academic Writing Skills and Motivation: Assessing the Efficacy of ChatGPT in AI-Assisted Language Learning for EFL Students. *Frontiers in Psychology*, 14, 1260843. https://doi.org/10.3389/fpsyg.2023.1260843 EDN: YPEEQR
7. Schmohl, T., Watanabe, A., Fröhlich, N., & Herzberg, D. (2020). How Artificial Intelligence can Improve the Academic Writing of Students. In: *Proceedings of* 10th International Conference the Future of Education (pp. 168–171).
8. Khabib, S. (2022). Introducing Artificial Intelligence (AI)-based Digital Writing Assistants for Teachers in Writing Scientific Articles. *Teaching English as a Foreign Language Journal*, *1*(2), 114–124. https://doi.org/10.12928/tefl.v1i2.249 EDN: PFNAFP
9. Storey, V.A., & Cunningham, B. (2024). Artificial Intelligence, Educational Leadership, Knowledge Production, and Transfer: A New Infrastructure. *Education Leadership Review*, *25*(2), 641–655. https://doi.org/10.46932/sfjdv4n2-001
10. Anson, C.M., & Straume, I. (2022). Amazement and Trepidation: Implications of AI-based Natural Language Production for the Teaching of Writing. *Journal of Academic Writing*, *12*(1), 1–9. https://doi.org/10.18552/joaw.v12i1.820 EDN: ИГЗЕКН
11. Floridi, L., & Chiriatti, M. (2020). GPT-3: Its Nature, Scope, Limits, and Consequences. *Minds and Machines*, *30*(4), 681–694. https://doi.org/10.1007/s11023-020-09548-1 EDN: УУЛУЛА
12. Safire, W. (2008). *Safire's Political Dictionary.* Oxford University Press.
13. Bender, E. M., & Koller, A. (2020, July). Climbing towards NLU: On Meaning, Form, and Understanding in the Age of Data. In: *Proceedings of the 58th Annual Meeting of the Association for Computational Linguistics* (pp. 5185–5198).
14. Gee, L., Rigutini, L., Ernandes, M., & Zugarini, A. (2023, December). Multi-Word Tokenization for Sequence Compression. In: *Proceedings of the 2023 Conference on Empirical Methods in Natural Language Processing: Industry Track* (pp. 612–621). https://doi.org/10.18653/v1/2023.emnlp-industry.58
15. Ji, Ziwei, Lee, N., Frieske, R., Yu, T., Su, D., Xu, Y., et al. (2023). Survey of Hallucination in Natural Language Generation. *ACM Computing Surveys*, *55*(12), 1–38. https://doi.org/10.1145/3571700
16. Peng, B., Narayanan, S., & Papadimitriou, C. (2024). On Limitations of the Transformer Architecture. *Corr*, Abs/2402.08164.
17. Bang, J., Lee, J., Lee, G.G., & Chung, M. (2014). Pronunciation Variants Prediction Method to Detect Mispronunciations by Korean Learners of English. *ACM Transactions on Asian Language Information Processing* (TALIP), *13*(4), 1–21. https://doi.org/10.1145/2629545

18. Gu, Y., Dong, L., Wei, F., & Huang, M. (2023). Minillm: Knowledge Distillation of Large Language Models. *Arxiv,* preprint: 2306.08543.
19. Lengo, N. (1995). What is an Error? *English Teaching Forum*, *33*(3), 20–24.
20. Lüdeling, A., Walter, M., Kroymann, E., & Adolphs, P. (2005). Multi-Level Error Annotation in Learner Corpora. In: *Proceedings of Corpus Linguistics* (Vol. 1. pp. 14–17).
21. Liu, Y., & Reichle, E.D. (2015). An Evolutionary Algorithm for Error-Driven Learning via Reinforcement. *arXiv* preprint:1503.07609.
22. Mollick, E.R., & Mollick, L. (2023). Using AI to Implement Effective Teaching Strategies in Classrooms: Five Strategies, Including Prompts. URL: https://teaching.uoregon.edu/ (accessed: 12.01.2025).
23. Kuleto, V., Ilić, M., Dumangiu, M., Ranković, M., Martins, O.M., Păun, D., & Mihoreanu, L. (2021). Exploring Opportunities and Challenges of Artificial Intelligence and Machine Learning in Higher Education Institutions. *Sustainability*, *13*(18), 10424. https://doi.org/10.3390/Su131810424 EDN: QRTPOF
24. Berrón, M. (2020). Aristotle's *Politics* I and the Method of Theanalytics. Rhizomata, *8*(1), 83–106. https://doi.org/10.1515/Rhiz-2020-0006 EDN: SFWMBH
25. Küng, L., Kröll, A. M., Ripken, B., & Walker, M. (1999). Impact of the Digital Revolution on the Media and Communications Industries. *Javnost-The Public*, *6*(3), 29–47. https://doi.org/10.1080/13183222.1999.11008717
26. De Saussure, F. (1916). Nature of the Linguistic Sign. In: *Course on General Linguistics* (pp. 65–70). New York : Philosophical Library.
27. Siertsema, B. (1955). The Linguistic Sign: the Sign in Itself. In: *A Study of Glossematics: Critical Survey of Its Fundamental Concepts* (pp. 128–145). Dordrecht: Springer Netherlands.
28. Church, A. (1951, July). The Need for Abstract Entities in Semantic Analysis. In: *Proceedings of the American Academy of Arts and Sciences* (Vol. 80. no. 1. pp. 100–112). American Academy of Arts & Sciences.
29. Harris, Z. (1976). On A Theory of Language. *The Journal of Philosophy*, *73*(10), 253–276. https://doi.org/10.2307/2025530
30. Chomsky, N. (1957). Logical Structure in Language. *Journal of The American Society for Information Science*, *8*(4), 284. https://doi.org/10.1002/asi.5090080406
31. Abbot-Smith, K., & Tomasello, M. (2006). Exemplar-Learning and Schematization in a Usage-Based Account of Syntactic Acquisition. *Linguistic Review*, *23*(3), 275–290. https://doi.org/10.1515/TLR.2006.011
32. Abusch, D. (1993). The Scope of Indefinites. *Natural Language Semantics*, *2*(2), 83–135. https://doi.org/10.1007/bf01250400 EDN: KQJLYA
33. Carlson, G.N. (1989). On the Semantic Composition of English Generic Sentences. In: *Properties, Types and Meaning*: Vol. II: Semantic Issues (pp. 167–192). Dordrecht: Springer Netherlands.
34. Grice, H.P. (1978). Further Notes on Logic and Conversation. *Syntax and Semantics*, 9, 113–127. https://doi.org/10.1163/9789004368873_006
35. Hiemstra, D. (2018). Language models. In: *Encyclopedia of Database Systems* (pp. 2061–2065). Springer.
36. Grice, H.P. (1975). *Grice-Logic and Conversation*. London: University College London Press.
37. Mirzadeh, I., Alizadeh, K., Shahrokhi, H., Tuzel, O., Bengio, S., & Farajtabar, M. (2024). GSM-Symbolic: Understanding the Limitations of Mathematical Reasoning in Large Language Models. *arXiv*, preprint:2410.05229.
38. Abendschein, B., Lin, X., Edwards, C., Edwards, A., & Rijhwani, V. (2024). Credibility and altered communication styles of AI graders in the classroom. *Journal of Computer Assisted Learning*, *40*(4), 1766—1776. https://doi.org/10.1111/jcal.12979 EDN: GVGEBJ
39. Galperin, I.R. (1973). About the Concepts of "Style" and "Stylistics". *Voprosy Jazykoznanija*, 3, 14–25. (In Russ.).
Гальперин И.Р. О понятиях «стиль» и «стилистика» // Вопросы языкознания. 1973. № 3. С. 14–25.

40. Shpetnyi, C. I. (2022). School of Linguistic Stylistics in Russia and Vista of Its Advance in the Paradigm of Modern Cognitive Knowledge. *Vestnik of Moscow State Linguistic University. Humanities*, *8*(863), 119–127. https://doi.org/10.52070/2542-2197_2022_8_863_119
41. Bennett, K. (2009). English Academic Style Manuals: a Survey. *Journal of English for Academic Purposes*, 8(1), 43–54. https://doi.org/10.1016/j.jeap.2008.12.003
42. Pagliawan, D.L. (2017). Feature Style for Academic and Scholarly Writing. *Academic Journal of Interdisciplinary Studies*, *6*(2), 35–41. https://doi.org/10.1515/ajis-2017-0004
43. Hayot, E. (2014). *The Elements of Academic Style: Writing for the Humanities*. Columbia University Press.
44. Rose, D., Rose, M., Farrington, S., & Page, S. (2008). Scaffolding Academic Literacy with Indigenous Health Sciences Students: an Evaluative Study. *Journal of English for Academic Purposes*, *7*(3), 165–179.
45. Kneale, P. (2008). *A Handbook for Teaching and Learning in Higher Education*. Routledge. https://doi.org/10.1016/j.jeap.2008.05.004
46. Jordan, R.R. (1997). *English for Academic Purposes*. Cambridge University Press.
47. Willems, S., Bollen, H., Van Der Veen, J., Sterpin, E., Crijns, W., Nuyts, S., & Maes, F. (2021). Learning from Mistakes: an Error-Driven Mechanism to Improve Segmentation Performance Based on Expert Feedback. In: *MICCAI Workshop on Distributed and Collaborative Learning* (pp. 68–77). Cham: Springer.

Supplement. Guidelines for evaluation the text appropriacy

Lexico-Semantic Level	**Phrasal verbs** – Are there direct verbs that could be used to increase clarity and precision? **Synonyms appropriateness** – Are the synonyms used appropriatly? – Do the meanings of synonyms maintain the equal emotional stress? **Spelling (Hyphenated/Solid)** – Have the spelling of hyphenated versus solid compound words been checked? – Are all compound words written correctly according to style guidelines?
Lexico-Grammatical Level	**Definiteness/Indefiniteness in Article Usage** – Are articles (a, an, the) used correctly to denote definiteness or indefiniteness? – Have noun phrases been reviewed to ensure proper article usage? **Nouns Capitalization** – Have capitalized proper nouns and other significant terms been used correctly? – Is there consistency with capitalization of terms used throughout the document? **Presence of Linking Words** – Are logical indicators present to guide the reader through the argument? **Sentence Structure** – Does the sentence structure predominantly follow the Subject-Predicate-Object sequence?
Stylistics	**Register** – Is the writing style consistent in maintaining an formal register? – Has ambiguity or vagueness in word choice been avoided? **Colloquial expressions** – Are there any colloquial expressions present? – Has informal language been replaced with formal alternatives? **Sentence style** – Are the sentences varied in structure while maintaining clarity?

Source: compiled by Anna V. Zabolotskikh, Elena E. Sokolova, Daria V. Tavberidze

Information about the authors:
Anna V. Zabolotskikh, Senior lecturer, Department of Foreign Languages, Faculty of Humanities and Social Sciences, Federal State Autonomous Educational Institution of Higher Education, RUDN University (6 Miklukho-Maklaya str., Moscow, Russian Federation, 117198); *Research interests:* lexicology, linguistics, TESOL, language acquisition, discourse analysis; *e-mail:* zabolotskikh-av@rudn.ru
ORCID: 0000-0002-5253-2733; Author ID 57217873153; SPIN-code: 6520-1445; Researcher ID: AAB-4687-2019.
Elena E. Sokolova, PhD in Philology, Associate Professor, Department of Foreign Languages, Moscow Institute of Physics and Technology (National Research University) (9 Institutsky lane, Moscow region, Dolgoprudny, Russian Federation, 141701); *Research interests:* cognitive linguistics, cultural studies, discourse analysis; *e-mail:* sokolova.ee@mipt.ru
ORCID: 0000-0002-1467-3455, Author ID: 56579818200; SPIN-code: 2013-1826; Researcher ID: G-4265-2019.
Daria V. Tavberidze, PhD in Philosophy, Associate Professor, Department of Foreign Languages, Faculty of Humanities and Social Sciences, RUDN University (6 Miklukho-Maklaya str., Moscow, Russian Federation, 117198); *Research interests:* linguistics, language acquisition, language learning, philosophy; *e-mail:* tavberidze-dv@rudn.ru
ORCID: 0000-0002-2727-6803; Author ID: 57195529929; SPIN-code: 3183-1668; Researcher ID: ABW-5226-2022.

Сведения об авторах:
Заболотских Анна Владимировна, старший преподаватель кафедры иностранных языков факультета гуманитарных и социальных наук, Российский университет дружбы народов (117198, Российская Федерация, г. Москва, ул. Миклухо-Маклая, д. 6); *сфера научных интересов:* лексикология, лингвистика, преподавание английского языка как иностранного, изучение языка, анализ дискурса; *e-mail*: zabolotskikh-av@rudn.ru
ORCID: 0000-0002-5253-2733; Author ID 57217873153; SPIN-код: 6520-1445; Researcher ID: AAB-4687-2019.
Соколова Елена Евгеньевна, кандидат филологических наук, доцент Департамента иностранных языков Московского физико-технического института (Национальный исследовательский университет) (141701, Российская Федерация, Московская область, г. Долгопрудный, Институтский переулок, 9); *сфера научных интересов:* когнитивная лингвистика, культурология, анализ дискурса; *e-mail*: sokolova.ee@mipt.ru
ORCID: 0000-0002-1467-3455; Author ID: 56579818200; SPIN-code: 2013-1826; Researcher ID: G-4265-2019.
Тавберидзе Дарья Владимировна, кандидат философских наук, доцент кафедры иностранных языков факультета гуманитарных и социальных наук, Российский университет дружбы народов (117198, Российская Федерация, г. Москва, ул. Миклухо-Маклая, д. 6); *сфера научных интересов:* лингвистика, освоение языка, изучение языка, философия; *e-mail*: tavberidze-dv@ rudn.ru
ORCID: 0000-0002-2727-6803; Author ID: 57195529929; SPIN-код: 3183-1668; Researcher ID: ABW-5226-2022.

RUDN Journal of Language Studies, Semiotics and Semantics
2313-2299 (Print), ISSN 2411-1236 (Online)
Вестник Российского университета дружбы народов.
Серия: ТЕОРИЯ ЯЗЫКА. СЕМИОТИКА. СЕМАНТИКА

2026 Vol. 17 No. 2 605–624
http://journals.rudn.ru/semiotics-semantics

ДИСКУРСИВНЫЕ ИССЛЕДОВАНИЯ

DISCOURSE STUDIES

DOI: 10.22363/2313-2299-2026-17-2-605-624
EDN: LUOMML

Research Article / Научная статья

Mitigation Strategies in South African English Corporate Communication

Elena N. Malyuga[1] ✉, **Maria P. Malkina**[2]

[1]RUDN University, *Moscow, Russian Federation*
[2]Moscow State Linguistics University, *Moscow, Russian Federation*
✉malyuga-en@rudn.ru

Abstract. This article examines how interactional tension is handled in the question-and-answer (Q&A) sections of corporate earnings calls conducted in South African English. The analysis uses transcripts from two events involving senior executives responding to evaluative questioning from analysts and investors. Attention is given to moments where questions introduce pressure, scepticism, or demands for justification, and to the linguistic resources executives use when responding in real time under public scrutiny. The study identifies a set of mitigation strategies, which use is examined in relation to what happens immediately after the response, using the following turn as evidence of whether the issue is treated as closed or remains interactionally active. The findings show that mitigation is a systemic part of answering in South African English earnings-call Q&A, closely tied to the interactional organisation of the genre. Responses that combine cautious formulation with collective positioning are more often followed by acceptance or movement to a new topic, while other responses leave room for continued questioning. Although the study does not pursue a comparative design, it extends interaction-based research on earnings calls to South African English, an institutional variety that has received little focused attention in this area. In doing so, it adds empirical detail to existing accounts of how disagreement is managed in high-stakes corporate communication.

Keywords: corporate communication, earnings calls, mitigation, management of disagreement

Authors' contribution: The authors contributed equally to this article and preparation of the article text**.**

Conflict of interests: The authors stated that there were no conflicts of interest.

Article history: Received: 20.01.2026; Accepted: 10.02.2026.

© Malyuga E.N., Malkina M.P., 2026
This work is licensed under a Creative Commons Attribution-NonCommercial 4.0 International License
https://creativecommons.org/licenses/by-nc/4.0/legalcode

For citation: Malyuga, E.N., & Malkina, M.P. (2026). Mitigation Strategies in South African English Corporate Communication. *RUDN Journal of Language Studies, Semiotics and Semantics, 17*(2), 605–624. https://doi.org/10.22363/2313-2299-2026-17-2-605-624 EDN LUOMML

Стратегии митигации в корпоративной коммуникации на материале южноафриканского варианта английского языка

Е.Н. Малюга[1], М.П. Малкина[2]

[1]Российский университет дружбы народов, *Москва, Российская Федерация*
[2]Московский государственный лингвистический университет, *Москва, Российская Федерация*
malyuga-en@rudn.ru

Аннотация. Анализ способов снижения коммуникативной напряженности в секциях вопросов и ответов (Q&A) корпоративных конференц-звонков по финансовой отчетности, проводимых на южноафриканском варианте английского языка, основан на расшифровках двух мероприятий, в рамках которых представители высшего руководства компаний в режиме реального времени отвечают на оценочные вопросы аналитиков и инвесторов. Особое внимание уделяется тем фрагментам общения, где вопросы содержат элементы давления, скепсиса или требования обоснования, а также языковым средствам, к которым прибегают руководители при ответе в условиях публичного контроля. Цель исследования — выявление набора стратегий митигации, использование которых рассматривается в соотнесении с последующим развитием диалога: следующий репликативный ход анализируется как показатель того, воспринимается ли вопрос как исчерпанный или сохраняет коммуникативную значимость. Результаты показывают, что митигация систематически проявляется в ответах в Q&A-сегментах конференц-звонков на южноафриканском английском и тесно связана с организацией речевого взаимодействия в данном жанре. Ответы, сочетающие осторожную формулировку с коллективным позиционированием, чаще сопровождаются принятием ответа или переходом к новой теме, тогда как иные типы ответов оставляют возможность для продолжения уточняющих вопросов. Хотя в работе не ставится задача сопоставительного анализа, исследование расширяет интеракционно ориентированные исследования конференц-звонков, включая в поле анализа южноафриканский вариант английского языка — институциональную разновидность, ранее практически не рассматривавшуюся в данном контексте. Тем самым работа дополняет существующие исследования, посвященные тому, каким образом несогласие регулируется в условиях высокорисковой корпоративной коммуникации.

Ключевые слова: корпоративная коммуникация, конференц-звонки по финансовой отчетности, митигация, регуляция несогласия

Вклад авторов: равнозначен на все этапах исследования и подготовки статьи к публикации.

Конфликт интересов: авторы заявили об отсутствии конфликтов интересов.

История статьи: дата поступления: 20.01.2026; дата приема в печать: 10.02.2026.

Для цитирования: *Malyuga E.N., Malkina M.P.* Mitigation Strategies in South African English Corporate Communication // Вестник Российского университета дружбы народов. Серия: Теория языка. Семиотика. Семантика. 2026. Т. 17. № 2. С. 605–624. https://doi.org/10.22363/2313-2299-2026-17-2-605-624 EDN LUOMML

Introduction

Conflict and disagreement are not exceptional in corporate communication. They are built into many of its most visible genres, particularly those that place senior executives in direct interaction with evaluative audiences. Board meetings, investor briefings, regulatory hearings, and media interviews all create conditions in which organisational decisions are questioned, defended, and re-justified in real time [1]. In such settings, disagreement is a predictable feature of institutional talk.

Earnings calls are a clear example. In the question-and-answer (Q&A) sections of these events, analysts and investors often return to the same pressure points, questioning strategic choices and pressing executives to account for performance or to clarify what lies ahead. These exchanges are public, time-constrained, and consequential for how the organisation is perceived by markets and other stakeholders [2]. Executives are expected to respond substantively, remain accountable, and keep the interaction moving forward, even when questions are sceptical or persistent [3]. The management of disagreement in this genre is therefore a practical necessity rather than a peripheral concern.

Linguistic research has long acknowledged that such interactional pressure is handled through mitigation. Available studies have described a range of linguistic resources that speakers draw on when addressing potential conflict, including modal verbs, epistemic markers, conditional formulations, and forms of collective reference [4–7]. In business and corporate communication, these resources have been linked to accountability management, reputational risk, and the need to respond to challenge without escalating tension [8–11]. More recent studies of financial communication treat earnings calls as interactional events in which questioning, justification, and stance-taking are negotiated turn by turn [12–14].

Despite this growing body of work, research on mitigation in earnings calls shows two recurring limitations. Empirically, most studies rely on a narrow range of English varieties, with British, American, or broadly defined international English serving as the default point of reference. Analytically, mitigation is often discussed in general terms, with attention focused on forms or broad pragmatic functions rather than on how mitigation operates within particular institutional varieties of English. As a result, established corporate varieties outside the Anglo-American core are seldom examined as empirical sites in their own right.

South African English occupies an ambiguous position in this respect. It is a long-standing institutional variety used in corporate governance, finance, and public communication [15], yet it remains underrepresented in interaction-based studies of corporate communication. While its phonological and lexical features are well documented, far less is known about how conflict and disagreement are managed in high-stakes corporate genres when South African English is the medium of interaction. In particular, earnings-call Q&A conducted by South African executives addressing an international audience has received little focused linguistic attention.

The present study addresses this gap via an examination of mitigation strategies in the Q&A sections of corporate earnings calls produced in South African English. The focus is deliberately narrow and empirical. The study does not aim to characterise South African English as a communicative style, nor to compare it explicitly with other varieties. Instead, we treat South African English as the linguistic material through which a well-defined corporate genre is realised, and ask how mitigation is used within that genre to manage conflict as it arises in interaction.

The study pursues two closely linked aims. It identifies and categorises mitigation strategies employed by executives when responding to conflict-relevant questions during earnings-call Q&A, and it assesses the effectiveness of these strategies by examining the immediate interactional development in the turn that follows. Effectiveness is understood in strictly local terms, based on whether a challenge is accepted and closed or whether pressure is renewed.

Accordingly, the study addresses the following research question: *Which mitigation strategies are employed in South African English earnings-call Q&A, and how effective are they in containing interactional conflict at the level of the immediate exchange?*

Literature Review

Conflict in corporate communication rarely takes the form of overt confrontation. More often, it emerges through questioning, through demands for explanation, or through the repeated return to issues that have not been settled to an interlocutor's satisfaction [16]. Earnings calls make this especially visible. In the Q&A sections, senior executives are placed in direct exchange with analysts and investors whose task is evaluative by design. Questions press on strategic choices, performance figures, or future direction, and they do so in a public setting where responses are recorded, circulated, and weighed carefully. Under these conditions, dealing with disagreement is a routine part of how the interaction is expected to unfold.

Linguistic research on mitigation offers a useful point of entry for understanding how speakers deal with moments of pressure in interaction. Mitigation is commonly treated as a range of linguistic practices through which commitment is adjusted and responsibility is managed when interaction risks becoming contentious [9]. Early pragmatic studies linked mitigation closely to politeness and face management, paying attention to how speakers soften potentially problematic utterances through devices such as hedging, modal expressions, or various forms of distancing [17—21]. Although this work was not concerned with earnings calls or corporate communication as such, it provided a descriptive vocabulary that still underpins analyses of how linguistic choices influence the subsequent development of an exchange.

More recent research places mitigation squarely within the study of corporate communication. Research on meetings, executive talk, and financial reporting

shows that mitigation is used repeatedly in situations where speakers are expected to account for their actions while remaining careful about how commitments are phrased [8; 11; 22–26]. In these studies, mitigation is less about interpersonal courtesy and more about managing exposure and responsibility in public-facing interaction. Linguistic choices reflect the fact that statements about strategy, performance, or future plans can carry tangible institutional consequences, which makes caution a routine part of how executives speak.

Earnings calls have drawn increasing attention within this line of research. They are often described as mixed-purpose events, combining the presentation of financial information with ongoing interactional work, most visibly in the Q&A sections where analysts' questions can be sceptical, persistent, or openly pressuring [27]. Corpus-based studies note that executive responses in these moments make extensive use of modal verbs and other epistemic devices, allowing speakers to answer while limiting exposure to reputational or regulatory risk [14; 28–32]. Other research has noted that executives often rely on inclusive reference and collective formulations. These formulations present decisions as belonging to the organisation rather than to the individual speaker, which matters in settings where responses are public and closely watched [33; 34].

Despite these developments, research on mitigation in earnings calls remains uneven in its empirical coverage. Most studies draw on data from Anglo-American corporate settings or rely on loosely defined corpora labelled as "international English". As a result, earnings calls are often treated as pragmatically similar across contexts, with little attention given to how mitigation practices might pattern in established varieties of English used outside the usual centres of corporate communication. This does not undermine the value of existing findings, but it does limit the range of settings in which they have been tested.

South African English stands out as an underused empirical site in this respect. It is a long-established institutional variety, widely used in corporate governance, finance, and other forms of public-facing communication. Sociolinguistic work has described its phonological, lexical, and stylistic features in considerable detail, situating it within a multilingual and historically layered setting [35–37]. By contrast, pragmatic research on South African English has tended to concentrate on everyday interaction, educational environments, or intercultural encounters [38–41], leaving high-stakes corporate genres largely unexamined.

Existing work on corporate communication in South African English has tended to focus on written material or on broad rhetorical patterns [42], with far less attention given to interaction-based genres such as earnings calls. As a result, little empirical work has examined how conflict and disagreement are handled linguistically in the Q&A sections of South African English earnings calls, or how mitigation functions when executives respond to an international audience under sustained evaluative pressure. This gap is all the more noticeable given the global reach of South African corporations and the public visibility of their earnings calls as moments of corporate accountability.

The present study addresses this gap through an analysis of mitigation strategies in the Q&A sections of earnings calls conducted in South African English. Rather than seeking to identify variety-specific traits in a typological sense, the study treats South African English as the linguistic medium through which a well-defined genre of corporate communication is realised. Attention is directed to the mitigation practices that recur in conflict-sensitive moments and to their immediate interactional consequences, as displayed in the turns that follow executive responses.

The study brings together research on mitigation and work on earnings calls, extending the empirical base of both strands to South African English. In doing so, it offers a closely delimited account of how disagreement is managed in a corporate genre where public scrutiny and accountability are built into the interaction itself.

Materials and Research Methods

Data and Material

The empirical material for this study consists of transcripts of publicly available earnings calls produced by Naspers Limited and its international subsidiary Prosus. The dataset includes two types of events that are central to executive–stakeholder interaction: a CEO transition earnings call and an earnings results webcast. Both events involve sustained Q&A sequences between senior executives and external interlocutors, primarily institutional investors and financial analysts.

The first text is the Prosus CEO Transition Conference Call held on 18 September 2023 (approximately 4,500 words). The second is the FY2025 Results Webcast conducted in November 2024 (approximately 8,000 words). The corpus thus amounts to roughly 12.500 words. These materials were selected for several reasons. First, they represent naturally occurring corporate communication with minimal editorial intervention, preserving the sequential organisation of turns and the interactional contingencies of live communication. Second, the topics under discussion (leadership change, capital allocation, long-term strategy, and financial performance) are intrinsically sensitive and frequently give rise to disagreement, scepticism, or pressure for justification. Third, the events are dominated by South African executives addressing an international audience, making them suitable for examining pragmatic practices associated with South African English in earnings-call Q&A within a global corporate setting.

No modifications were made to the transcripts apart from the removal of irrelevant procedural segments (opening remarks, technical announcements) that did not contain interactional material.

Analytical Orientation

The analysis is grounded in linguistic pragmatics, understood here as the study of how speakers use linguistic resources to act in interaction and to manage meaning

in relation to specific communicative situations. This orientation treats utterances as contributions to an unfolding exchange in which participants orient to one another's roles, expectations, and prior turns [43].

A pragmatic perspective allows attention to how the same linguistic forms contribute differently to an interaction depending on where they occur and what they respond to. This view assumes that linguistic resources do not have a single, fixed interactional role when considered in isolation — their relevance emerges from how they are used in response to challenge, how they adjust degrees of commitment, and how they position the speaker in relation to other participants. This is particularly evident in earnings-call Q&A as a genre of corporate communication, where speakers are required to handle pressure while remaining publicly accountable.

The analysis remains interaction-oriented without extending to full conversation-analytic modelling. Attention is restricted to local developments that can be observed directly in the transcript, without reference to inferred intentions or longer-term strategic outcomes.

Identification of Conflict-Relevant Turns

Conflict is not treated as overt confrontation but as interactional tension arising from diverging evaluations, expectations, or interests. Such tension is identified independently of mitigation effectiveness and prior to any coding of responses.

Conflict-relevant turns are identified on the basis of analyst or investor questions in the Q&A sections of earnings calls that introduce or sustain pressure on the executive speaker. These include sceptical questioning, requests for justification, challenges to decisions or projections, and reformulations that signal dissatisfaction with a prior answer. Identification relies on the wording and positioning of the question, interpreted with reference to the immediately preceding talk.

This broader sequential view is used only to establish that a given executive response is conflict-relevant. It does not enter the analysis of effectiveness and is not itself evaluated.

Units of Analysis and Coding

The analysis is turn-based. A turn is understood in the standard interactional sense as a continuous contribution by one speaker bounded by a change of speaker [11; 44; 45]. In the transcripts used here, turns are clearly delimited by speaker attribution and sequential order.

Two turns are examined for each case. The first is the *response turn*: the executive's turn produced in direct response to a conflict-relevant question in the earnings-call Q&A. The second is the *follow-up turn*: the next turn produced by the questioner (or another external interlocutor) immediately after the response. The response turn is used to identify and code mitigation practices. The follow-up turn

is used only to assess what happens next interactionally; it is not coded for mitigation.

Mitigation coding is manual and based on linguistic evidence in the response turn rather than on inferred intention. Response turns were read closely to identify linguistic resources associated in the literature with mitigation in corporate communication, including modalisation, epistemic marking, conditional constructions, and choices that adjust alignment (e.g., inclusive reference). No fixed typology was imposed at the identification stage. Where a response turn contained more than one mitigation practice, each instance was coded separately. Ambiguous cases that could not be assigned consistently on the basis of linguistic form and local sequential placement were excluded from quantitative counts.

Effectiveness Criterion

Effectiveness is assessed through what the interaction does immediately after the response turn, using the follow-up turn as the observable indicator. This restriction is deliberate. In corporate Q&A, sufficiency or insufficiency of a response is typically displayed right away: the interlocutor either treats the answer as adequate and moves on, or treats it as inadequate and keeps pressure on the same point. Four follow-up patterns are distinguished: (1) acceptance, i.e. acknowledgement, thanks, or explicit confirmation that treats the response as sufficient); (2) topic shift, i.e. movement to a different issue without returning to the contested point; (3) reformulation — the same challenge is re-raised in altered wording, and (4) renewed pressure — repetition, insistence, or escalation focused on the same issue.

Acceptance and topic shift are treated as effective, since they close the issue locally and allow progression to a new matter. Reformulation and renewed pressure are treated as ineffective, since they show that the response has not been accepted as sufficient and the point remains interactionally active.

This criterion keeps the evaluation of effectiveness tied to visible interactional development rather than to inferred success, intention, or longer-term consequences.

Results

Overview of the Dataset

The analysed dataset comprises 78 executive response turns produced in reply to conflict-relevant questions posed by analysts or investors during two corporate communication events. Each response turn directly addresses a challenge, request for justification, or sceptical evaluation identified in the preceding question turn, in line with the criteria set out in the Methodology section. All response turns contain at least one mitigation instance. In total, 78 primary mitigation instances were identified, with some response turns containing additional embedded instances.

Where multiple mitigation practices occurred within a single response turn, each instance was recorded separately.

Across the dataset, mitigation instances fall into three analytically defined groups identified during coding: *non-confrontational hedging, collaborative problem-solving,* and *combined use* of both within the same response turn.

The distribution is as follows:

- non-confrontational hedging: 37 instances (47%);
- collaborative problem-solving: 30 instances (38%);
- combined use: 11 instances (15%).

Non-confrontational hedging is the most frequent pattern when mitigation occurs in isolation. Collaborative problem-solving is slightly less frequent but remains a stable and recurrent practice. In 11 response turns, both practices occur together, either in successive stretches of the same turn or in overlapping form.

Other mitigation practices identified during initial reading occurred only sporadically (fewer than five instances each) and did not recur with sufficient consistency to warrant separate quantitative reporting.

Response turns classified as non-confrontational hedging rely primarily on modal verbs, epistemic adverbs, and conditional constructions. Modal verbs occur in 33 out of 37 instances (89%). The most frequent modal verb is *would* (34 occurrences), followed by *can* (28), *will* (21), and *might* (12). Multiple modal verbs often appear within the same response turn. Epistemic adverbs such as *perhaps*, *certainly*, and *generally* co-occur with modal verbs in 61% of hedging instances. Conditional constructions, including *if*-clauses and hypothetical framing, are present in 43% of cases. In many response turns, hedging devices appear early, preceding evaluative statements or explanatory content. In several cases, hedged formulations are followed by explicit statements of institutional position or justification within the same turn.

Collaborative problem-solving is realised through inclusive reference and collective framing. Inclusive pronouns (*we*, *our*, *us*) occur in 93% of instances classified under this category. Across the 30 response turns in this group, *we* appears 127 times and *our* appears 43 times. These response turns frequently include references to shared objectives, joint responsibility, or future-oriented collective action. In 24 out of 30 cases (80%), at least one formulation presents the issue as a shared task rather than as a contested point. References to continued engagement, reassessment, or joint review occur in 18 response turns. Collaborative problem-solving appears both in responses to external pressure and in exchanges involving speakers of comparable institutional standing.

In 11 response turns (15%), non-confrontational hedging and collaborative problem-solving co-occur. These cases typically involve a hedged formulation followed by a shift to collective framing, or the embedding of modalised statements within inclusive constructions. Response turns in this group are generally longer and contain a higher density of mitigation markers. Within a single turn, shifts in

speaker positioning are observable as speakers move between cautious formulation and collective positioning.

Distribution and Interactional Outcomes

The distribution of mitigation practices varies with interactional configuration. In response turns addressing questions from external interlocutors in positions of evaluative authority, non-confrontational hedging accounts for 62% of mitigation instances. In exchanges among speakers of comparable institutional standing, collaborative problem-solving is more frequent, accounting for 67% of instances in that subset. Combined use occurs in both configurations but is slightly more frequent in response turns addressing repeated or sustained questioning.

Interactional outcomes were assessed on the basis of the follow-up turn immediately after each response turn. Outcomes are classified as acceptance, topic shift, reformulation, or renewed pressure, as defined in the Methodology section. The distribution of effective outcomes (acceptance or topic shift) is as follows:

- non-confrontational hedging: 31 out of 37 cases (84%);
- collaborative problem-solving: 27 out of 30 cases (91%);
- combined use: 10 out of 11 cases (95%).

Non-effective outcomes involve reformulation of the challenge or renewed pressure on the same issue. These occur most frequently in response turns relying exclusively on hedging, particularly where the initial question is repeated or intensified.

The results thus show that mitigation in the analysed interactions is concentrated around two recurring practices, with a smaller but consistent set of response turns combining both. These practices differ in frequency, linguistic composition, and immediate interactional outcome.

The observed distributions provide an empirical basis for addressing the research question concerning how disagreement is managed through mitigation in earnings-call Q&A within corporate communication. An overview of the identified mitigation practices, their linguistic realisation, distribution, and interactional outcomes is presented in Table.

Discussion

Mitigation Strategies in Earnings-Call Q&A

The findings show that mitigation in South African English earnings-call Q&A relies on a limited and recurrent set of practices. Across the response turns analysed, most mitigation work is carried out through non-confrontational hedging and collaborative problem-solving, with a smaller but stable group of responses combining both within a single turn. This concentration reflects the interactional constraints of the genre, where executives are required to respond to evaluative questioning in real time, under conditions of public accountability and limited opportunity for repair.

Mitigation Practices in Executive Response Turns

Strategy	Attribute	Description
Non-confrontational hedging	Frequency	37 instances (47%)
	Linguistic realisation	Modal verbs *(would, can, might),* epistemic adverbs *(perhaps, generally)*, conditional constructions
	Typical setting	Executive responses to analyst or investor questioning
	Illustrative features	Early modalised framing followed by institutional justification
	Effective outcomes	31 / 37 (84%)
Collaborative problem-solving	Frequency	30 instances (38%)
	Linguistic realisation	Inclusive pronouns *(we, our)*, collective goals, future-oriented formulations
	Typical setting	Both asymmetric and symmetric exchanges
	Illustrative features	Framing issues as shared tasks or joint review
	Effective outcomes	27 / 30 (91%)
Combined use	Frequency	11 instances (15%)
	Linguistic realisation	Co-occurrence of hedging and inclusive framing
	Typical setting	Responses to repeated or sustained questioning
	Illustrative features	Modalised statements embedded in collective formulations
	Effective outcomes	10 / 11 (95%)

Source: compiled by Elena N. Malyuga & Maria P. Malkina.

Previous studies of earnings calls have described similar patterns in Anglo-American and international English data. Modalisation, epistemic caution, and calibrated formulations have been repeatedly identified as routine features of executive responses to analyst questioning, particularly in moments of scepticism or pressure [2; 3; 13; 46]. In the South African English data, executives make frequent use of cautious, qualified formulations. This mirrors what has been reported for earnings-call Q&A elsewhere and suggests that the mitigation practices observed here are part of the ordinary communicative work of the genre.

What is notable is the narrowness of the repertoire actually used. The response turns examined here return to the same practices with notable regularity. Mitigation appears to draw on a small set of forms that have become interactionally recognisable within the genre of earnings calls. As a result, it appears predictable in form even as the substantive content of questions varies. Executives seem oriented less to selecting among alternatives and more to producing answers that conform to established expectations about how scrutiny is to be handled.

This pattern is visible across different levels of pressure. The same resources recur in responses to brief clarification requests and to more sustained questioning. What varies is the amount of mitigation and its placement within the turn, not the type of practice itself. In this sense, mitigation functions as a stabilising feature of earnings-call interaction, providing a familiar way of managing exposure across a range of topics.

Notably, general pragmatic research on mitigation has long emphasised the diversity of linguistic resources available to speakers across interactional settings. Studies grounded in politeness theory, stance, and interpersonal pragmatics describe mitigation as an umbrella term covering a wide range of devices, from modalisation and epistemic marking to reformulation, distancing, vagueness, and shifts in footing. This body of work is primarily concerned with mapping what speakers can draw on in principle, across genres and institutional environments, and it therefore tends to foreground the breadth of the repertoire rather than its patterned restriction in any one setting.

Work that takes earnings calls as its object of analysis paints a more delimited picture. Research in this area repeatedly notes that, despite the theoretical availability of many mitigating options, executive responses in earnings-call Q&A tend to rely on a small number of recurrent practices. Modal verbs, epistemic qualifiers, and collective formulations appear with high frequency, while other potential resources described in the mitigation literature occur rarely or not at all. This concentration has been linked to the institutional design of the genre, where responses are produced under time pressure, recorded for public circulation, and evaluated by both the immediate interlocutor and a wider, unseen audience. Within these conditions, linguistic variability appears constrained by expectations of recognisability and procedural adequacy rather than expressive range.

For example, Alanko's [47] analysis of persuasion in earnings calls provides a clear illustration of this point. The author uses a large corpus of calls across companies and reporting periods and shows that executives routinely recycle the same interactional techniques to produce responses that are structurally and pragmatically similar even when the substantive content differs. The genre emerges as standardised in its interactional solutions, as speakers orient to established ways of managing challenge and accountability. This observation builds on earlier descriptions of earnings calls as governed by a shared linguistic protocol [48], in which deviation from familiar response formats may carry interactional or reputational risk.

The South African English material analysed here can thus be viewed as aligning with genre norms: South African English earnings-call Q&A relies on the same limited set of interactionally established solutions that previous studies have documented in other settings.

Interactional Effectiveness and Local Containment of Conflict

The study assesses effectiveness through immediate interactional development, using the follow-up turn as the indicator of whether a response is treated as sufficient or remains open to challenge.

A long line of interactional research takes the next turn as the most reliable place to see how a prior contribution has landed. In conversation-analytic work, uptake is treated as something participants display to one another, often without comment or evaluation. A response that is taken as sufficient is allowed to recede,

and one that is treated as lacking is brought back into focus, reformulated, or pressed again. This way of reading interaction does not depend on assumptions about intention or success. It rests on what happens next, and on how participants organise continuation or persistence in real time [49–51]. In institutional settings, where interaction is task-oriented and time-sensitive, this logic becomes especially visible. Progression itself does much of the analytic work.

Studies of workplace disagreement adopt this perspective with little hesitation. Handford and Koester [6], for instance, describe how contested matters remain interactionally active when interlocutors do not take up the direction implied by a prior response. Vine [52] makes a similar observation in her work on professional interaction, noting that agreement need not be declared for an exchange to move on. What matters is whether the issue is treated as closed for the purposes of the ongoing activity. Research on earnings calls follows the same reasoning. Analyses of analyst–executive exchanges read follow-up questions, acknowledgements, and topic shifts as displays of how an answer has been received under conditions of scrutiny and limited time [2; 3; 13; 48; 53; 54].

The present study stays within this tradition and treats next-turn behaviour as a narrow but grounded point of observation. It captures how mitigation works in the moment where an answer is either allowed to stand or becomes the basis for continued pressure. This criterion is intentionally narrow, but it is well suited to the genre under study. In earnings-call Q&A, acceptance, topic shift, reformulation, or renewed pressure are typically displayed without delay, which makes next-turn uptake a reliable site for assessing local containment of conflict.

The distribution of outcomes suggests that mitigation practices differ in their likelihood of closing an issue at the level of the immediate exchange. Responses built around collaborative problem-solving, either alone or in combination with hedging, are more often followed by acceptance or topic shift than responses relying exclusively on hedging. This does not imply that hedging is ineffective in general. In most cases, hedged responses still lead to local resolution. At the same time, hedging alone accounts for the majority of cases in which questioning persists.

Similar observations have been made in studies of workplace and institutional interaction, where continued disagreement is often traced to how a prior turn leaves the direction of the exchange insufficiently settled [6; 52]. Where a response moderates commitment without offering a shared frame for moving forward, the issue may remain interactionally active. Collective framing, by contrast, can recast the matter as institutionally managed or jointly addressed, thus reducing the relevance of further challenge even when disagreement has not been explicitly resolved.

The combined use of hedging and collaborative problem-solving appears particularly effective in this respect. In these response turns, modalised formulations regulate exposure while inclusive reference situates the issue within a collective remit. The two practices work together to produce responses that are cautious yet interactionally sufficient. The relatively high rate of immediate

resolution associated with this pattern suggests that stacking mitigation resources within a single turn is one reliable way of containing sustained questioning in earnings-call Q&A.

What matters for the present analysis is how closure is achieved in practice. Work on cognitive engagement has shown that linguistic formulation affects how information is processed and whether further engagement is prompted [55]. A similar dynamic is visible in earnings-call Q&A, where mitigation appears to influence whether a response invites continuation or supports progression. In many cases, follow-up turns acknowledge a response or move the exchange to a new topic without any explicit statement of agreement. Progress becomes visible when a response provides a workable point from which the interaction can continue. Effectiveness, in this sense, is observable in how a response supports the onward movement of the exchange and allows the next turn to proceed without renewed focus on the same issue.

Differences between mitigation practices take on clearer interactional form when seen in these terms. Hedged responses frequently remain institutionally acceptable and defensible, and they often satisfy the immediate demand for an answer. At the same time, such responses may leave the future handling of the issue open. In cases where questioning continues, the response has typically managed commitment while keeping the issue available for further treatment. Collective formulations, by contrast, tend to present the matter as already situated within shared procedures or ongoing institutional work. This framing appears sufficient to support continuation of the exchange, even when specific points remain to be addressed later.

Using the follow-up turn as the point of observation also places clear boundaries on what the analysis can claim. The focus remains on immediate interactional development, without extending to later assessments or external consequences. Within this narrow window, responses either enable the interaction to move on or keep the issue in play. Across the material examined here, mitigation shows a consistent and limited effect at this level: it influences whether a point remains active in the exchange or recedes enough for the interaction to proceed. This form of local containment is a routine feature of earnings-call Q&A, where time constraints and public exposure favour progression over extended dwelling on a single issue.

South African English as Empirical Material

The study was not designed as a comparative investigation of pragmatic variation across Englishes, and the findings do not support strong claims about variety-specific mitigation strategies. The practices identified here are well attested in earnings-call research more generally. What the analysis demonstrates is that these practices are also central in South African English earnings-call Q&A and that they function in comparable ways when executives address an international evaluative audience.

This observation matters because South African English has rarely been examined as an empirical site for interaction-based research on earnings calls. Existing studies overwhelmingly rely on Anglo-American data or on mixed corpora labelled as international English, often treating the genre as pragmatically uniform. The present findings do not challenge those accounts, but they extend their empirical base by showing that the same mitigation practices are integral to South African corporate communication in this genre.

At the same time, the results caution against attributing the observed patterns to national or cultural communicative styles. The strategies identified here are best understood as responses to the interactional organisation of earnings-call Q&A rather than as markers of South African English in the abstract. The contribution of the study lies in documenting how mitigation operates in this material, not in claiming pragmatic distinctiveness.

Scope of Contribution

The scope of the study is deliberately limited. The dataset comprises two events from a single corporate group, and the analysis focuses on turn-level developments visible in transcript data. Within these boundaries, the study provides interaction-based evidence of how mitigation strategies are deployed in South African English earnings-call Q&A and how they relate to immediate interactional outcomes.

The contribution is therefore empirical and descriptive. It adds underrepresented material to earnings-call research and demonstrates a clear link between mitigation practices and local containment of conflict. It does not attempt to generalise across varieties of English, to model long-term persuasive effects, or to account for prosodic and multimodal resources. Future work can extend this line of inquiry via broadening the dataset, incorporating audio data, or placing South African English earnings calls alongside other varieties in a comparative design.

Conclusions

This study examined how executives manage interactional tension in the Q&A sections of earnings calls conducted in South African English. Focusing on conflict-relevant questions and on what happens immediately after executive responses, the analysis traced which mitigation strategies recur in this genre and how they function at the level of the local exchange.

The findings show that mitigation work in this material is concentrated within a small and stable set of practices. Non-confrontational hedging and collaborative problem-solving account for most executive responses to pressure, with a smaller group of turns combining both within a single response. These strategies recur across events and are deployed systematically in moments where questions introduce scepticism, demand justification, or return to unresolved issues. In this sense, mitigation appears as an ordinary part of answering in earnings-call Q&A.

Assessment of immediate interactional uptake indicates that mitigation strategies differ in how effectively they contain tension at the next turn. Responses framed through collaborative problem-solving, particularly when combined with hedging, are more likely to be treated as sufficient and followed by acceptance or topic shift. Hedging on its own often succeeds in closing an issue, but it also accounts for most cases in which pressure is renewed. This suggests that containment of conflict in earnings-call Q&A depends both on regulating commitment, and on how responses position the issue within a shared institutional trajectory.

Although the study does not intend to compare varieties of English, its findings remain relevant for work on South African English in corporate communication. The mitigation practices observed here are well documented in earnings-call research based on Anglo-American data, yet they have rarely been examined in South African material. The analysis shows that executives using South African English rely on the same interactional resources when addressing an international evaluative audience, and that these resources function in comparable ways at the level of immediate exchange. At the same time, the data do not warrant claims about distinctive national styles. The observed patterns are better understood as responses to the interactional organisation of earnings-call Q&A than as properties of South African English in the abstract.

The study is necessarily limited in scope. It draws on two events from a single corporate group and relies on transcript data, which means that prosodic and multimodal resources remain outside the analysis. Its contribution lies in close, interaction-based observation rather than in generalisation. Future research could extend this work through an examination of a broader range of South African firms, incorporating audio data, or placing South African English earnings calls alongside other varieties in a comparative design. Such work would make it possible to see whether the patterns described here recur in other corporate settings, without losing sight of how conflict is handled turn by turn.

References

1. Scollon, R., Scollon, S.W., & Jones, R.H. (2011). *Intercultural Communication: A Discourse Approach.* Amsterdam: John Wiley & Sons.
2. Crawford Camiciottoli, B. (2011). Ethics and Ethos in Financial Reporting: Analyzing Persuasive Language in Earnings Calls. *Business Communication Quarterly, 74*(3), 298–312. https://doi.org/10.1177/1080569911413810
3. Crawford Camiciottoli, B. (2010). Earnings Calls: Exploring an Emerging Financial Reporting Genre. *Discourse & Communication*, *4*(4), 343–359. https://doi.org/10.1177/1750481310381681
4. Bella, S. (2025). Mitigation, Rapport, and Identity Construction in Workplace Requests. *Languages, 10*(8), Art. 179. https://doi.org/10.3390/languages10080179
5. Faber, B. (2003). Creating Rhetorical Stability in Corporate University Discourse: Discourse Technologies and Change. *Written Communication*, *20*(4), 391–425. https://doi.org/10.1177/0741088303259869

6. Handford, M., & Koester, A. (2019). The Construction of Conflict Talk Across Workplace Contexts: (Towards) A theory of conflictual compact. *Language Awareness*, *28*(3), 186–206. https://doi.org/10.1080/09658416.2019.16368027.
7. Jehn, K.A. (1997). A Qualitative Analysis of Conflict Types and Dimensions in Organizational Groups. *Administrative Science Quarterly*, *42*(3), 530–557. https://doi.org/10.2307/2393737 EDN: CNVKBH
8. Angouri, J., & Locher, M.A. (2012). Theorising Disagreement. *Journal of Pragmatics*, *44*(12), 1549–1553. https://doi.org/10.1016/j.pragma.2012.06.011
9. Caffi, C. (2013). Mitigation. In: M. Sbisà & K. Turner (Eds.), *Pragmatics of Speech Actions* (pp. 257–286). De Gruyter Mouton. https://doi.org/10.1515/9783110214383.257
10. Harutyunyan, L., & Baghramyan, R. (2022). Identification of Hedging Strategies in Business Negotiations. *Wisdom*, *2*(22), 170–181. https://doi.org/10.24234/wisdom.v22i2.706 EDN: SBLSOY
11. Malyuga, E., & McCarthy, M. (2018). English and Russian Vague Category Markers in Business Discourse: Linguistic Identity Aspects. *Journal of Pragmatics*, *135*, 39–52. https://doi.org/10.1016/J.PRAGMA.2018.07.011 EDN: YBXXLN
12. Al Fajri, M.S., & Oktavianti, I.N. (2024). Stance Expressions in Applied Linguistics Research Articles: A Corpus-Based Contrastive Study. *Training, Language and Culture*, *8*(1), 54–65. https://doi.org/10.22363/2521-442X-2024-8-1-54-65 EDN: TUGGRP
13. Hirsto, H., Koskela, M., & Jokipii, A. (2023). Performing Financial Communication as Professional Practice: The Interplay of Consensus and Tension in Earnings Calls. *Journal of Professions and Organization*, *10*(2), 165–181. https://doi.org/10.1093/jpo/joad011 EDN: IQXUJV
14. Yang, Z., & Ai, B. (2025). "We're doing well in virtually every corner of the world": A Corpus-Assisted Discourse Study of Persuasiveness in Apple's Earnings Conference Calls. *Journal of Business and Technical Communication*, *39*(3), 275–301. https://doi.org/10.1177/10506519251326577
15. Roberts-Lombard, M., Mpinganjira, M., Wood, G., & Svensson, G. (2019). South African Corporate Ethics Codes: Establishment and Communication. *European Business Review*, *31*(3), 379–396. https://doi.org/10.1108/EBR-08-2017-0150
16. Malyuga, E.N. (2023). A Corpus-Based Approach to Corporate Communication Research. *Russian Journal of Linguistics*, *27*(1), 152–172. https://doi.org/10.22363/2687-0088-33561 EDN: AARCSO
17. Bargiela-Chiappini, F. (2003). Face and Politeness: New (Insights) for Old (Concepts). *Journal of Pragmatics*, *35*(10-11), 1453–1469. https://doi.org/10.1016/S0378-2166(02)00173-X EDN: EJYPRX
18. Bargiela-Chiappini, F., & Harris, S. (2006). Politeness at Work: Issues and Challenges. *Journal of Politeness Research: Language, Behavior, Culture*, *2*(1), 7–33. https://doi.org/10.1515/PR.2006.002
19. Brown, P., & Levinson, S.C. (1987). *Politeness: Some Universals in Language Usage*. Cambridge: Cambridge University Press.
20. Fraser, B. (1990). Perspectives on Politeness. *Journal of Pragmatics*, *14*(2), 219–236. https://doi.org/10.1016/0378-2166(90)90081-N
21. Tracy, K. (2008). "Reasonable hostility": Situation-Appropriate Face-Attack. *Journal of Politeness Research: Language, Behavior, Culture*, *4*(2), 169–191. https://doi.org/10.1515/JPLR.2008.009
22. Lebedeva, I.S., & Ponomarenko, E.V. (2023). Manifestations of Politeness and Power in Problem-Solving Business Telephone Calls. *Training, Language and Culture*, *7*(2), 64–71. https://doi.org/10.22363/2521-442X-2023-7-2-64-71 EDN: OQDRLE
23. Leontovich, O.A., & Nikitina, A.V. (2024). Communication Monitoring as a Politeness Mechanism. *Training, Language and Culture*, *8*(3), 62–72. https://doi.org/10.22363/2521-442X-2024-8-3-62-72 EDN: GMYCZZ

24. Litvinova, A.V., & Larina, T.V. (2023). Mitigation Tools and Politeness Strategies in Invitation Refusals: American and Russian Communicative Cultures. *Training, Language and Culture*, *7*(1), 116–130. https://doi.org/10.22363/2521-442X-2023-7-1-116-130 EDN: OVCWYT
25. Malyuga, E.N., & Tomalin, B. (2025). A Functional-Pragmatic Approach to Wealth-Related Euphemisms Study in Business Discourse. *Russian Journal of Linguistics*, *29*(3), 473–491. https://doi.org/10.22363/2687-0088-44130 EDN: BETFKR
26. Marra, M., & Dawson, S. (2021). Workplace and Institutional Discourse. In: M. Haugh, D. Z. Kádár, & M. Terkourafi (eds.), *Cambridge Handbook of Sociopragmatics* (pp. 475–495). Cambridge: Cambridge University Press. https://doi.org/10.1017/9781108954105.025
27. Nagengast, M. (2025). Rhetoric And Language in Strategic Financial Communication: Earnings Announcements as Means of Impression Management. *International Journal of Strategic Communication*, 1–21. https://doi.org/10.1080/1553118X.2025.2583142
28. Boginskaya, O.A. (2022). Cross-Disciplinary Variation in Metadiscourse: A Corpus-Based Analysis of Russian-Authored Research Article Abstracts. *Training, Language and Culture*, *6*(3), 55–66. https://doi.org/10.22363/2521-442X-2022-6-3-55-66 EDN: PWPCDQ
29. Cheng, W. (2014). Corpus Analyses of Professional Discourse. In: V. Bhatia & S. Bremner (Eds.), *The Routledge handbook of Language and Professional Communication* (pp. 13–25). London: Routledge. https://doi.org/10.4324/9781315851686
30. Hyland, K., & Jiang, F. (2024). Metadiscourse: The Evolution of an Approach to Texts. *Text & Talk*, *44*(3), 411–433. https://doi.org/10.1515/TEXT-2021-0156 EDN: NLWUNW
31. Lee, W.W.L. (2020). Impression Management Through Hedging and Boosting: A Cross-Cultural Investigation of the Messages of US and Chinese Corporate Leaders. *Lingua*, *242*, Art. 102872. https://doi.org/10.1016/j.lingua.2020.102872 EDN: ELXSZH
32. Lorson, A., Rohde, H., & Cummins, C. (2023). Epistemicity and Communicative Strategies. *Discourse Processes*, *60*(8), 556–593. https://doi.org/10.1080/0163853X.2023.2255494 EDN: SLSBVS
33. Fu, X., Wu, X., & Zhang, Z. (2021). The Information Role of Earnings Conference Call Tone: Evidence from Stock Price Crash Risk. *Journal of Business Ethics*, *173*(3), 643–660. https://doi.org/10.1007/s10551-019-04326-1
34. Palmieri, R., Rocci, A., & Kudrautsava, N. (2015). Argumentation in Earnings Conference Calls. Corporate Standpoints and Analysts' Challenges. *Studies in Communication Sciences*, *15*(1), 120–132. https://doi.org/10.1016/j.scoms.2015.03.014
35. Connell, B., & Zeitlyn, D. (2023). Sociolinguistic Studies of West and Central Africa. In: M. J. Ball (Ed.), *The Routledge Handbook of Sociolinguistics Around the World* (pp. 371–381). London: Routledge. https://doi.org/10.4324/9780203869659
36. Weldon, T.L. (2022). Visibly Invisible: The Study of Middle Class African American English. *Language and Linguistics Compass*, *16*(11), Art. e12477. https://doi.org/10.1111/lnc3.12477 EDN: BPRCOH
37. Wigdorowitz, M., Pérez, A.I., & Tsimpli, I.M. (2022). Sociolinguistic Context Matters: Exploring Differences in Contextual Linguistic Diversity in South Africa and England. *International Multilingual Research Journal*, *16(*4), 345–364. https://doi.org/10.1080/19313152.2022.2069416 EDN: YJQUGV
38. Ameka, F.K. (2023). *Pragmatics in African Context: Introduction.* Leiden: Leiden University.
39. Huddlestone, K., & Fairhurst, M. (2013). The Pragmatic Markers *Anyway, Okay,* and *Shame*: A South African English Corpus Study. *Stellenbosch Papers in Linguistics Plus*, *42*, 93–110. https://doi.org/10.5842/42-0-170
40. Schroeder, A., & Sickinger, P. (2025). From Observation to Elicitation: An Ethnographically Grounded Approach to Pragmatic Variation in Namibian English. *Journal of Pragmatics*, *235*, 99–111. https://doi.org/10.1016/j.pragma.2024.11.004 EDN: HWKMLK

41. Unuabonah, F., & Mtembu, N. (2023). Multilingual Pragmatic Markers in South African English. *Southern African Linguistics and Applied Language Studies*, *41*(3), 264–279. https://doi.org/10.2989/16073614.2022.2123366 EDN: GJPVCH
42. Nyangiwe, B., & Tappe, H. (2021). Politeness Constructions in Written Business Communication: A Plea for African Politeness Strategies. *South African Journal of African Languages*, *41*(1), 44–54. https://doi.org/10.1080/02572117.2021.1902133 EDN: VTOYRB
43. Malyuga, E.N. (2024). *The Language of Corporate Communication: Functional, Pragmatic and Cultural Dimensions*. Springer. https://doi.org/10.1007/978-3-031-58905-8
44. Hutchby, I., & Wooffitt, R. (2008). *Conversation Analysis.* Cambridge: Polity Press.
45. Sacks, H., Schegloff, E.A., & Jefferson, G. (2025). A Simplest Systematics for the Organization of Turn-Taking for Conversation. *Language*, *101*(1), 151–188. https://doi.org/10.1353/lan.1974.0010 EDN: UHNNBQ
46. Doey, B., & de Jong, P. (2025). How Negative Tones in Earnings Calls Shape Media Narratives. *Review of Behavioral Finance*, *17*(3), 406–423. https://doi.org/10.1108/RBF-08-2024-0228 EDN: KJRRXH
47. Alanko, M. (2024). *Persuasive Language in Earnings Calls.* Tampere: Tampere University.
48. Crawford Camiciottoli, B. (2019). Multimodality and Financial Communication: The Case of Earnings Calls. *Ibérica*, *37*, 17–38.
49. Drew, P., & Heritage, J. (1992). *Talk at Work.* Cambridge: Cambridge University Press.
50. Grishechko, E. G. (2023, April 7). Enhancing Language Learning with Podcasts: Evidence from Feedback and Practice. In: *Proceedings of the 2nd Pan-Russian Research and Practical Conference Topical Issues of Intercultural Communication and Foreign Language Teaching Methodology* (pp. 69–72). Rostov-on-Don: Akadem-Lit. EDN: DJONAB
51. Schegloff, E. A. (2007). *Sequence Organization in Interaction: A Primer in Conversation Analysis.* Cambridge: Cambridge University Press.
52. Vine, B. (2020). *Introducing Language in the Workplace.* Cambridge: Cambridge University Press.
53. Crawford Camiciottoli, B. (2018). Persuasion in Earnings Calls: A Diachronic Pragmalinguistic Analysis. *International Journal of Business Communication*, *55*(3), 275–292. https://doi.org/10.1177/2329488417735644
54. Peluso, M. (2021). The Functional Approach, Semiotics and Professional Discourse. *Training, Language and Culture*, *5*(2), 62–72. https://doi.org/10.22363/2521-442X-2021-5-2-62-72
55. Grishechko, E.G., & Tomalin, B. (2025). Cognitive Engagement in Scientific Writing: Empirical Evidence on Research Data Processing Variation. *Research Result. Theoretical and Applied Linguistics*, *11*(2), 54–79. EDN: QUBPTA

Information about the authors:

Elena N. Malyuga, Dr. Sc. (Philology), Professor, Head of Foreign Languages Department at the Faculty of Economics, RUDN University (6 Miklukho-Maklaya Str., Moscow, Russian Federation, 117198); *Research interests:* theory and practice of intercultural professional and business communication, pragmatics, corpus studies, discourse analysis; *e-mail*: malyuga-en@rudn.ru
ORCID: 0000-0002-6935-0661; SPIN-code: 8612-2186; AuthorID: 376376; ResearcherID: J-8927-2017; Scopus Author ID: 56587911800.
Maria P. Malkina, Senior Lecturer, Moscow State Linguistics University (38 build 1 Ostozhenka Str., Moscow, Russian Federation, 119034); *Research interests:* hedges, politeness theory, multimodal communication; *e-mail*: maria020602@mail.ru
ORCID: 0009-0003-4854-5034; SPIN-code: 2628-3691; AuthorID: 1295946; ResearcherID: JJE-2633-2023; Scopus Author ID: 58193066600.

Сведения об авторах:
Малюга Елена Николаевна, доктор филологических наук, профессор, заведующая кафедрой иностранных языков экономического факультета, Российский университет дружбы народов (117198, Российская Федерация, г. Москва, ул. Миклухо-Маклая, д. 6); *сфера научных интересов*: теория и практика межкультурной профессиональной и деловой коммуникации, прагматика, корпусные исследования, дискурс-анализ; *e-mail*: malyuga-en@rudn.ru
ORCID: 0000-0002-6935-0661; SPIN-код: 8612-2186; AuthorID: 376376; ResearcherID: J-8927-2017; Scopus ID: 56587911800.
Малкина Мария Павловна, ассистент, Московский государственный лингвистический университет (119034, Российская Федерация, г. Москва, ул. Остоженка, д. 38 с. 1); *сфера научных интересов*: хеджи, теория вежливости, мультимодальная коммуникация; *e-mail*: maria020602@mail.ru
ORCID: 0009-0003-4854-5034; SPIN-код: 2628-3691; AuthorID: 1295946; ResearcherID: JJE-2633-2023; Scopus ID: 58193066600.

RUDN Journal of Language Studies, Semiotics and Semantics
2313-2299 (Print), ISSN 2411-1236 (Online)
Вестник Российского университета дружбы народов.
Серия: ТЕОРИЯ ЯЗЫКА. СЕМИОТИКА. СЕМАНТИКА

2026 Vol. 17 No. 2 625–647
http://journals.rudn.ru/semiotics-semantics

DOI: 10.22363/2313-2299-2026-17-2-625-647
EDN: LYHRVN

Научная статья / Research Article

Взаимосвязь метафоры и оценки в репрезентации образа России в англоязычном медиадискурсе Великобритании и Индии

О.И. Калинин

Московский государственный лингвистический университет, *Москва, Российская Федерация*
Военный университет имени князя Александра Невского Министерства обороны РФ, *Москва, Российская Федерация*
Южно-Уральский государственный университет (национальный исследовательский университет), *Челябинск, Российская Федерация*
okalinin.lingua@gmail.com

Аннотация. В условиях глобальной информационной поляризации национальный образ государства превращается в стратегический ресурс мягкой силы и объект дискурсивного противоборства. Данная работа посвящена сравнительному лингвопрагматическому анализу метафорической и оценочной репрезентации России в англоязычном медиадискурсе Великобритании и Индии. Актуальность исследования обусловлена контрастными геополитическими позициями стран, что позволяет выявить степень прагматической и идеологической детерминированности выбора оценочных средств и метафорических моделей при конструировании одного и того же референта. Целью исследования являлся сравнительный анализ особенностей метафорической и оценочной репрезентации образа России в двух сопоставимых подкорпусах англоязычных медиатекстов с количественным выявлением характера их взаимодействия и зависимости от экстралингвистических факторов. Для достижения цели был проведен качественный лингвопрагматический анализ оценочных и метафорических средств; разработаны и применены интегральные индексы оценочности и метафоричности, выполнен корреляционный анализ взаимосвязи между плотностью, интенсивностью и языковой формой оценки и метафоры в каждом подкорпусе, установлена статистическая связь между характером общей оценки России и параметрами метафоричности на всем массиве данных. Полученные результаты показали, что британский медиадискурс характеризуется высокой плотностью и эмоциональной интенсивностью как оценочных средств, так и метафор, реализуя стратегию открытой негативной оценки России, тогда как индийский дискурс демонстрирует умеренные показатели и преобладание рациональных форм оценки, соответствующих политике стратегической автономии. Корреляционный анализ подтвердил сильную отрицательную связь между негативной оценкой России и интенсивностью/плотностью метафор при статистически незначимой роли языковой формы метафоры. Таким образом, именно эмоционально насыщенные и частотные метафоры выступают ключевым инструментом формирования антироссийского нарратива в медиадискурсе.

© Калинин О.И., 2026
This work is licensed under a Creative Commons Attribution-NonCommercial 4.0 International License
https://creativecommons.org/licenses/by-nc/4.0/legalcode

Ключевые слова: лингвопрагматика, национальный образ, корреляционный анализ, репрезентация России

Финансирование и благодарности: Исследование выполнено за счет гранта Российского научного фонда № 24-18-00049, https://rscf.ru/project/24-18-00049/

Конфликт интересов: автор заявил об отсутствии конфликтов интересов.

История статьи: дата поступления: 23.12.2025; дата приема в печать: 15.02.2026.

Для цитирования: *Калинин О.И.* Взаимосвязь метафоры и оценки в репрезентации образа России в англоязычном медиадискурсе Великобритании и Индии // Вестник Российского университета дружбы народов. Серия: Теория языка. Семиотика. Семантика. 2026. Т. 17. № 2. С. 625–647. https://doi.org/10.22363/2313-2299-2026-17-2-625-647 EDN LYHRVN

The Interrelation of Metaphor and Evaluation in the Representation of Russia's Image in English-Language Media Discourse in Great Britain and India

Oleg I. Kalinin

Moscow State Linguistic University, *Moscow, Russian Federation*
Military University*, Moscow, Russian Federation*
South Ural State University, *Chelyabinsk, Russian Federation*
okalinin.lingua@gmail.com

Abstract. In the context of global information polarization, a nation's image becomes a strategic resource of soft power and an object of discursive confrontation. This study is devoted to a comparative linguopragmatic analysis of the metaphorical and evaluative representation of Russia in English-language media discourse in Great Britain and India. The relevance of the research stems from the contrasting geopolitical positions of these countries, which makes it possible to reveal the extent to which the choice of evaluative means and metaphorical models is pragmatically and ideologically determined when constructing representations of the same referent. The aim of the study was to perform a comparative analysis of the specific features of metaphorical and evaluative representation of Russia's image in two comparable subcorpora of English-language media texts, with quantitative identification of the nature of their interaction and dependence on extralinguistic factors. To achieve this aim, a qualitative linguopragmatic analysis of evaluative and metaphorical devices was conducted; integrated indices of evaluativity (EDI, EII, ELFI) and metaphoricity (MDI, MII, MLFI) were developed and applied; correlation analysis of the relationships between density, intensity, and linguistic form of evaluation and of metaphor was carried out for each subcorpus; and the statistical relationship between the overall evaluative tone towards Russia (+/–) and metaphoricity parameters across the entire dataset was established. The findings demonstrate that the British media discourse is characterised by high density and emotional intensity of both evaluative means and metaphors, thereby implementing a strategy of overt negative evaluation and demonisation of Russia. In contrast, the Indian discourse exhibits moderate indices and a predominance of rational and latent evaluative forms consistent with a policy of strategic autonomy. Correlation analysis confirmed a strong negative relationship between negative evaluation of Russia and the intensity/density of metaphors, whereas the linguistic form of metaphors proved statistically insignificant. Thus, emotionally charged and frequent metaphors emerge as the key instrument for constructing an anti-Russian narrative in media discourse.

Keywords: linguopragmatics, national image, correlation analysis, representation of Russia

Financing: The research is financially supported by Russian Science Foundation № 24-18-00049, https://rscf.ru/project/24-18-00049/.

Conflict of interests: The author declares no conflict of interest.

Article history: Received: 23.12.2025; Accepted: 15.02.2025.

For citation: Kalinin, O.I. (2026). The Interrelation of Metaphor and Evaluation in the Representation of Russia's Image in English-Language Media Discourse in Great Britain and India. *RUDN Journal of Language Studies, Semiotics and Semantics, 17*(2), 625–647. (In Russ.). https://doi.org/10.22363/2313-2299-2026-17-2-625-647 EDN LYHRVN

Введение

В условиях современного глобального информационного пространства национальный образ страны превратился в стратегический нематериальный актив, который во многом определяет ее позиционирование на мировой арене. Он выступает одним из центральных элементов публичной дипломатии и концепции «мягкой силы». Как показали научные исследования, позитивный или негативный национальный образ напрямую коррелирует с объемом прямых иностранных инвестиций, туристическими потоками, экспортом высокотехнологичной продукции, дипломатической поддержкой в международных организациях и даже с поддержкой третьих стран в геополитических конфликтах [1; 2].

Позиционирование другого государства в национальных и международных СМИ является дискурсивно опосредованным процессом, в котором медиа выступают активным агентом идеологической стратификации мирового порядка. Через систематическое использование определенных фреймов, топосов и оценочно-маркированных нарративов медиадискурс конструирует иерархию государств-акторов, распределяя их по шкале от «цивилизованных» и «ответственных» до «ревизионистских», «авторитарных» или «угрожающих глобальной стабильности» [3].

В этом смысле репрезентация иностранного государства перестает быть нейтральным описанием и превращается в акт концептуального осмысления, который напрямую влияет на восприятие суверенитета страны, права на собственную внешнеполитическую субъектность и допустимость применения силы в отношении нее [4]. Таким образом, борьба за доминирующий образ той или иной страны в глобальном информационном поле становится одной из ключевых форм современного геополитического соперничества, где дискурсивные ресурсы приобретают статус стратегического оружия гибридных войн [5].

Сравнительный анализ британского и индийского англоязычного медиадискурса в контексте репрезентации России обладает высокой эвристической ценностью, поскольку позволяет эмпирически зафиксировать контрастные модели концептуализации и аксиологической интерпретации одного и того же референта в рамках двух принципиально разных цивилизационно-идеоло-

гических контекстов: «коллективного Запада» и Глобального Юга. Как известно, Великобритания, один из ключевых центров западного мира, в значительной степени воспроизводит нарративы, в которых Россия предстает ревизионистской державой и вызовом либеральному миропорядку [6]. Индия, напротив, будучи ведущей державой Глобального Юга и активным участником БРИКС и ШОС, демонстрирует большую дискурсивную автономию. Индийские СМИ, несмотря на использование английского языка и отдельных западных стандартов журналистики, во многом сохраняют национальную перспективу и независимость в оценке международных событий [7].

Именно поэтому изучение механизмов метафорической и оценочной репрезентации России в англоязычных медиадискурсах разных геополитических полюсов позволяет установить, в какой степени выбор и комбинация метафорических моделей и оценочных маркеров детерминированы прагматическими целями дискурса, идеологической позицией продуцента и целевой аудиторией. С сугубо лингвистической точки зрения такое исследование выявляет корреляцию между метафорическими сценариями и аксиологической нагрузкой языковых единиц, демонстрируя, как прагматический контекст определяет полярность оценки, интенсивность эмоционального воздействия и степень идеологической маркированности образа референта.

Так, в данной статье с использованием квантитативно-корпусного подхода исследуется взаимосвязь между частотностью и типологией средств выражения оценки референта «Россия», с одной стороны, и употреблением метафор, с другой, в зависимости от экстралингвистических (геополитических) факторов и идеологической ориентации медиадискурса. На материале двух сопоставимых подкорпусов британских и индийских англоязычных СМИ устанавливается, в какой мере идеологическая позиция продуцента определяет не только полярность оценки и выбор метафорических сценариев, но и степень их синергетического взаимодействия.

Целью настоящего исследования является сравнительный анализ особенностей метафорической и оценочной репрезентации образа России в англоязычном медиадискурсе Великобритании и Индии. В работе количественно выявляется характер оценочной метафорики в контексте прагматической обусловленности двух национальных дискурсивных сообществ.

Материал и методы исследования

Материалом исследования послужили два специализированных сопоставимых подкорпуса англоязычных медиатекстов, собранных методом целенаправленного отбора с элементами стратифицированной выборки в период февраля 2022 — июня 2025 гг.:

1) Подкорпус «Великобритания»: 50 новостных и аналитических текстов из The Economist, The Guardian, The Daily Telegraph, BBC News, Reuters UK, Financial Times. Общий объем — 37 469 словоупотреблений.

2) Подкорпус «Индия»: 50 текстов из The Times of India, The Indian Express, The Hindu, India Today, Daily Pioneer. Общий объем — 23 841 словоупотребление.

Для достижения поставленной цели была разработана и применена комплексная методика, включающая три последовательных этапа.

1. Формирование корпуса и подготовка материала

Определение тематических, хронологических и источниковедческих параметров; обеспечение сопоставимости подкорпусов по жанру, объему и релевантности.

2. Качественный лингвопрагматический анализ

а) Анализ оценочности:

— определение общей тональности текста (комбинация нейросетевого сентимент-анализа и экспертной оценки);

— сплошной текстологический разбор с фиксацией эксплицитных и имплицитных средств выражения оценки;

— классификация средств выражения оценки по осям «эксплицитность / имплицитность» и «рациональность / эмоциональность».

б) Анализ метафоричности:

— идентификация языковых метафор и реконструкция лежащих в их основе концептуальных моделей;

— классификация метафор по форме (прямые / непрямые), интенсивности и прагматической функции.

3. Количественный анализ и статистическая обработка

Расчет плотности и интенсивности оценочных и метафорических единиц; введение индексов метафоричности и оценочной маркированности; корреляционный анализ взаимосвязи между частотностью метафор и степенью оценочности в каждом подкорпусе; сравнение полученных показателей между британским и индийским дискурсом.

Используемые квантитативные индексы, формулы их подсчета и описание представлены в табл. 1.

Таблица 1. **Формулы квантитативного анализа оценочности и метафоричности**
Table 1. **Formulas for the quantitative analysis of evaluativity and metaphor power**

Параметр / Parameter	Формула / Formula	Описание / Description
Оценочность / Evaluation		
Плотность **Evaluativeness Density Index**	**EDI = (nelm*100)/nw**	**nelm —** количество всех языковых средств выражения оценки; **nw —** общее количество слов
Интенсивность **Evaluativeness Intensivity Index**	**EII = (ER*1 + EE*3)/nelm**	**ER —** количество проявлений рациональной оценки; **EE —** количество средств проявления эмоциональной оценки

Окончание табл. 1 / Ending of the Table 1

Параметр / Parameter	Формула / Formula	Описание / Description
Языковая форма **Evaluativeness Language Form Index**	**(EIm*1 + EEx*3)/nelm**	**EIm** — количество средств выражения скрытой, имплицитной оценки; **EEx** — количество средств выражения явной эксплицитной оценки
Метафоричность / Metaphor power		
Плотность **Metaphor Density Index**	**MDI= nme *100/nwords**	**nwords** — общее количество слов; **nme** — общее количество метафор
Интенсивность **Metaphor Intensivity Index**	**MII = (1w+2a+3s)/nme**	**w** — количество низкоинтенсивных метафор; **a** — количество метафор средней интенсивности**;** **s** — количество высокоинтенсивных метафор
Языковая форма **Metaphor Language Form Index**	**MLFI= (1I+3E)/nme**	**I** — количество непрямых метафор; **E** — количество прямых метафор

Источник: составлено О.И. Калининым. / *Source:* compiled by Oleg I. Kalinin.

Примененный комплексный подход позволил не только описать качественные различия в репрезентации России, но и количественно подтвердить степень прагматической обусловленности метафоры и оценки в каждом из исследуемых медиадискурсов.

Оценочность как прагматическая характеристика медиадискурса

Медиадискурс выполняет роль доминирующего канала конструирования и циркуляции коллективных ментальных репрезентаций о чужой стране. Особенно значима роль медиадискурса в ситуациях асимметричного информационного обмена, когда большинство реципиентов не имеет прямого опыта взаимодействия с репрезентируемой страной и полностью зависит от вторичных, опосредованных источников. В таких условиях языковые и дискурсивные средства (метафора, оценочная лексика, синтаксические конструкции имплицитной оценки) превращаются в инструменты идеологической интерпелляции и когнитивного манипулирования, способные закреплять в массовом сознании устойчивые образы «друга», «врага», «партнера» или «угрозы». Таким образом, национальный образ перестает быть лишь отражением реальности и становится дискурсивно сконструированной семиотической сущностью, чья структура во многом детерминирована аксиологическими и идеологическими параметрами продуцирующего его медиадискурса.

Как отмечает М. Беднарек, «оценка, или языковое выражение мнения говорящего / пишущего, лишь недавно стала объектом пристального лингвистического анализа, однако она пронизывает медиатексты, выступая важным инструментом формирования общественного восприятия» [8. С. 241]. Коммуникативно-прагматический потенциал медиатекста реализуется через строгую ориентацию на достижение заданного перлокутивного эффекта: формирование общественного мнения, конструирование социальных стереотипов, регулирование коллективных эмоций и ценностных установок.

Оценка в медиадискурсе проявляется в систематическом использовании языковых средств субъективно-модальной интерпретации фактов, событий и действующих лиц, что позволяет автору или медиаинституту эксплицировать собственную аксиологическую позицию и транслировать ее реципиенту. Будучи встроенной в ткань текста на лексическом, синтаксическом, стилистическом и риторическом уровнях, оценка не просто сопровождает изложение информации, а активно структурирует восприятие адресата, формируя устойчивые когнитивные и эмоциональные схемы. Таким образом, оценочный компонент выступает одним из главных инструментов идеологического конструирования реальности в поляризованном медиапространстве, определяя не только интерпретацию отдельного медиатекста, но и более широкие общественно-политические дискурсивные практики, что позволяет использовать в научном обороте термин оценочность применительно к тексту и дискурсу в целом, а не только к отдельной языковой единице.

Классически в отечественном языкознании под термином оценочность понимается «способность языковой или речевой единицы эксплицировать положительные и отрицательные свойства объекта, его место на оценочной шкале и в аксиологическом пространстве» [9. С. 101]. Однако в рамках настоящего исследования оценочность трактуется прежде всего как интегративная имманентная характеристика, присущая медиатексту на всех уровнях его организации и определяющая способность дискурса выражать субъективно-аксиологическое отношение к описываемой реальности.

Оценочность как свойство дискурса проявляется в перманентном присутствии субъективного модуса в сообщении, благодаря которому в рамках дискурсивной коммуникации свойства референта неизбежно подвергаются аксиологической интерпретации. Она пронизывает ткань текста, реализуясь одновременно на лексическом (эпитеты, оценочная лексика), синтаксическом (инверсии, парцелляция, риторические вопросы, модальные конструкции) и стилистическом (метафоры, ирония, гипербола, аллюзии) уровнях. При этом оценочность включает в себя тональность и модальность как свои частные проявления, но не сводится к ним, выступая более широкой дискурсивной категорией, организующей весь аксиологический каркас текста [10. С. 185].

Будучи свойством дискурса, оценочность может быть эксплицитной (прямо закрепленной в семантике языковых единиц, например, brilliant leader / dangerous extremist) и имплицитной (скрытой, выводимой только из

контекста, пресуппозиций, интертекстуальных связей и фоновых знаний реципиента). Эксплицитная оценочность опирается на формализованные языковые средства, такие как модальные глаголы, оценочная лексика или стандартизированные речевые клише, обеспечивая когнитивную доступность и иллокутивное воздействие, в то время как имплицитная активирует латентные смыслы через апелляцию к интуиции, традиции или культурным сценариям, требуя от адресата сложных когнитивных операций интерпретации. Как отмечается в проведенных исследованиях, различие между эксплицитной и имплицинтой формами оценочности в дискурсе коренится в отсутствии явных лингвистических маркеров в имплицитных конструкциях, что делает их особенно уязвимыми для идеологической манипуляции в реальных сценариях коммуникации [11].

Кроме того, оценочность как свойство дискурса всегда двойственна по своей природе: она сочетает эмоциональный (аффективный) и рациональный (когнитивно-оценочный) компоненты, которые, взаимопроникая, формируют у адресата устойчивые ценностные ориентации и эмоционально-оценочные фреймы. В основе этой двойственности лежит сложное взаимодействие языковых и логических форм выражения оценки: апелляция к образцу или интуиции активирует субъективное, эмоциональное восприятие, создавая аффективный фрейм [12], в то время как использование норм, стандартов и аналогий структурирует дискурс на рациональной основе, обеспечивая логическую когерентность. В реальном дискурсивном взаимодействии эти формы редко функционируют изолированно, они пересекаются, взаимно усиливаясь и образуя гибридные конструкции, где эмоциональное воздействие дополняется рациональным обоснованием. Разграничение этих компонентов, несмотря на их взаимопроникновение, позволяет глубже понять механизмы субъективного позиционирования и когнитивной интерпретации, где рациональные апелляции к нормам и аналогиям усиливают доверие к эмоциональным импликациям.

Стоит отметить, что оценочность может быть реализована на всех уровнях языковой системы, от графического и фонетического (интонационные акценты, типографические выделения) до лексического, грамматического и стилистического, где каждая единица вносит вклад в общее аксиологическое содержание. Логические формы, такие как апелляция к традиции, образцу или интуиции, интегрируют субъективное отношение говорящего в объективные критерии, повышая убедительность и доверие аудитории.

Таким образом, рассматривая оценочность именно как неотъемлемое свойство медиадискурса, а не как набор разрозненных языковых маркеров, мы получаем возможность выявить глубинные механизмы идеологического конструирования реальности и целенаправленного прагматического воздействия на массовую аудиторию. Комплексное оценочное воздействие медиатекста складывается из взаимодействия эксплицитных и имплицитных, эмоциональных и рациональных, лексических, синтаксических и риторических

средств, которые в совокупности формируют устойчивые аксиологические фреймы и перлокутивные эффекты.

Для выявления и измерения этого комплексного оценочного воздействия наиболее эффективным оказывается квантитативный подход, сочетающий корпусный анализ с вычислительными методами. Такой подход сосредоточен не на фиксации отдельных оценочных лексем или синтаксических конструкций, а предполагает расчет интегральных индексов, отражающих общий оценочный потенциал текста и его идеологическую направленность. В рамках предлагаемой методики выделяются два последовательных этапа:

1. Выявление оценочного потенциала всего текста (от негативного до позитивного), осуществляемое с помощью комбинации методов анализа при помощи нейросетей и экспертной разметки.

2. Подсчет и классификация оценочных средств с использованием специализированных количественных индексов: Evaluativeness Density Index (EDI) — плотность оценочных маркеров на 100 лексических единиц текста; Evaluativeness Intensity Index (EII) — соотношение эмоциональных и рациональных компонентов оценки; Evaluativeness Language Form Index (ELFI) — доля эксплицитной и имплицитной форм выражения оценки.

Совокупное применение этих индексов дает объективную, воспроизводимую и сравнимую картину оценочного профиля медиатекста, позволяя прогнозировать степень и характер идеологической поляризации в сопоставимых подкорпусах национальных СМИ. Тем самым квантитативный подход превращает оценочность из качественной интуитивной категории в измеримую дискурсивную переменную, открывая путь к строгому сравнительному анализу механизмов формирования общественного мнения в современном англоязычном медиапространстве.

Метафора как средство выражения оценки

Метафора занимает особое место среди средств выражения оценочности в медиадискурсе, поскольку сочетает когнитивную, прагматическую и аксиологическую функции, позволяя одновременно структурировать восприятие реальности и транслировать субъективное отношение к ней. Будучи одновременно единицей языка, мышления и культуры, метафора формирует устойчивый оценочный фрейм, задающий позитивную или негативную тональность объекту описания [13].

Оценочная сила метафоры обусловлена ее способностью активировать имплицитные культурные сценарии и эмоциональные реакции, часто оставаясь за пределами прямого лексического маркирования. Исследования последних лет убедительно показывают, что «метафоры способны передавать более интенсивные эмоциональные реакции, чем их буквальные аналоги» [14], оказывая комплексное метафорическое речевое воздействие, понимаемое как «целенаправленное или нецеленаправленное, концептуальное или

эмоциональное воздействие посредством использования метафор в тексте и дискурсе» [15. С. 135].

В медиатекстах оценочный потенциал метафоры реализуется через перенос аксиологических характеристик сферы-источника на сферу-цель: концептуальные метафоры ВОЙНЫ, БОЛЕЗНИ, СТИХИЙНОГО БЕДСТВИЯ систематически используются для негативной репрезентации политических оппонентов, миграционных процессов или экономических кризисов, тогда как метафоры ПУТЕШЕСТВИЯ, СВЕТА, СЕМЬИ — для позитивного фрейминга «своих» акторов и инициатив [16; 17]. Подобное распределение подтверждается и другими исследованиями: «среди метафор с высокой частотностью преобладают негативные оценки (64,4 %), в то время как положительные составляют только 35,6 %. Наибольшую негативную оценочность имеют метафоры, связанные с войной, медициной и трагическими событиями» [18]. Так, очевидно, что метафора является инструментом идеологического конструирования, где выбор сферы-источника предопределяет аксиологическую окраску всего нарратива.

Особую роль в реализации оценочной функции играет интенсивность (новизна) метафоры. Согласно теории «карьеры метафоры» Б. Боудл и Д. Гентнера [19], а также современным разработкам О.И. Калинина [15] конвенциональные (стертые) метафоры обрабатываются через быстрый процесс категоризации и чаще выражают рациональную, социально закрепленную оценку, тогда как новые (авторские) метафоры требуют активного сравнения доменов и обладают значительно большим эмоциональным и персуазивным потенциалом. Исследования политического и военного дискурсов подтверждают линейную зависимость: чем выше интенсивность метафоры, тем сильнее выражен аффективный компонент оценки и тем выше перлокутивный эффект [15].

Отдельно следует выделить различие между прямыми и непрямыми метафорами как носителями эксплицитной и имплицитной оценочности. Прямые (преднамеренные) метафоры («Эта политика — это рак общества», «Демократия — это маяк надежды») эксплицитно указывают на сравнение и, согласно теории преднамеренной метафоры Дж. Стина [20], целенаправленно привлекают внимание реципиента к самому акту метафорического переноса, усиливая иллокутивную силу оценки. Непрямые метафоры («общество заражено ложью», «идеи расцвели») воздействуют имплицитно, через атрибутивные и глагольные конструкции, создавая суггестивный эффект и позволяя автору сохранять формальную «объективность» [21].

Таким образом, метафора выступает одним из средств выражения оценочности в медиадискурсе, позволяя одновременно конструировать социальную реальность, эмоционально воздействовать на аудиторию и маскировать идеологическую позицию под образно-ассоциативной формой. Для выявления комплексного оценочного воздействия метафор в сопоставимых подкорпусах либеральных и консервативных СМИ целесообразно применять

многоуровневый квантитативный подход, включающий расчет трех взаимодополняющих индексов (15): Metaphor Density Index (MDI) — количество метафор на 100 лексических единиц текста; Metaphor Intensity Index (MII) — уровень новизны всех метафор в тексте, что отражает эмоциональное и рациональное воздействие; Metaphor Language Form Index (MLFI) — соотношение прямых и непрямых метафор, что показывает эксплицитность и имплицитность метафорического речевого воздействия.

Совокупное применение этих индексов дает объективную, воспроизводимую картину метафорического фрейминга и позволяет статистически подтвердить различия в оценочных стратегиях национальных медиа.

Результаты исследования

Настоящее исследование посвящено сравнительному анализу оценочности и метафоричности в англоязычном медиадискурсе о России в двух национальных подкорпусах (Великобритания и Индия), сформированных на материале ведущих изданий за период 2022–2025 гг. Экстралингвистические характеристики подкорпусов предопределяют принципиально различные оценочные и метафорические стратегии.

Гипотеза исследования строилась на том, что в условиях острой глобальной поляризации национальные медиа не могут оставаться нейтральными: они неизбежно транслируют внешнеполитическую линию государства и его место в системе международных альянсов. Великобритания, будучи одним из наиболее последовательных и активных участников антироссийской коалиции, демонстрирует максимально негативную и идеологически заряженную риторику. Индия, напротив, сохраняя стратегическую автономию и исторически выстроенные партнерские отношения с Россией при одновременном давлении Запада, вынуждена выстраивать дискурс, избегающий резких оценочных полюсов.

Таким образом, контраст между этими двумя подкорпусами позволил проверить, в какой степени экстралингвистические переменные (позиция в конфликте, историческая память, экономические интересы) предопределяют выбор оценочных средств и метафорических сценариев, превращая медиатекст в инструмент мягкой силы и отражение государственной стратегии.

Первый этап исследования был посвящен измерению общей оценочности медиатекстов по пятибалльной шкале (от –2 «отрицательная» до +2 «положительная»). Для обеспечения объективности и воспроизводимости применялась гибридная методика, сочетающая нейросетевой анализ и опрос экспертной группы. Современным генеративным моделям (ChatGPT-4.0, Grok-4, DeepSeek) предъявлялся детализированный промпт, содержащий полное теоретическое обоснование категории оценочности (структура, явные и имплицитные маркеры, семантические, прагматические и логические механизмы), что позволило моделям учитывать не только лексические индикаторы, но и латентные оценочные стратегии, включая пресуппозиции, импликатуры и

фрейминговые эффекты. Результаты автоматической разметки проверялись группой из трех независимых экспертов-лингвистов (включая одного доктора филологических наук). Эксперты оценивали каждый текст по той же шкале, отвечая на вопрос: «Какую общую оценку России и ее действий формирует данный текст у типичного читателя?». Итоговая оценка принималась по правилу большинства; при трех разных оценках бралось среднее арифметическое (табл. 2).

Таблица 2. **Результаты анализа общей оценочности**
Table 2. **Overall evaluativity analysis results**

Оценка / Estimate	**Значение / Value**	**Британия (*n* = 50) / Britain (*n* = 50)**	**Индия (*n* = 50) / India (*n* = 50)**
Резко отрицательная	–2	15 (30 %)	8 (16 %)
Умеренно отрицательная	–1	26 (52 %)	22 (44 %)
Нейтральная	0	5 (10 %)	8 (16 %)
Умеренно положительная	+1	3 (6 %)	9 (18 %)
Положительная	+2	1 (2 %)	3 (6 %)
Итого отрицательная	**(–2 / –1)**	**41 (82 %)**	**30 (60 %)**
Среднее значение		**–1,22**	**–0,50**

Источник: составлено О.И. Калининым. / *Source:* compiled by Oleg I. Kalinin.

Проведенный анализ выявил резкий контраст в общей оценочности двух подкорпусов. Британский корпус демонстрирует ярко выраженный негативный профиль: 82 % текстов имеют отрицательную тональность, 30 % — резко отрицательную; средний показатель оценочности составляет –1,22, что отражает систематическое морально-этическое осуждение России. Индийский корпус, напротив, при преобладании отрицательной оценки (60 %) остается значительно более сдержанным и амбивалентным: доля резко отрицательных текстов в два раза ниже (16 %), а нейтральных и умеренно положительных — в три раза выше, чем в британском; средний показатель –0,50 свидетельствует о стратегии прагматического баланса и избегании резких оценочных полюсов. Разница в 0,72 балла между подкорпусами статистически значима и полностью подтверждает исходную гипотезу о том, что экстралингвистический контекст, то есть внешнеполитическая позиция страны, прямо определяет общую оценочность медиадискурса.

Количественный анализ трех индексов оценочности по двум подкорпусам выявил четко выраженные различия, также полностью соответствующие экстралингвистическому контексту (табл. 3).

Британский подкорпус демонстрирует значительно более высокую плотность оценочных средств (EDI = 7,86 против 6,58 в подкорпусе «Индия»), что отражает высокий уровень оценочного воздействия при подаче информации. Индекс интенсивности EII = 2,38 указывает на явное доминирование эмоционально окрашенных средств выражения оценки над рациональными, усиливая эффект осуждения и даже «демонизации» России. Индекс языковой

формы ELFI = 2,04 подтверждает преобладание эксплицитной оценки, то есть британские тексты в большей степени используют прямые, открытые оценочные суждения, не оставляющие места для неоднозначности и направленные на максимальное персуазивное воздействие.

Таблица 3. **Значение индексов оценочности**
Table 3. **Values of evaluativity indices**

Индекс / index	Великобритания / The United Kingdom	Индия / India	Разница / Difference
EDI (плотность)	7,86	6,58	+1,28
EII (интенсивность)	2,38	1,87	+0,51
ELFI (языковая форма)	2,04	1,96	+0,08

Источник: составлено О.И. Калининым. / *Source:* compiled by Oleg I. Kalinin.

В примере «I would say that Putinism is one more incarnation of Stalinism. He behaves like Stalin. His power is personalised, like in Stalin's time»[1] («Я бы сказал, что путинизм — это еще одно воплощение сталинизма. Он ведет себя как Сталин. Его власть персонифицирована, как во времена Сталина») автор напрямую отождествляет современный российский политический режим с эпохой Сталина, используя исторический образ тоталитарного лидера для формирования однозначно негативного отношения к нынешней власти в России. Основные лингвопрагматические средства формирования оценки включают лексику с ярко выраженной негативной коннотацией и сильным прагматическим потенциалом аналогии с историческим стереотипом («incarnation of Stalinism», активирующее пресуппозицию «воскрешения тоталитарного режима»; номинации «Stalinism» / «Stalin» как устоявшиеся негативные идеологемы, автоматически вызывающие ассоциативный ряд «тоталитаризм — репрессии — культ личности»). Также используются эксплицитные компаративные конструкции с высокой степенью категоричности («behaves like Stalin», «like in Stalin's time»), обеспечивающие прямой перенос негативного оценочного стереотипа с исторического прототипа на современный референт. С точки зрения синтаксиса используются короткие предложения, создающие иллюзию объективного констатирования фактов и одновременно повышающие перлокутивный эффект безапелляционной оценки.

В примере «Russia may well expand its sabotage, espionage and disinformation operations in Europe»[2] («Россия вполне может расширить свои диверсионные, шпионские и дезинформационные операции в Европе») автор имплицитно конструирует образ России как агрессивного дестабилизирую-

[1] Владимир Путин: современный царь России вступил в должность в 5 раз. Режим доступа: https://www.theguardian.com/world/2024/sep/22/putin-regime-will-collapse-without-warning-says-freed-gulag-dissident (дата обращения: 25.06.2025).
[2] Война на Украине: Россия начала новые атаки после обещания мира. Режим доступа: https://www.bbc.com/news/world-europe-60925713 (дата обращения: 26.06.2025).

щего актора мировой политики, систематически использующего нелегитимные и угрожающие методы воздействия на европейскую безопасность. Перлокутивный эффект усиливается за счет иллюзии объективного прогноза, который создается лексемами с негативной коннотацией («sabotage — espionage — disinformation operations»), и модальным глаголом «may», выполняющим функцию предупреждения и угрозы и повышающим воспринимаемую вероятность негативного сценария. Глагол «expand» создает пресуппозицию уже существующей агрессии и добровольного эскалационного выбора субъекта.

Индийский подкорпус, напротив, характеризуется меньшей плотностью (EDI = 6,58), умеренной интенсивностью (EII = 1,87) и почти сбалансированным соотношением эксплицитных и имплицитных форм (ELFI = 1,96). Это свидетельствует о стратегии сдержанного, дипломатически выверенного воздействия: оценка присутствует в большом объеме, но чаще маскируется под фактическую информацию без резких эмоциональных всплесков и прямых обвинений.

В примере из новости «Russians press assault on eastern Ukrainian city»[3] в предложении «It [the British Defence Ministry] said Russia likely was using such missiles because of a shortage of more accurate weapons» («Министерство обороны Великобритании заявило, что Россия, вероятно, использует такие ракеты из-за нехватки более точного оружия») автор опирается на заявление британского министерства, чтобы подчеркнуть предполагаемую неэффективность и ограниченность российского военного арсенала. Это формирует негативную оценку России как страны, вынужденной прибегать к устаревшим или неподходящим средствам в конфликте, подразумевая технологическую отсталость, дефицит ресурсов и, косвенно, стратегическую слабость. Лингвопрагматические средства здесь направлены на создание иллюзии объективного анализа, но с акцентом на дискредитацию российского военного потенциала.

Ключевыми элементами являются модальные средства выражения неопределенности и оценки, например, наречие «likely» (вероятно) вводит эпистемическую модальность, смягчая утверждение и позиционируя его как обоснованное предположение, а не как абсолютный факт. Это позволяет избежать прямых обвинений в фальсификации, сохраняя видимость аналитической объективности, но одновременно усиливает перлокутивный эффект, так как читатель воспринимает информацию как правдоподобную и авторитетную, поскольку она исходит от «экспертного» источника (British Defence Ministry). Также номинация «shortage» (нехватка) несет семантику дефицита и недостаточности, подразумевая системные проблемы в российской армии (экономические санкции, технологическое отставание или неэффективное планирование), а сравнительная конструкция «more accurate weapons» (более

[3] Российские войска продолжают наступление на город на востоке Украины. Режим доступа: https://indianexpress.com/article/world/russians-press-assault-on-eastern-ukrainian-city-8005452/ (дата обращения: 01.09.2023).

точное оружие) усиливает контраст, где Россия предстает в невыгодном свете по сравнению с «современными» стандартами (имплицитно — западными). Это усиливает идеологему «российской отсталости», часто встречающуюся в западных медиатекстах, и способствует формированию стереотипа о «варварских» методах ведения войны (как в контексте ударов по гражданским объектам, упомянутым в статье).

В примере «Russia is in favour of peace, but with the conditions that its stability is ensured and the objectives of the special military operation are fulfilled»[4] («Россия выступает за мир, но при условии обеспечения стабильности и выполнения задач специальной военной операции») автор опирается на заявление пресс-секретаря Кремля Дмитрия Пескова, чтобы представить российскую позицию как конструктивную и обоснованную в контексте переговоров о мире. Это формирует позитивную оценку России как ответственной державы, стремящейся к миру, но не готовой к компромиссам, угрожающим ее суверенитету, и противопоставляет ее «разумный» подход агрессивным требованиям оппонентов (например, президента Украины В. Зеленского, чьи слова отвергаются как «роковая ошибка»). Лингвопрагматические средства здесь направлены на создание образа России как значимого актора в международных отношениях, принципиального, ориентированного на стабильность и защиту национальных интересов, что усиливает идеологему «суверенной державы».

Ключевыми элементами являются контрастные конструкции. Так, союзная конструкция «but with the conditions» (но при условии) вводит антитезу, где первая часть («in favour of peace») активирует позитивную пресуппозицию о миролюбии России, а вторая определяет военные действия как необходимую меру для обеспечения («ensured») стабильности, перенося оценочный стереотип «защитника интересов» на референт. Лексика с позитивной коннотацией усиливает это: номинации «favour of peace» (выступает за мир) и «stability is ensured» (стабильность обеспечена) несут значение активности и стабильности, вызывая ассоциации с рациональностью и предсказуемостью, в то время как «objectives... are fulfilled» (задачи… выполнены) апеллирует к пресуппозиции законности применения вооруженных сил. В итоге такие средства способствуют формированию стереотипа о России как о державе, чьи действия мотивированы не экспансией, а необходимостью, что в медиатекстах часто служит инструментом для смягчения негативного восприятия конфликта.

Таким образом, количественные индексы подтверждают качественные различия: британский медиадискурс использует оценочность как инструмент открытой идеологической конфронтации, тогда как индийский — как механизм поддержания стратегического баланса, в котором критика России

[4] Кремль заявил, что принуждение России к миру станет «роковой ошибкой». Режим доступа: https://timesofindia.indiatimes.com/world/europe/kremlin-says-forcing-russia-into-peace-would-be-fatal-mistake/articleshow/113663921.cms (дата обращения: 01.11.2024).

присутствует, но подается в смягченной, рациональной и частично имплицитной форме, с рациональным объяснением позиции России в этом вооруженном конфликте.

Количественное измерение метафоричности по трем индексам выявило существенные различия между подкорпусами, полностью соответствующие общей оценочной картине и экстралингвистическому контексту (табл. 4).

Таблица 4. **Значения индексов метафоричности**
Table 4. **Values of metaphor power indices**

Индекс / Index	Великобритания / The United Kingdom	Индия / India	Разница / Difference
MDI (плотность метафор)	4,09	3,30	+0,79
MII (интенсивность)	2,01	1,33	+0,68
MLFI (языковая форма)	1,66	1,35	+0,31

Источник: составлено О.И. Калининым. / *Source:* compiled by Oleg I. Kalinin.

Британский подкорпус характеризуется значительно более высокой метафорической насыщенностью (MDI = 4,09) и заметно большей интенсивностью метафорических образов (MII = 2,01). Это означает преобладание ярких новых метафор, обладающих сильным эмоциональным и персуазивным потенциалом. Индекс языковой формы MLFI = 1,66 показывает, что хотя непрямые метафоры все еще доминируют, прямые и структурные конструкции используются существенно чаще, чем в индийском корпусе, усиливая открытость идеологического воздействия.

В примере «Russia's advances are coming at a horrifying cost. [...] "They are really pushing everyone into the grinder: drivers, cooks, builders"»[5] («Продвижение России дается ужасающей ценой. […] Они буквально бросают всех в мясорубку: водителей, поваров, строителей») центральная метафора «pushing everyone into the grinder» относится к концептуальной модели «ВОЙНА — ЭТО МЯСОРУБКА / ПРОМЫШЛЕННАЯ ПЕРЕРАБОТКА МЯСА». Сфера-источник — бытовая/промышленная мясорубка, сфера-цель — современные боевые действия на линии фронта. Метафора активирует фрейм «индустриального уничтожения живой силы», солдаты представляются «безликим сырьем, которое автоматически и массово перерабатывается в «отходы» ради конечного продукта». Перечисление «drivers, cooks, builders» (водители, повара, строители) усиливает эффект: в «мясорубку» бросают не профессиональных военных, а всех, в том числе явно непригодных для этого военнослужащих, что подчеркивает варварскую неразборчивость и цинизм командования. В итоге Россия концептуализируется как «бездушный, жестокий оператор индустриальной машины смерти,

[5]Украина находится в обороне — в военном, экономическом и дипломатическом плане. Режим доступа: https://www.economist.com/briefing/2024/09/26/ukraine-is-on-the-defensive-militarily-economically-and-diplomatically (дата обращения: 26.11.2025).

сознательно жертвующий огромным количеством собственных граждан ради минимального результата», что является классическим образом антигу манного агрессора, отсылающим к самым негативным историческим стереотипам. Метафора вызывает физиологическое отвращение и моральное осуждение, резко усиливая негативную оценку России.

В примере «Putin is inflicting appalling devastation and horror on Ukraine — bombing apartment buildings and maternity wards... these are atrocities. It is an outrage to the world» («Путин сеет ужасающие разрушения и страх в Украине — бомбит жилые дома и родильные отделения... это зверства. Это вызов всему миру»)[6] две ключевые оценочные метафоры создают предельно негативный образ В.В. Путина и России. Первая — это непрямая глагольная метафора «is inflicting appalling devastation and horror» (сеет ужасающие разрушения и ужас), где сфера-источник — апокалиптическое бедствие — переносится на сферу-цель «военные действия». Россия в таком метафорическом контексте предстает демонической фигурой, сознательно распространяющей хаос и страдания. Вторая метафора-номинация «these are atrocities» (это зверства), в рамках которой «преступления против человечности / зверстав» переносятся на способ ведения Россией боевых действий. Упоминание бомбардировок роддомов и жилых домов усиливает отсылки к самым страшным историческим примерам, выводя действия России за рамки войны в область абсолютного морального зла.

Индийский подкорпус демонстрирует умеренную плотность метафор (MDI = 3,30), минимальную интенсивность (MII = 1,33, близко к конвенциональным, «стертым» образам) и ярко выраженное преобладание непрямых форм (MLFI = 1,35). Такая конфигурация свидетельствует о сознательной стратегии сдерживания: метафоры активно присутствуют, но выбираются максимально нейтральные, культурно-универсальные и имплицитные, что позволяет сохранять видимость объективности и институционального нейтралитета при сохранении скрытого оценочного эффекта.

В примере «India considers Russia a time-tested ally from the Cold War era with key cooperation in defense, oil, nuclear energy and space exploration»[7] («Индия рассматривает Россию как проверенного временем союзника со времен Холодной войны, с ключевым сотрудничеством в сфере обороны, нефти, ядерной энергетики и космоса») оценка передается через прямую структурную метафору «time-tested ally» (проверенный временем союзник). Метафора создает ярко позитивную оценку России, которая предстает надежным, выдержавшим все исторические испытания партнером, чья цен-

[6] Конфликт в Украине: Байден назвал Путина «военным преступником». Режим доступа: https://www.bbc.com/news/world-us-canada-60773626 (дата обращения: 26.11.2025).

[7] Что нужно знать об отношениях Индии и России. Режим доступа: https://www.thehindu.com/news/national/what-to-know-about-indias-ties-with-russia/article66996480.ece (дата обращения: 26.11.2025).

ность не зависит от текущей конъюнктуры и западного давления. Упоминание «from the Cold War era» усиливает образ России как великой державы, сохраняющей свои позиции на протяжении долгого времени.

Перечисление стратегических сфер (оборона, нефть, ядерная энергетика, космос) подчеркивает глубину и незаменимость партнерства. В итоге метафора конструирует образ России как долгосрочного союзника, которому можно доверять в условиях глобальной турбулентности, что прямо противостоит западному нарративу об «изоляции» России и формирует ее позитивный имидж надежного и проверенного временем глобального игрока.

В примере «The Czechs are looking to acquire from countries outside the European Union at least 800,000 artillery shells that Ukraine badly needs. The war has drained stockpiles in Europe, the United States and Russia»[8] («Чехи ищут возможность приобрести у стран за пределами Европейского Союза как минимум 800 тыс. артиллерийских снарядов, которые крайне необходимы Украине. Война истощила запасы в Европе, США и России») — ключевая оценочная метафора — это «drained stockpiles» (истощила запасы), реализующая концептуальную модель «ВОЙНА — ЭТО ИСТОЩЕНИЕ РЕСУРСОВ».

Сфера-источник реализуется глаголом «drain» (истощать, иссушать, дренаж), кого подразумевает постепенное, неконтролируемое выкачивание жизненно важной энергии, приводящее к ослаблению и истощению. Метафора создает нейтральную оценку России, подчеркивая универсальность последствий войны, которая предстает как процесс, который «высасывает» ресурсы одинаково из Европы, США и России, активируя пресуппозицию о ее разрушительной природе для всех участников.

Таким образом, сравнительный анализ медиадискурсов Великобритании и Индии выявил существенные различия в использовании оценочных средств и метафор при репрезентации России. Британские СМИ используют метафору как мощное оружие открытой эмоциональной и идеологической мобилизации, тогда как индийские издания превращают ее в тонкий инструмент рационализированного, латентного фрейминга, соответствующий политике стратегической автономии и дипломатического баланса.

Корреляционный анализ показателей индексов метафоричности и оценочности по каждому из 50 текстов подтверждает тесную связь между выражением оценки и использованием метафор в британском дискурсе (соотношение EDI–MDI = 0,76; соотношение EII–MI = 0,58), свидетельствующую о том, что яркие и прямые метафоры выступают основным инструментом усиления эмоциональной негативной оценки России. В индийском дискурсе сохраняется корреляция плотности оценки и использования метафор (соотношение EDI–MDI = 0,51), однако связь по интенсивности оценочности

[8] Россия продолжает оказывать давление на линии фронта до получения Украиной помощи от Запада. Режим доступа: https://www.dailypioneer.com/2024/world/russia-keeps-up-frontline-pressure-before-ukraine-receives-a-boost-from-western-military-aid.html (дата обращения: 26.11.2025).

и метафоричности и по языковой форме выражение оценки и метафор слабая (соотношение EII–MII = 0,10, соотношение ELFI–MLFI = 0,13), что отражает компенсаторную стратегию: метафоры используются не для эмоционального воздействия, а для рационально-оценочного, при сохранения общей позитивно-нейтральной тональности по отношению к России.

Первый этап исследования, предоставивший количественные показатели характера общей оценки в изученных текстах (от отрицательной (–2) до положительной (+2)) позволил также провести корреляционный анализ соотношения разных аспектов использования метафор и характером оценки (отрицательная / положительная) (табл. 5).

Таблица 5. **Результаты корреляционного анализа индексов метафоричности и оценки**
Table 5. **Results of the correlation analysis between metaphoricity indices and evaluation**

	MDI	**MII**	**MLFI**
Характер оценки (+/--)	–0,5734	–0,6367	–0,1477

Источник: составлено О.И. Калининым. / *Source:* compiled by Oleg I. Kalinin.

Проведенное исследование показало, что характер оценки России в медиадискурсе (положительный или отрицательный) имеет ярко выраженную и статистически достоверную связь с количественными и эмоционально-смысловыми характеристиками метафор, а не с их грамматической или структурной формой.

1. Плотность метафор (MDI): чем чаще в тексте встречаются метафорические выражения, тем более негативной оказывается общая оценка России (коэффициент корреляции r = –0,5734). Это умеренная отрицательная связь: при переходе от нейтрально-положительных материалов к критическим количество метафор в расчете на 100 словоупотреблений заметно возрастает.

2. Интенсивность метафор (MII): наблюдается наиболее значимый показатель корреляции (r = –0,6367), который указывает на сильную обратную зависимость: чем ярче, экспрессивнее и эмоционально заряженнее метафоры, тем более негативной становится оценка. Именно этот параметр выступает главным лингвистическим индикатором антироссийской направленности текста.

3. Языковая форма метафор (MLFI) практически не коррелирует с характером оценки (r = –0,1477). Это означает, что для создания негативного образа России одинаково используются как прямые, эксплицитные метафоры, так и непрямые, скрытые, встроенные в синтаксис или лексику.

Таким образом, негативная репрезентация России в медиадискурсе достигается прежде всего за счет массированного и эмоционально насыщенного метафорического представления референта. Чем больше метафор и чем сильнее они связаны с воздействием на эмоциональную сферу, тем более однозначно отрицательным становится образ страны и ее руководства. При

этом языковая форма метафоры играет второстепенную роль и используется скорее в сугубо стилистических целях.

Заключение

Проведенное исследование показало, что метафора выступает одним из главных лингвопрагматических инструментов формирования оценки образа России в англоязычном медиадискурсе, при этом ее роль и характер существенно различаются в зависимости от геополитической позиции продуцента. Британский медиадискурс демонстрирует высокую плотность и интенсивность оценочных средств, преимущественно эксплицитных и эмоционально насыщенных, что служит реализации стратегии открытой идеологической конфронтации и демонизации референта. Индийский дискурс, напротив, характеризуется умеренными показателями и преобладанием скрытых, рациональных и конвенциональных метафор, что соответствует политике стратегической автономии и дипломатического баланса.

Корреляционный анализ подтвердил, что негативная оценка России прямо связана с количеством и эмоциональной интенсивностью метафор, тогда как языковая форма метафоры не оказывает значимого влияния на знак оценки. Таким образом, именно массированное использование ярких, аффективно заряженных метафор является ключевым маркером антироссийской направленности текста.

Разработанная методика интегральных индексов и гибридного анализа позволила количественно подтвердить качественные различия и установить степень синергетического взаимодействия оценочности и метафоричности. Теоретическая значимость работы заключается в уточнении механизмов метафорического речевого воздействия в медиадискурсе, в расширении понятийного аппарата оценочности как интегративной дискурсивной категории и в эмпирическом подтверждении ее зависимости от экстралингвистических факторов. Полученные результаты вносят вклад в развитие лингвопрагматики, критического анализа дискурса и теории мягкой силы.

Перспективы дальнейших исследований связаны с расширением корпуса на другие страны Глобального Юга (Бразилия, Турция, ЮАР, ОАЭ), включением неанглоязычных медиа (русскоязычных, китайских, арабских), а также с применением разработанной методики к другим конфликтным дискурсам, что позволит построить типологию национальных дискурсивных стратегий в условиях современного информационного противоборства.

Список литературы

1. *Dinnie K.* Nation Branding: Concepts, Issues, Practice. London; New York : Routledge, 2015. https://doi.org/10.4324/9781315773612
2. *Papadopoulos N., Hamzaoui-Essoussi L.* Place Images and Nation Branding in the African Context: Challenges, Opportunities, and Questions for Policy and Research // Africa Journal of Management. 2015. Vol. 1. Iss. 1. P. 54–77. https://doi.org/10.1080/23322373.2015.994423

3. *van Dijk T.A.* Critical Discourse Analysis // The Handbook of Discourse Analysis / ed. by D. Tannen, H. E. Hamilton, D. Schiffrin. Malden, MA; Oxford : Wiley-Blackwell, 2015. P. 466–485. https://doi.org/10.1002/9781118584194.ch22
4. *Wodak R.* Critical Discourse Analysis, Discourse-Historical Approach // The International Encyclopedia of Language and Social Interaction / ed. by K. Tracy, T. Sandel, C. Ilie. Malden, MA; Oxford : Wiley-Blackwell, 2015. P. 1–9. https://doi.org/10.1002/9781118611463.wbielsi116
5. Информационная война: формы ведения и методы лингвистического анализа / Н.А. Ахренова, Е.В. Голубцова, Н.В. Зененко [и др.]. М. : ФЛИНТА, 2025. EDN: NGWNQO.
6. Forging the World: Strategic Narratives and International Relations / ed. by A. Miskimmon, B. O'Loughlin, L. Roselle. Ann Arbor : University of Michigan Press, 2017. https://doi.org/10.3998/mpub.6504652
7. New Directions in India's Foreign Policy: Theory and Praxis / ed. by H. V. Pant. Cambridge : Cambridge University Press, 2019. https://doi.org/10.1017/9781108562850
8. *Bednarek M.* Investigating evaluation and news values in news items that are shared through social media // Corpora. 2016. Vol. 11. Iss. 2. P. 227–257. https://doi.org/10.3366/cor.2016.0093
9. *Марьянчик В.А.* Аксиологичность и оценочность медиа-политического текста : монография. М. : URSS, 2013.
10. *Зубова Т.Б., Калинин О.И.* Оценочность как дискурсивная характеристика: свойства и типология // Филология: научные исследования. 2025. № 2. С. 172–187. https://doi.org/10.7256/2454-0749.2025.2.72801 EDN: OHYPEH
11. *Ma M.D., Bowden K., Wu J., Walker M.* Implicit Discourse Relation Identification for Open-domain Dialogues // Proceedings of the 57th Annual Meeting of the Association for Computational Linguistics (Florence, Italy, July 28 — August 2, 2019). Stroudsburg, PA : Association for Computational Linguistics, 2019. P. 666–672. https://doi.org/10.18653/v1/P19-1065
12. *Заппеттини Ф., Понтон Д.М., Ларина Т.В.* Эмоционализация современного медиадискурса: исследовательская повестка дня // Russian Journal of Linguistics. 2021. Т. 25. № 3. С. 586—610. https://doi.org/10.22363/2687-0088-2021-25-3-586-610 EDN: NYDDJB
13. *Brugman B.C., Burgers C., Vis B.* Metaphorical framing in political discourse through words vs. concepts: a meta-analysis // Language and Cognition. 2019. Vol. 11. Iss. 1. P. 41–65. https://doi.org/10.1017/langcog.2019.5
14. *Citron F.M.M., Goldberg A.E.* Metaphorical Sentences Are More Emotionally Engaging than Their Literal Counterparts // Journal of Cognitive Neuroscience. 2014. Vol. 26. Iss. 11. P. 2585–2595. https://doi.org/10.1162/jocn_a_00654
15. *Калинин О.И., Игнатенко А.В.* Сравнительный анализ использования метафор в русских, английских и китайских медиатекстах информационного и воздействующего характера // Вестник Российского университета дружбы народов. Серия: Теория языка. Семиотика. Семантика. 2022. Т. 13. № 4. С. 1062–1082. https://doi.org/10.22363/2313-2299-2022-13-4-1062-1082 EDN: IKOTXX
16. *Zibin A., Solopova O.A.* Metaphors across languages, cultures and discourses: A research agenda // Russian Journal of Linguistics. 2024. Vol. 28. Iss. 1. P. 7–32. https://doi.org/10.22363/2687-0088-37837 EDN: LRWDXE
17. Ammar M. Migration Narrative(s) in the Arab Media: A Metaphorical Framing Analysis // Linguæ & — Journal of Modern Languages and Cultures. 2023. Vol. 23. Iss. 1. P. 87–105. https://doi.org/10.14276/l.v23i1.3535
18. *Ананченко О.Г.* Оценочные метафоры в газетных текстах начала XXI в. : лингвопрагматический аспект : автореф. дис. … канд. филол. наук : 10.02.01. Майкоп, 2011. EDN: ZOGRLB
19. *Bowdle B.F., Gentner D.* The Career of Metaphor // Psychological Review. 2005. Vol. 112. Iss. 1. P. 193–216. https://doi.org/10.1037/0033-295X.112.1.193

20. *Steen G.J.* Slowing Metaphor Down: Elaborating Deliberate Metaphor Theory. Amsterdam; Philadelphia : John Benjamins Publishing Company, 2023. xv. // Metaphor and Symbol. 2024. Vol. 39. Iss. 3. P. 217–221. https://doi.org/10.1075/celcr.26 EDN: KCGNZT
21. *Калинин О.И.* Функциональный потенциал метафоры как средства речевого воздействия в дискурсе // Когнитивные исследования языка. 2022. № 2(49). С. 134–139. EDN: WUHMOL

References

1. Dinnie, K. (2015). *Nation Branding: Concepts, Issues, Practice*. London; New York: Routledge. https://doi.org/10.4324/9781315773612
2. Papadopoulos, N., & Hamzaoui-Essoussi, L. (2015). Place Images and Nation Branding in the African Context: Challenges, Opportunities, and Questions for Policy and Research. *Africa Journal of Management*, *1*(1), 54–77. https://doi.org/10.1080/23322373.2015.994423
3. van Dijk, T.A. (2015). Critical Discourse Analysis. In: D. Tannen, H.E. Hamilton, & D. Schiffrin (eds.) *The Handbook of Discourse Analysis* (pp. 466–485). Wiley-Blackwell. https://doi.org/10.1002/9781118584194.ch22
4. Wodak, R. (2015). Critical Discourse Analysis, Discourse-Historical Approach. In: K. Tracy, T. Sandel, & C. Ilie (eds.) *The International Encyclopedia of Language and Social Interaction* (pp. 1—9). Wiley-Blackwell. https://doi.org/10.1002/9781118611463.wbielsi116
5. Akhrenova, N.A., Golubtsova, E.V., Zenenko, N.V., et al. (2025). *Information Warfare: Forms of Conduct and Methods of Linguistic Analysis.* Moscow: Flinta. (In Russ.). EDN: NGWNQO
6. Miskimmon, A., O'Loughlin, B., & Roselle, L. (eds.) (2017). *Forging the World: Strategic Narratives and International Relations*. University of Michigan Press. https://doi.org/10.3998/mpub.6504652
7. Pant, H.V. (Ed.). (2019). *New Directions in India's Foreign Policy: Theory and Praxis.* Cambridge: Cambridge University Press. https://doi.org/10.1017/9781108562850
8. Bednarek, M. (2016). Investigating Evaluation and News Values in News Items that are Shared through Social Media. *Corpora*, *11*(2), 227–257. https://doi.org/10.3366/cor.2016.0093
9. Maryanchik, V. A. (2013). *Axiologicality and Evaluativity of Media-Political Text.* Moscow: URSS. (In Russ.).
10. Zubova, T.B., & Kalinin, O.I. (2025). Evaluativity as a Discursive Characteristic: Properties and Typology. *Filologiya: Nauchnye issledovaniya*, 2, 172–187. (In Russ.). https://doi.org/10.7256/2454-0749.2025.2.72801 EDN: OHYPEH
11. Ma, M.D., Bowden, K., Wu, J., & Walker, M. (2019). Implicit Discourse Relation Identification for Open-Domain Dialogues. In: *Proceedings of the 57th Annual Meeting of the Association for Computational Linguistics* (pp. 666—672). Association for Computational Linguistics. https://doi.org/10.18653/v1/P19-1065
12. Zappettini, F., Ponton, D.M., & Larina, T.V. (2021). Emotionalisation of Contemporary Media Discourse: A Research Agenda. *Russian Journal of Linguistics*, *25*(3), 586–610. (In Russ.). https://doi.org/10.22363/2687-0088-2021-25-3-586-610 EDN: NYDDJB
13. Brugman, B.C., Burgers, C., & Vis, B. (2019). Metaphorical Framing in Political Discourse through Words vs. Concepts: A Meta-Analysis. *Language and Cognition*, *11*(1), 41–65. https://doi.org/10.1017/langcog.2019.5
14. Citron, F.M.M., & Goldberg, A.E. (2014). Metaphorical Sentences are More Emotionally Engaging than Their Literal Counterparts. *Journal of Cognitive Neuroscience*, *26*(11), 2585–2595. https://doi.org/10.1162/jocn_a_00654
15. Kalinin, O.I., & Ignatenko, A.V. (2022). Comparative Analysis of Metaphor Use in Russian, English and Chinese Media Texts of Informative and Persuasive Nature. *RUDN Journal of Language Studies, Semiotics and Semantics*, *13*(4), 1062–1082. (In Russ.). https://doi.org/10.22363/2313-2299-2022-13-4-1062-1082 EDN: IKOTXX

16. Zibin, A., & Solopova, O.A. (2024). Metaphors Across Languages, Cultures and Discourses: A Research Agenda. *Russian Journal of Linguistics*, *28*(1), 7–32. https://doi.org/10.22363/2687-0088-37837 EDN: LRWDXE
17. Ammar, M. (2023). Migration Narrative(S) in the Arab Media: A Metaphorical Framing Analysis. *Linguæ & – Journal of Modern Languages and Cultures*, *23*(1), 87–105. https://doi.org/10.14276/l.v23i1.3535
18. Ananchenko, O.G. (2011). Evaluative Metaphors in Newspaper Texts of the Early 21st Century: Linguopragmatic Aspect [PhD Thesis]. Maykop: Adyghe State University publ. (In Russ.). EDN: ZOGRLB
19. Bowdle, B.F., & Gentner, D. (2005). The Career of Metaphor. *Psychological Review*, *112*(1), 193–216. https://doi.org/10.1037/0033-295X.112.1.193
20. Steen, G.J. (2023). Slowing Metaphor Down: Elaborating Deliberate Metaphor Theory. Philadelphia: John Benjamins Publishing Company. *Metaphor and Symbol*, *39*(3), 217–221. https://doi.org/10.1075/celcr.26 EDN: KCGNZT
21. Kalinin, O.I. (2022). Functional Potential of Metaphor as a Means of Speech Influence in Discourse. *Cognitive Studies of Language*, *2*(49), 134–139. (In Russ.). EDN: WUHMOL

Сведения об авторах:
Калинин Олег Игоревич, доктор филологических наук; доцент, профессор кафедры китайского языка переводческого факультета, Московский государственный лингвистический университет (119034, Российская Федерация, г. Москва, ул. Остоженка, д. 38 с. 1); доцент 36 кафедры института иностранных языков, Военный университет им. князя Александра Невского (111033, Российская Федерация, г. Москва, ул. Волочаевская, д. 3); старший научный сотрудник Управления научной и инновационной деятельности, Южно-Уральский государственный университет (национальный исследовательский университет) (454080, Российская Федерация, г. Челябинск, просп. В.И. Ленина, д. 76.); *сфера научных интересов*: теория дискурса, когнитивная лингвистика, метафорология, теория речевого воздействия; *e-mail:* okalinin.lingua@gmail.com
ORCID: 0000-0002-1807-8370; Scopus AuthorID: 57223031724; SPIN-код: 7494-7512.

Information about the authors:
Oleg I. Kalinin, Dr.Sc. (Philology); Professor at the Chinese Language Department, Faculty of Translation and Interpreting, Moscow State Linguistic University (38 build 1 Ostozhenka Str., Moscow, Russian Federation, 119034); Associate Professor at the 36th Department, Institute of Foreign Languages, Military University (3 Volochayevskaya Str., Moscow, Russian Federation, 111033); Senior Researcher of the Department of Scientific and Innovative Activity, South Ural State University (National Research University) (76 Lenin prospekt, Chelyabinsk, Russian Federation, 454080); *Research interests*: theory of discourse, cognitive linguistics, metaphorology, theory of speech influence; *e-mail:* okalinin.lingua@gmail.com
ORCID: 0000-0002-1807-8370; Scopus AuthorID: 57223031724; SPIN-code: 7494-7512.

RUDN Journal of Language Studies, Semiotics and Semantics
2313-2299 (Print), ISSN 2411-1236 (Online)
Вестник Российского университета дружбы народов.
Серия: ТЕОРИЯ ЯЗЫКА. СЕМИОТИКА. СЕМАНТИКА

2026 Vol. 17 No. 2 648–667
http://journals.rudn.ru/semiotics-semantics

DOI: 10.22363/2313-2299-2026-17-2-648-667
EDN: LYRRPL

Научная статья / Research Article

Национально-культурные коды в мексиканском политическом дискурсе: инаугурация сквозь призму лингвосемиотики

М.В. Ларионова ✉, С.К. Мосина

Московский государственный институт международных отношений (университет)
МИД РФ, *Москва, Российская Федерация*
✉larionova.m@list.ru

Аннотация. *Актуальность* исследования определяется необходимостью изучения кодов национальной мексиканской культуры как неотъемлемой части цивилизационной идентичности латиноамериканского мира, объединяющего через доминанту «единства в многообразии» — *unidad en la diversidad* — вековую мудрость, исторические и социальные ценности и традиции автохтонной, иберийской и африканской цивилизационных матриц как основы латиноамериканского цивилизационного и социокультурного единства. Сохранение, возрождение, овладение знаниями о народной культуре социума, о духовных концептах и материальных основах лингвокультурного сообщества находятся в центре внимания современных научных ономасиологических и семиологических исследований, рассматривающих культуру как знаковую и коммуникативную систему. *Цель* статьи — сквозь призму политического дискурса провести интегральный анализ лингвосемиотической составляющей вербальных и невербальных кодов церемонии инаугурации президента Мексики К. Шейнбаум, отражающих особую знаковую систему культурно-исторических и социальных ценностей мексиканского национально-культурного сознания, которую понимают и разделяют его носители. *Научная новизна* рассматриваемой в статье проблематики обусловлена системным подходом к изучению в рамках лингвосемиотической и лингвокультурологической парадигм дискурсивных исследований вербальных и невербальных культурных кодов мексиканской политической коммуникации, оказывающих воздействие на дискурсивное и социальное поведение аудитории-адресата в соответствии с интенцией политического актора-адресанта. Незнание особенностей менталитета и национального характера, обычаев, традиций, специфических языковых номинаций и устойчивых выражений препятствует успешной коммуникации, служит причиной для ошибочного понимания и интерпретации знаков культуры, приводит к коммуникативным и профессиональным неудачам специалистов-международников. Особое внимание в исследовании уделяется невербальным компонентам политического дискурса, использование которых становится устойчивой тенденцией современного информационно-коммуникационного пространства. Артефакты, символы, вестиментарный код, цвет, а также визуальные, звуковые, графические составляющие в конденсированной форме транслируют дополнительные смыслы, оказывающие устойчивый суггестивный эффект и

© Ларионова М.В., Мосина С.К., 2026
This work is licensed under a Creative Commons Attribution-NonCommercial 4.0 International License
https://creativecommons.org/licenses/by-nc/4.0/legalcode

имеющие особое значение для реципиентов, разделяющих общие для этнокультурного сообщества знания.

Ключевые слова: Латинская Америка, Мексика, языковая картина мира, семиотический код, вербальные и невербальные компоненты дискурса

Вклад авторов: равнозначен на все этапах исследования и подготовки статьи к публикации.

Конфликт интересов: авторы заявили об отсутствии конфликтов интересов.

История статьи: дата поступления: 20.01.2026; дата приема в печать: 10.02.2026.

Для цитирования: *Ларионова М.В.*, *Мосина С.К.* Национально-культурные коды в мексиканском политическом дискурсе: инаугурация сквозь призму лингвосемиотики // Вестник Российского университета дружбы народов. Серия: Теория языка. Семиотика. Семантика. 2026. Т. 17. № 2. С. 648–667. https://doi.org/10.22363/2313-2299-2026-17-2-648-667 EDN LYRRPL

National-Cultural Codes in Mexican Political Discourse: Inauguration through the Prism of Linguosemiotics

Marina V. Larionova ✉, Sofia K. Mosina

MGIMO University, *Moscow*, *Russian Federation*
✉larionova.m@list.ru

Abstract. The relevance of the study is explained by the need to consider the codes of national Mexican culture as an integral part of the Latin America civilization identity, merging historical and social values and traditions of indigenous, Iberian and African civilization matrices as the basis of Latin American civilization, society and culture by means of an ancient dominant — unity in diversity (*unidad en la diversidad*). The preservation, revival, and mastery of knowledge about the folk culture of society, its spiritual concepts, and material foundations of the linguistic and cultural community are the focus of modern scientific onomasiological and semiological studies that consider culture as a communicative system of signs. The purpose of the article is to carry out an integral analysis of the linguistic and semiotic component of the verbal and non-verbal codes of the inauguration ceremony of C. Sheinbaum, the President of Mexico, in the light of political discourse, reflecting a special sign system of cultural, historical and social values of the Mexican national identity, which is understood and shared by the Mexicans. The scientific novelty of the issues considered in the article is due to a systematic approach to the study of the linguistic and semiotic and linguocultural paradigms of discursive research on the verbal and non-verbal cultural codes of Mexican political communication, which affect the discursive and social behavior of the audience in accordance with the intention of the political actor. Ignorance of the peculiarities of mentality and national identity, customs, traditions, specific language nominations and idiomatic expressions hinders successful communication, causes misunderstanding and wrong interpretation of cultural signs, which leads to communicative and professional failures of international specialists. The study pays special attention to non-verbal components of political discourse, the use of which is becoming a trend in the modern communication. Artifacts, symbols, vestimentary code, color, as well as visual, sound, and graphic components convey additional meanings in a condensed form that have a powerful suggestive effect and are of particular importance to recipients who share cultural community values.

Keywords: Latin America, Mexico, linguistic worldview, semiotic code, verbal and non-verbal components of discourse

Authors' contribution: The authors contributed equally to this article and preparation of the article text.

Conflict of interests: The authors stated that there were no conflicts of interest.

Article history: Received: 20.01.2026; Accepted: 10.02.2026.

For citation: Larionova, M.V., & Mosina, S.K. (2026). National-Cultural Codes in Mexican Political Discourse: Inauguration through the Prism of Linguosemiotics. *RUDN Journal of Language Studies, Semiotics and Semantics, 17*(2), 648–667. (In Russ.). https://doi.org/10.22363/2313-2299-2026-17-2-648-667 EDN LYRRPL

Введение

Лингвокультурное пространство, термин, предложенный В.В. Красных [1. С. 69], включает в себя не только язык как инструмент общения и познания, но и национальные обычаи, культурные традиции, религиозные нормы и ценности, общественно-политическую, правовую и экономическую системы, способы мышления и формы бытия, а также иные цивилизационные коды, которые разделяет национальное сообщество. Ценностно-смысловые доминанты лингвокультурного пространства Мексики как одной из самых больших и значимых стран латиноамериканского цивилизационного сообщества имплицируют множество культурно-исторических, социально-политических, языковых факторов и отражают менталитет и особенности национального характера ее жителей.

Изучению вербальных и невербальных проявлений культуры народа в дискурсивной среде и художественных текстах, особенностям миропонимания и способам концептуализации действительности, а также установлению механизмов связи языка и культуры и их функционированию в рамках лингвокультурологии как научной парадигмы посвящены исследования Н.Ф. Алефиренко [2], В.В. Воробьева и др. [3], В.А. Масловой [4], Ю.Л. Оболенской [5. С. 39–45], И.Г. Ольшанского [6], А.Т. Хроленко [7] и других ученых.

Актуальность исследования определяется необходимостью изучения кодов национальной мексиканской культуры как неотъемлемой части цивилизационной идентичности латиноамериканского мира, объединяющего через доминанту «единства в многообразии» — *unidad en la diversidad*, которая представляет собой «сущностную характеристику» латиноамериканского мировидения [8. С. 35; 9. С. 39–63], вековую мудрость, исторические и социальные ценности и традиции автохтонной, иберийской и африканской цивилизационных матриц, составляющие основу латиноамериканского цивилизационного и социокультурного единства.

Сохранение, возрождение и овладение знаниями о народной культуре социума, о духовных концептах и материальных основах лингвокультурного сообщества, воплощенных в языке, находятся в центре внимания современных научных ономасиологических и семиологических исследований, «открывших перспективы познания культуры как знаковой и коммуникативной системы» [10. С. 27]. Следует особо отметить, что для стран латиноамериканского региона культура как детерминационная модель мышления, поведения,

восприятия и бытия через систему знаков и значений функционирует как коммуникационный механизм, который благодаря полифонии языковых и невербальных семиотических кодов сохраняет духовные ценности и смыслы, значимые для латиноамериканской цивилизации [8; 11. С. 16–40; 12. С. 41–67; 13. P. 56–76; 14. P. 15–58].

Цель статьи — сквозь призму политического дискурса провести интегральный анализ лингвосемиотической составляющей вербальных и невербальных кодов церемонии инаугурации президента Мексики К. Шейнбаум, отражающих особую знаковую систему культурно-исторических и социальных ценностей мексиканского мифологического сознания, которую понимают и разделяют его носители.

Научная новизна рассматриваемой в статье проблематики обусловлена системным подходом к изучению в рамках лингвосемиотической и лингвокультурологической парадигм дискурсивных исследований вербальных и невербальных культурных кодов мексиканской политической коммуникации, оказывающих воздействие на дискурсивное и социальное поведение аудитории-адресата в соответствии с интенцией политического актора-адресанта. Незнание особенностей менталитета и национального характера, обычаев, специфических языковых номинаций и устойчивых выражений препятствует успешной коммуникации, служит причиной ошибочного понимания и интерпретации знаков культуры, приводит к коммуникативным и профессиональным неудачам специалистов-международников.

Особое внимание в исследовании уделяется невербальным компонентам политического дискурса, использование которых становится устойчивой тенденцией современного информационно-коммуникационного пространства. Артефакты, символы, вестиментарный код, цвет, а также визуальные, звуковые, графические составляющие в конденсированной форме транслируют дополнительные смыслы, оказывающие устойчивый суггестивный эффект и имеющие особое значение для реципиентов, разделяющих общие для этнокультурного сообщества знания. Эвристика невербальных семиотических кодов «обусловливает синтез рационального и образного мышления, что наделяет [невербальные знаки] уникальными свойствами с точки зрения лингвопрагматики и семиотики» [15. С. 1024].

Методы, выбранные для проведения исследования, *общенаучные* — анализа и синтеза, сравнения и аналогии, а также *специальные* — метод социолингвистической и социокультурной интерпретации в сочетании с семантическим, лингвокогнитивным и дискурс-анализом, — обеспечивают системное изучение семиотического и лингвокультурного измерения вербальных и невербальных компонентов политической коммуникации и раскрывают их символьную, концептуальную и прагматическую нагрузку.

Материалом для статьи послужили публикации в мексиканских и латиноамериканских СМИ, как текстовые, так и видеорепортажи, трансляции, отражающие состоявшуюся 1 октября 2024 г. церемонию инаугурации

(рис. 1) — вступления в должность К. Шейнбаум, первой женщины-президента Мексики в истории страны, а также информация, размещенная на сайтах президента и правительства Мексики, включая публикации на других официальных мексиканских интернет-порталах.

Рис. 1. К. Шейнбаум[1] на церемонии инаугурации[2]
Источник: исследование М.В. Ларионовой, С.К. Мосиной.
Figure 1. C. Sheinbaum at the inauguration ceremony
Source: compiled by Marina V. Larionova & Sofia K. Mosina.

Церемония инаугурации, то есть официальная процедура торжественного вступления в должность политика, победившего на выборах, представляет собой важнейший акт политической кампании в любом государстве, так как утверждает преемственность и стабильность власти в стране, ее легитимность и заключает в себе глубинные символьные смыслы. Кульминационным моментом инаугурации является принесение присяги новым главой государства, олицетворяющее собой обещание соблюдать конституцию, самоотверженно и честно служить народу.

Несмотря на то, что концептуальное содержание акта инаугурации, репрезентирующее начало нового политического цикла, остается

[1] Клаудия Шейнбаум одержала убедительную победу на прошедших 2 июня 2024 г. в Мексике президентских выборах как кандидат от альянса *Sigamos Haciendo Historia* — Продолжим творить историю, куда входит политическая партия *Morena* (***Mo**vimiento de **Re**generación **Na**cional* — Движение национального возрождения), основанная самой К. Шейнбаум. Кандидат получила в свою поддержку почти 60% голосов избирателей — за нее проголосовали 36 миллионов мексиканцев — и в соответствии с конституцией страны будет занимать пост президента со дня официального вступления в должность 1 октября 2024 г. по 30 сентября 2030 г.

[2] РСМД. Режим доступа: https://russiancouncil.ru/activity/policybriefs/itogi-vyborov-v-meksike-kakim-budet-prezidentstvo-klaudii-sheynbaum/ (дата обращения: 09.11.2025).

тождественным почти во всех странах мира, сама церемония, включающая в себя знаковые моменты, неповторяющиеся у разных народов, является отражением национально-государственной идентичности. Мексиканская президентская инаугурация через языковые, вербальные и неязыковые культурные коды раскрывает особенности национального менталитета и бытия жителей Мексики.

Результаты исследования

Toma de protesta *как национальный вербальный код инаугурации*

Языковая картина мира как совокупность знаний о мире, воплощенных в языке, и как «специфика отражения бытия через язык» [16. С. 99] воспроизводит национально-культурное своеобразие мировидения народа (термин В. фон Гумбольдта [17. С. 370]) и семиотически организует дискурсивное пространство политической коммуникации, репрезентированное церемонией вступления в должность президента страны.

В мексиканском национальном варианте испанского языка для обозначения процедуры инаугурации президента в качестве языкового кода используется устойчивое сочетание *toma de protesta*, для корректного перевода которого на русский язык требуются дополнительные знания лингвокультурной специфики страны. Ни один из традиционных для испаноязычного языкового пространства терминов — *juramento/juramentación, investidura, declaración, inauguración, asunsión*, которые называют торжественный акт принесения избранным президентом присяги народу при вступлении в должность, не используется в Мексике. Президентская клятва, состоящая из 61 слова, открывается формой 1 лица единственного числа глагола *protestar* и звучит так: «***Protesto*** *guardar y hacer guardar la Constitución Política de los Estados Unidos Mexicanos y las leyes que de ella emanen, y desempeñar leal y patrióticamente el cargo de presidente de la República que el pueblo me ha conferido, mirando en todo por el bien y prosperidad de la Unión; y si así no lo hiciere, que la Nación me lo demande*»[3] ‘*Обещаю* соблюдать и обеспечивать соблюдение политической Конституции Мексиканских Соединенных Штатов и законов, из нее проистекающих, а также преданно и патриотично исполнять обязанности Президента Республики, возложенные на меня народом, неустанно заботясь о благе и процветании Союза. Если я этого не сделаю, то пусть народ строго спросит с меня’.

Соответственно, использование перформативного глагола *protestar* послужило основанием для того, чтобы церемония инаугурации президента Мексики получила название *toma de protesta* ‘принесение клятвы’.

[3] *Melgar Adalid M*. La protesta del presidente de la república. ¿rito o requisito constitucional? // Cuestiones Constitucionales. 2008. № 18. Режим доступа: https://www.scielo.org.mx/scielo.php?script=sci_arttext&pid=S1405-91932008000100004 (дата обращения: 10.09.2025).

Когнитивный диссонанс у тех, кто недостаточно разбирается в особенностях мексиканской языковой картины мира, возникает из-за того, что глагол *protestar* и отглагольное существительное *protesta* реализуют не свое базовое, основное лексическое значение ‘протестовать, выражать несогласие, жаловаться’ и соответственно, ‘протест, несогласие, упрек, жалоба’, хорошо известные всем носителям испанского языка, а национально специфичное: *protestar* ‘обещать, брать обязательство’, *protesta* ‘клятва, обещание, обязательство’.

Согласно академическому словарю испанского языка *Diccionario de la Lengua Española* (DLE), ***protesta*** в своем первом значении интерпретируется как ‘действие и результат глагола протестовать’, поэтому при незнании мексиканского лингвокультурного кода номинация *toma de protesta* может быть ошибочно воспринята как ‘акт протеста’, что совершенно не соответствует действительности. Вторым значением слова указано ‘обещание с подтверждением или засвидетельствованием выполнить что-либо’ и только третьим ‘торжественное обещание официального лица при вступлении в должность’. Четвертое толкование существительного *protesta* терминологическое: юридическое заявление, которое делается с целью не допустить ущемления, а, наоборот, гарантировать обеспечение права какого-либо лица.

protesta

1. f. Acción y efecto de protestar.

Sin.: queja, reproche, crítica, condena, protesto, demanda, petición, abucheo, pitada, pita4, rechifla, desaprobación, disconformidad, descontento, disgusto.

2. f. Promesa con aseveración o atestación de ejecutar algo.

Sin.: promesa, declaración.

3. f. Promesa solemne de un alto dignatario al tomar posesión de su cargo.

4. f. Der. Declaración jurídica que se hace para que no se perjudique, antes bien se asegure, el derecho que alguien tiene[4].

В мексиканском культурно-историческом контексте использование лексемы *protesta*, которая семантически является синонимом существительных *обещание*, *клятва*, обусловлено стремлением нивелировать возможный религиозный подтекст. В современной Мексике в ходе инаугурации особо подчеркивается светский характер политического действия, поэтому президент не приносит клятву (*jurar*), а берет на себя обязательство, обещает (*protestar*). В тексте конституции 1824 г., первой после обретения Мексикой независимости, указывается, что президентская присяга при вступлении в должность включала в себя глагол *jurar* и звучала следующим образом: «*Yo, N., nombrado presidente (o vicepresidente) de los Estados Unidos Mexicanos,* ***juro por Dios y los Santos Evangelios****, que ejerceré fielmente el encargo que los mismos Estados Unidos me han confiado, y que guardaré y haré guardar exactamente la*

[4] Protesta // Diccionario de la lengua española RAE — ASALE Режим доступа: https://dle.rae.es/protesta?m=form (дата обращения: 10.09.2025).

Constitución y leyes generales de la Federación»[5] ‘Я, (имя), назначенный президентом (или вице-президентом) Мексиканских Соединенных Штатов, клянусь перед Богом и на Святом Евангелии, что буду добросовестно исполнять миссию, порученную мне самими Соединенными Штатами, и что буду соблюдать и в точности исполнять Конституцию и генеральные законы Федерации’. В конституции 1857 г. в тексте президентской присяги сохраняется глагол *jurar*, но исчезает упоминание религиозных символов: «***Juro*** *desempeñar leal y patrióticamente el encargo de presidente de los Estados Unidos Mexicanos, conforme a la Constitución y mirando en todo por el bien y prosperidad de la Unión*»[6] ‘Клянусь преданно и патриотично исполнять обязанности Президента Мексиканских Соединенных Штатов в соответствии с Конституцией, неустанно стремясь к благу и процветанию Союза.’

В начале XX в., после целого ряда реформ и принятия в 1917 г. новой, действующей по сей день конституции, церковь была отделена от мексиканского государства, поэтому все вербальные компоненты, которые так или иначе имели отношение к религиозной клятве, были заменены нейтральной терминологией: *jurar* ‘клясться’ превратился в *protestar* ‘обещать’, что по своей сути и стало своеобразным протестом. Изменения в языковом коде отразили тот факт, что Мексика превратилась в светское государство. В связи с этим сегодня на мексиканской церемонии вступления президента в должность употребляется номинация *toma de protesta* и отсутствуют религиозные атрибуты-символы, такие как Библия или крест.

Еще одной прецедентной номинацией в мексиканском варианте испанского языка, активно использующейся в пространстве политического дискурса, является топоним *El Zócalo* — народное название центральной и самой большой площади столицы страны, на которой расположен президентский дворец и где отмечают важнейшие национальные праздники и события, откуда вступивший в должность президент произносит речь перед гражданами страны. Хотя официально это место именуется Площадью Конституции, мексиканцы используют исторически возникший топоним *El Zócalo* ‘Сокало’ (в переводе — цоколь). Это название появилось еще в первой половине XIX в. благодаря плану возвести в центре столицы монумент-колонну в честь независимости страны (соответствующий приказ издал в 1843 г. президент Мексики Антонио Лопес де Санта-Ана). Удалось построить только постамент-основание, цоколь — *zócalo*, а сам памятник так и не был воздвигнут. Тем не менее неофициальное название продолжает жить. Более того, оно стало прецедентным: сложилась традиция и в других мексиканских городах называть главную площадь *El Zócalo* — Сокало.

[5] *Melgar Adalid M*. La protesta del presidente de la república. ¿rito o requisito constitucional? // Cuestiones Constitucionales. 2008. № 18. Режим доступа: https://www.scielo.org.mx/scielo.php?script=sci_arttext&pid=S1405-91932008000100004 (дата обращения: 10.09.2025).
[6] Ibid.

Символьные смыслы, заложенные в номинацию *Zócalo* национально-языковым сознанием, ее метафорическая трактовка оказались востребованными в политическом дискурсе. Победившая на выборах К. Шейнбаум вела свою электоральную кампанию, ее финальную часть, под лозунгом *Del Zócalo a la Victoria* ‘От основания к победе’, что имело дополнительную прагматическую силу *in situ*, так как политик проводила завершающий митинг именно на площади Конституции (рис. 2): *con un Zócalo capitalino completamente lleno, la doctora Claudia Sheinbaum afianzó el respaldo del Pueblo de México para así asegurar su victoria este próximo 2 de junio* ‘на переполненной столичной площади К. Шейнбаум получила внушительную поддержку мексиканского народа, столь необходимую для победы’.

Рис. 2. К. Шейнбаум на центральной площади г. Мехико[7]
Источник: исследование М.В. Ларионовой, С.К. Мосиной.
Figure 2. C. Sheinbaum at the Mexico Central Square
Source: compiled by Marina V. Larionova & Sofia K. Mosina.

Национальные артефакты и символы как невербальные коды инаугурации

В своих работах по семиотике Ю.М. Лотман неоднократно подчеркивает, что знаковые системы являются ключом к пониманию различных культур и организации человеческого мира: знаки «заменяют сущности, явления и вещи и позволяют людям обмениваться информацией» [18. С. 9]. Используя знаки, люди получают возможность формировать свое восприятие реальности, передавать накопленный опыт, сохраняя культурное наследие и формируя свою картину мира: «Знаки с особенной активностью аккумулируют социальный опыт: ордена и звания, гербы и ритуалы, деньги и обряды — все это знаки, отражающие разные принципы организации человеческого коллектива. Но

[7] El Soberano. Режим доступа: https://elsoberano.mx/portadas/del-zocalo-a-la-victoria/ (дата обращения: 09.11.2025).

знаки используются и системами, накапливающими духовный опыт людей» [18. С. 5–10]. О.С. Чеснокова интерпретирует языковую картину мира народа как «один из семиотических кодов культуры, ею же обусловленный» [10. С. 31]. Артефакты как невербальные семиотические знаки являются обязательной частью политического дискурсивного пространства Мексики. В качестве феноменов культуры артефакты обладают символьным знаковым содержанием, передающим информацию о культуре их создателей и пользователей. Для мексиканского культурно-языкового пространства в рамках церемонии вступления в должность президента страны обращение к культурным кодам автохтонной цивилизационной парадигмы имеет особый смысл.

Ритуальным артефактом президентской власти в Мексике является жезл — *bastón de mando*, передача которого президенту имеет особый символьный смысл. Жезл создают лучшие ремесленники из индихенистских коммун и вручают представители 70 индихенистских и афро-мексиканских сообществ как знак власти, которую национальный лидер обязан осуществлять для народа, подчиняясь воле народа, творя справедливость в интересах коренного населения страны — *el mandar obedeciendo al pueblo y hacer justicia en las comunidades originarias*[8] (рис. 3–4). Глубинный символьный смысл и мудрость предков, которые заключают в себе жезл-артефакт, транслируют философию управления со смирением и в соответствии с потребностями народа — *mandar obedeciendo* ‘править через подчинение’. Этот древний артефакт-символ могущества, используемый различными коренными народами, служит напоминанием о том, что власть должна осуществляться на благо тех, кто исторически был дискриминирован.

Жезл олицетворяет духовную силу народов Мексики и общинную власть. По традиции, передавая жезл, коренные народы признают президента Мексики своим законным правителем. Форма жезла — вертикальная и прямая — символизирует справедливость, равновесие и гармонию (*la justicia, el balance y la armonía*) как имманентные свойства, присущие времени и жизни (*el tiempo y la vida*)[9]. В верхней части жезла изображен внеземной мир (*el supramundo*), где обитают боги, затем — традиционный мир, с четырьмя сторонами света и центром вселенной (*con los cuatro rumbos y el centro del universo*), где живут все люди, а внизу — подземный мир (*el inframundo*), где обитают души предков[10]. Таким образом, жезл-артефакт отражает мировидение коренных народов, где соединяется воедино мир богов, мир людей и мир ушедших и где правители должны действовать справедливо и сбалансированно, руководствуясь ценностями предков, определяющими благополучие общества.

[8] Claudia Sheinbaum: Detalles sobre el bastón de mando que recibirá la Presidenta de México // El Universal Режим доступа: https://www.eluniversal.com.mx/nacion/claudia-sheinbaum-detalles-sobre-el-baston-de-mando-que-recibira-la-presidenta-de-mexico/ (дата обращения: 09.11.2025).
[9] Ibid.
[10] Ibid.

Рис. 3. Артефакты на церемонии инаугурации[11]
Источник: исследование М.В. Ларионовой, С.К. Мосиной.
Figure 3. The artefacts at the ceremony of inauguration
Source: compiled by Marina V. Larionova & Sofia K. Mosina.

Рис. 4. Обряд очищения перед инаугурацией[12]
Источник: исследование М.В. Ларионовой, С.К. Мосиной.
Figure 4. The purification ritual before the inauguration
Source: compiled by Marina V. Larionova & Sofia K. Mosina.

Семиотика мексиканских инаугурационных мероприятий раскрывает необходимость понимания дискурсивных кодов, языковых и невербальных, для адекватной интерпретации ее политических и культурных смыслов.

Непосредственно перед актом передачи жезла К. Шейнбаум прошла еще один ритуал предков — особый обряд очищения, благословения и наделения силой, традиционный для автохтонных культур страны, который призван обеспечить успешное правление президента многонациональной Мексики: *Tú eres la voz de las que no tuvimos voz por mucho tiempo, eres la voz de nuestros pueblos, eres la esperanza que nosotros teníamos. (…) Claudia, que los elementos sagrados te acompañen, que el agua bendita purifique siempre tu alma, que el aire siempre esté contigo, que nuestra madrecita tierra te bendiga siempre. (…) Invocamos a los espíritus divinos que habitan este lugar, pedimos vida, iluminación y sabiduría para la presidenta constitucional, doctora Claudia Sheinbaum Pardo, pedimos para ella tu luz divina, que la acompañes, que la ilumines a cada paso que dé, en cada decisión que tome. (…) Todo el poder para que ella se llene de fortaleza y en estos momentos bajamos la energía para que ella esté llena de virtudes y poderes, y camine por el bien de los 70 pueblos originarios y de todo el pueblo de México*[13] — Ты — голос тех из нас, кто так долго был безмолвен. Ты — голос

[11] Официальный сайт Правительства Мексики Gobierno de México. Режим доступа: https://www.gob.mx/inpi/articulos/mujeres-de-70-pueblos-indigenas-y-afromexicano-entregan-baston-de-mando-y-servicio-ala-presidenta-claudia-sheinbaum-pardo (дата обращения: 09.11.2025).

[12] El Heraldo de Saltillo. Режим доступа: https://elheraldodesaltillo.mx/2024/10/01/entregan-baston-de-mando-a-claudia-sheinbaum-y-le-hacen-una-limpia/ (дата обращения: 09.11.2025).

[13] Entregan Bastón de Mando a Claudia Sheinbaum, y le hacen una “limpia” Режим доступа: https://elheraldodesaltillo.mx/2024/10/01/entregan-baston-de-mando-a-claudia-sheinbaum-y-le-hacen-una-limpia/ (дата обращения: 09.11.2025).

наших народов. Ты — наша надежда. (…) Клаудия, да пребудут с тобой священные стихии, пусть святая вода всегда очищает твою душу, да пребудет с тобой воздух, да благословит тебя наша земля-матушка. (…) Мы призываем божественных духов, обитающих в этом месте, и молим о жизни, просветлении и мудрости для конституционного президента, доктора Клаудии Шейнбаум Пардо. Мы молим о божественном свете, чтобы он сопровождал ее, чтобы озарял каждый ее шаг, каждое ее решение. (…) Призываем всю силу, чтобы она могла зарядиться ею, и призываем снизойти энергию, чтобы Клаудия наполнилась добродетелями и обрела уверенность на своем пути во благо 70 коренных народов и всего народа Мексики.

Таким образом, важное место, которое отводится на церемонии инаугурации президента автохтонным традициям, ритуалам и артефактам, символам как неотъемлемой части национального менталитета, свидетельствует о том, что индихенизм остается значимой культурной доминантой цивилизационной парадигмы Мексики, источником, откуда ее народ черпает энергию, силу и коллективную мудрость — *energía, fuerza y sabiduría colectiva*[14].

Необходимо обратить внимание еще на один важнейший артефакт-символ, ставший частью инаугурационной церемонии: изображение орла, сидящего на кактусе и держащего в лапе и клюве змею (рис. 5). Образ орла представляет собой центральную часть национального герба Мексики. Изображение разместили на трибуне, откуда произносила речь К. Шейнбаум, на креслах в президиуме, а также на президентской ленте как части инаугурационного наряда президента страны. Также орел в виде серебряной статуэтки располагался в центре стола заседаний.

Рис. 5. Орел в государственной символике Мексики[15]
Источник: исследование М.В. Ларионовой, С.К. Мосиной.
Figure 5. An eagle in Mexican national symbols
Source: compiled by Marina V. Larionova & Sofia K. Mosina.

[14] Entregan Bastón de Mando a Claudia Sheinbaum, y le hacen una “limpia” Режим доступа: https://elheraldodesaltillo.mx/2024/10/01/entregan-baston-de-mando-a-claudia-sheinbaum-y-le-hacen-una-limpia/ (дата обращения: 09.11.2025).
[15] Campaigns & Elections: México. Режим доступа: https://ceonline.com.mx/post/inicia-era-de-sheinbaum-rinde-protesta-como-presidenta-de-mexico/ (дата обращения: 09.11.2025).

С доиспанских времен орел является частью мексиканской мифологии, обладает и транслирует глубокие культурные смыслы. Священный орел — посланник Уитцилопочтли (Huitzilopochtli), бога войны в мифологии ацтеков, а также покровителя города Теночтитлан, основанного на месте современного Мехико. Бога Уитцилопочтли отождествляют с солнцем, а орла считают птицей, летающей на самой большой высоте, выше облаков, как и звезда-солнце: *el águila sagrada representaba a Huitzilopochtli, dios de la guerra, a quien también se identifica con el sol, al ser el ave que vuela a mayor altura, tal como lo hace el astro*[16]. В геральдике Мексики орел символизирует стойкий характер мексиканцев, силу, свет и добро — *simboliza el carácter aguerrido del mexicano, la fuerza, la luz y el bien*[17], мощь, превосходство над другими, воинственный характер, агрессивность и храбрость: *la connotación del águila real es de fuerza, poderío, dominio sobre los otros, el carácter guerrero, la agresividad y la valentía*[18]. Изображение орла, пожирающего змею, в официальной символике Мексики, уходящей корнями в ацтекскую мифологию, олицетворяет победу солнца над его врагами, сил добра и созидания над силами разрушения и зла. Во время войны за независимость Мексики орел изображался на знаменах как символ стремления к свободе и победы. В политическом дискурсивном контексте коннотативные смыслы, которые визуально репрезентирует изображение орла как национального символа, опираются на стереотипы национально-культурного мировидения народа, обладают особым прагматическим потенциалом и легко «считываются», транслируя идею победы, независимости и божественного покровительства, необходимых успешному политику.

Национальная одежда как вестиментарный код инаугурации

Продуманный имидж является непременным условием достижения успеха политическим деятелем. Важную роль играет внешность и одежда, создающие публичный образ политика, заинтересованного в получении и удержании поддержки электората. Вестиментарный код, который продемонстрировала президент Мексики на церемонии инаугурации, опирается на имплицитное использование пресуппозитивных знаний национальной аудитории, которые хранятся в коллективном бессознательном и являются частью культурно-исторической памяти народа. Элементы наряда, выбранного К. Шейнбаум для церемонии вступления в должность, несут дополнительную смысловую нагрузку, добавляя новые смыслы, которые формируют и поддерживают определенную картину мира в сознании национальной аудитории.

[16] Официальный сайт Правительства Мексики Gobierno de México. Режим доступа: https://www.gob.mx/semarnat/articulos/el-aguila-real-en-nuestro-escudo-nacional (дата обращения: 09.11.2025).

[17] Официальный сайт Правительства Мексики Gobierno de México. Режим доступа: https://www.gob.mx/conanp/articulos/el-aguila-real-un-simbolo-de-identidadnacional?idiom=es (дата обращения: 09.11.2025).

[18] Ibid.

Для инаугурации К. Шейнбаум выбирает платье цвета слоновой кости, украшенное традиционной мексиканской вышивкой — цветочным орнаментом в этническом стиле, которую вручную выполнила лучшая мастерица из штата Оахака (рис. 6). Подобным поступком К. Шейнбаум выражает уважение к культурным традициям, демонстрирует поддержку народных промыслов, а также подчеркивает единство с народом Мексики[19].

Рис. 6. Платье К. Шейнбаум на церемонии инаугурации[20]
Источник: исследование М.В. Ларионовой, С.К. Мосиной.
Figure 6. C. Sheinbaum's dress at the inauguration ceremony
Source: compiled by Marina V. Larionova & Sofia K. Mosina.

Национальные элементы в виде вышивки цветочных узоров на инаугурационном наряде призваны актуализировать значение, которое К. Шейнбаум придает в рамках своей политической программы правам коренных жителей и национальному вопросу, а также привлечь на свою сторону электорат, считывающий близкий к своей национальной идентичности культурный код. Белое (оттенка слоновой кости) платье и украшающая его цветочная гирлянда воспринимаются как холст, на котором визуально, не вербально, отражена многовековая история мексиканского сообщества и мировоззрение народа — *se convierte en un lienzo donde se entrelazan historias y simbolismos, se comunica la conexión de los pueblos indígenas con su entorno y su cosmovisión*[21]. Кроме того, выбор одежды призван отвлечь внимание прессы от этнической

[19] Animal MX. Режим доступа: https://animalpolitico.com/tendencias/actualidad/vestido-de-claudia-shein baum-toma-de-protesta (дата обращения: 09.11.2025).
[20] México News Daily. Режим доступа: https://mexiconewsdaily.com/lifestyle/the-story-behind-sheinbaum-dress/ (дата обращения: 09.11.2025).
[21] РежиMexico Historico. Режим доступа: https://www.mexicohistorico.com/paginas/El-significado-del-color-blanco-en la-cultura-ind--gena-mexicana.html (дата обращения: 09.11.2025).м доступа: https://www.mexicohistorico.com/paginas/El-significado-del-color-blanco-en-la-cultura-ind--gena-mexicana.html (дата обращения: 09.11.2025).

принадлежности К. Шейнбаум, которую критиковали как кандидата, не обладающего коренной мексиканской национальностью[22].

Выбор цвета инаугурационного платья не был случайным. Следует помнить о том, что цвет является культурным знаком, имплицирующим символическое значение [19; 20]. При этом, по мнению А.Н. Леонтьева, цвет обладает сигнальной функцией и сам по себе выступает лишь как означающее, а означаемым становится некое эмоциональное состояние, и только в сочетании с каким-либо состоянием или ситуацией цвет образует знак [21]. Иначе говоря, для того чтобы цвета стали знаками, должна сложиться традиция их связанности с определенными ментальными состояниями или событиями.

В мексиканской культуре белый цвет пронизывает ритуалы и церемонии, придавая им сакральный и значимый характер, является символом жизни, гармонии и возрождения, воплощает сущность бытия с точки зрения коренных народов. У ацтеков и майя белый цвет отождествляется с сотворением и плодородием, отражает глубокую связь с циклом жизни, символизирует чистоту и обновление. Белый цвет играет важную роль в ритуалах и церемониях коренных народов Мексики, где он используется для призыва защиты, чистоты и благословения духовных сил — *El color blanco tiene un papel destacado en los rituales y ceremonias indígenas, donde se utiliza para invocar la protección, la pureza y la bendición de las fuerzas espirituales*[23]. Во многих случаях участники надевают белую одежду во время важных церемоний, таких как свадьбы, фестивали и ритуалы перехода в мир иной, что символизирует у мексиканцев новое начало и надежду.

Обращение К. Шейнбаум к белому цвету в рамках важнейшей государственной церемонии не только подчеркивает автохтонную историческую традицию его использования, но реализует дополнительный коннотативный и метафорический смысл: символизирует переход к новому политическому курсу, не запятнанному коррупцией.

После вручения символов президентской власти на К. Шейнбаум торжественно возложили ленту в цветах мексиканского флага, которую может носить только президент. Символьные цвета — белый, зеленый и красный, соответственно, представляют религию, независимость и единство мексиканцев — [*con*] *sus colores, blanco, verde y rojo, respectivamente, fueron representadas la religión, la independencia y la unión de los mexicanos*[24]. Согласно закону «О национальных гербе, флаге и гимне»[25] лента должна

[22] DAILY BULLETIN. Режим доступа: https://www.dailybulletin.com/2024/04/19/claudia-sheinbaum-at-the-inter section-of-sexism-and-antisemitism/amp/ (дата обращения: 09.11.2025).

[23] Mexico Historico. Режим доступа: https://www.mexicohistorico.com/paginas/El-significado-del-color-blanco-en-la-cultura-ind--gena-mexicana.html (дата обращения: 09.11.2025).

[24] Официальный сайт Правительства Мексики Gobierno de México. Режим доступа: https://www.gob.mx/conanp/articulos/el-aguila-real-un-simbolo-de-identidad-nacional?idiom=es (дата обращения: 09.11.2025).

[25] Режим доступа: https://www.diputados.gob.mx/LeyesBiblio/pdf/LEBHN.pdf (дата обращения: 09.11.2025).

надеваться строго под пиджак, однако К. Шейнбаум подчеркнула тот факт, что она является первой женщиной-президентом Мексики, надев на церемонию *toma de protesta* платье.

Семиотика политического дискурса, объединяя вербальные и невербальные знаковые коды, отражает социальную действительность. Реальность того, что К. Шейнбаум стала первой в истории Мексики женщиной-президентом, объективируется через невербальную коммуникацию. Это историческое для Мексики событие было подчеркнуто на церемонии вступления в должность официальным сопровождением К. Шейнбаум: в ходе церемонии ее охраняли кадеты-женщины.

Таким образом, семиотический механизм невербального воздействия одежды политика на восприятие его образа в политическом дискурсивном пространстве основывается на пресуппозитивных знаниях аудитории-адресата, разделяющей совокупную когнитивную картину мира. Дискурсивная риторика вестиментарного кода интерпретируется как отличительный социально-культурный знак, актуализирующий дихотомию «*свой — чужой*», обозначающий границы между этническими группами, политическими сообществами, идеологическими объединениями, репрезентирующий и разъясняющий ценностные приоритеты и убеждения облаченного в определенный костюм политика: одежда «всегда обладает дополнительным значением, не связанным с ее прямым назначением. Это коннотативное значение может иметь характер идеологический, этический, религиозный, культурный или политический, но всегда направлено на формирование определенного восприятия человека» [22. С. 11], который в нее одет.

Заключение

Понимание мексиканского мира как лингвокультурного семиотического пространства включает в себя знания кодов национальной культуры, ее вербальных и невербальных знаков, что имеет особое прагматическое значение в политической дискурсивной среде. К. Шейнбаум во время церемонии инаугурации, именуемой в Мексике национально маркированным сочетанием *toma de protesta*, которое отражает мексиканское языковое членение мира, заложенное в системных значениях слов, активно использовала в качестве невербальных дискурсивных компонентов традиционные мексиканские артефакты, символы и ритуалы, послужившие актуализации ее политической позиции и убеждений, продемонстрировав тем самым глубокое уважение к традициям мексиканского народа.

Цифровые возможности новой дискурсивной среды, существенным образом изменившие и расширившие границы политической коммуникации, трансформировали способы кодирования — подачи и декодирования — восприятия, информации, значительно увеличив удельный вес невербальных дискурсивных компонентов: видео и аудио контента, цветовых, графических, иконических и пр. семиотических кодов. Невербальная коммуникация,

дополняющая, а иногда и замещающая вербальную, представляется особенно информативной: она позволяет передать дополнительные смыслы, легко декодируемые реципиентами, которые недоступны вербальной знаковой системе или требуют подробных и длинных языковых описаний.

Семиотическими доминантами невербальной политической коммуникации являются материальные артефакты, символы власти, одежда, цвет, визуальные изображения, фото и политические плакаты, визуальные метафоры и логотипы, которые обладают символьным знаковым содержанием, создают имидж и позволяют аудитории сформировать мнение о политике, не только основываясь на его вербальных заявлениях, но и руководствуясь собственными ощущениями, которые суггестивно транслируют экстралингвистические компоненты дискурса. К. Шейнбаум, первая женщина-президент Мексики, политически эффективно использует культурно-значимые национальные символы, артефакты и традиции, формируя свой публичный имидж и поведение с учетом цивилизационных импликатур, разделяемых мексиканским народом. Это позволяет ей выстраивать сбалансированный политический курс, который ориентирован на поддержание исторической, духовной и культурной самобытности коренных народов, на укрепление роли женщин в развитии страны, включая более заметное участие представительниц автохтонных культур, долгие годы остававшихся «невидимыми» — *invisibles*, а также на защиту национальных интересов Мексики в ее противостоянии США и на международной арене. Таким образом, политический дискурс, имплицирующий семиотически поликодовые и полимодальные знаковые системы, функционирует не только как способ обмена информацией и достижения взаимопонимания между политиками и их электоратом, не просто репрезентирует политику, а сам становится ее частью.

Список литературы

1. *Красных В.В.* «Свой» среди «чужих»: миф или реальность? М. : Гнозис, 2003.
2. *Алефиренко Н.Ф.* Лингвокультурология. Ценностно-смысловое пространство языка. М. : Флинта: Наука, 2010.
3. *Воробьев В.В., Виноградов В.В., Саяхова Л.Г., Хайруллина Р.Х.* Сопоставительная лингвокультурология: теория и принципы анализа языковых. Уфа: РИЦ БашГУ, 2015.
4. *Маслова В.А.* Лингвокультурология. М. : Академия, 2004. EDN: QRFEZH
5. *Оболенская Ю.Л.* Иберийский мир: пограничные цивилизации в эпоху глобализации // Russian Journal of Linguistics. 2013. № 4. С. 39–45. EDN: RKUXIB
6. *Ольшанский И.Г.* Лингвокультурология в конце XX века: итоги, тенденции, перспективы // Лингвистические исследования в конце XX века. М. : ИНИОН, 2000. С. 26–55.
7. *Хроленко А.Т.* Основы лингвокультурологии. М. : Флинта-Наука, 2005. EDN: QRWUVD
8. *Ларионова М.В.* Латиноамериканский цивилизационный дискурс: топонимы как паттерны культурной идентичности // Язык и культура. 2024. № 67. С. 32–57. https://doi.org/10.17223/19996195/67/2 EDN: FCNSRM
9. *Ларионова М.В.* Свой среди чужих, чужой среди своих: о латиноамериканской идентичности сквозь призму этнонимов latinoamericano/latin, hispanic // Иберoамериканские тетради. 2025. Т. 13. № 3. С. 39–63. https://doi.org/10.46272/2409-3416-2025-13-3-39-63 EDN: BFWUFZ

10. *Чеснокова О.С.* Испанский язык Мексики: языковая картина мира. М. : Ленанд, 2021. EDN: DOCBPD
11. *Кофман А.Ф.* «Что гринго здорово, то латиноамериканцу смерть». Особенности выстраивания латиноамериканской картины мира // Иберoамериканские тетради. 2023. № 1. С. 16–40. https://doi.org/10.46272/2409-3416-2023-11-1-16-40 EDN: JBTLVO
12. *Шемякин Я.Г.* «Великие географические открытия»: смысл и историческое содержание термина в свете опыта взаимодействия цивилизаций в Новом Свете // Иберoамериканские тетради. 2023. № 1. С. 41–67. https://doi.org/10.46272/2409-3416-2023-11-1-41-67 EDN: FWXAKH
13. *Correa Henao J. D.* Panamericanismo versus latinoamericanismo: tensión geopolítica y civilizacional // Analecta Política. 2020. Vol. 10. № 19. P. 56–76. https://doi.org/10.18566/apolit.v10n19.a03 EDN: NYRFLR
14. *Ramos V.H.* La identidad latinoamericana: proceso contradictorio de su construcción- deconstrucción-reconfiguración dentro de contextos globales // Universitas Humanística. 2012. № 73. P. 15–58.
15. *Ларионова М.В.*, *Мосина С.К.* Семиотика визуальной метафоры в мексиканском электоральном дискурсе: на примере президентской избирательной кампании 2024 г. // Вестник Российского университета дружбы народов. Серия: Теория языка. Семиотика. Семантика. 2024. Т. 15. № 3. С. 1021–1039. https://doi.org/10.22363/2313-2299-2024-15-3-1021-1039 EDN: KFYAXU
16. *Карасик В.И.* Языковой круг: личность, концепты, дискурс. М. : Гнозис, 2004.
17. *Гумбольдт В.* Язык и философия культуры. М. : Прогресс, 1985.
18. *Лотман Ю.М.* Семиосфера. С.-Петербург : Искусство–СПБ, 2000.
19. *Тонквист Г.* Аспекты цвета. Что они значат и как могут быть использованы. Проблема цвета в психологии. М. : Академкнига, 1993.
20. *Липская В.М.* Символика цвета в культуре и ее роль в костюме // Теория и практика сервиса: экономика, социальная сфера, технологии. 2012. № 2 (12). С. 67–73.
21. *Леонтьев А.Н.* Проблемы развития психики: монография. М. : Издательство Московского университета, 1981. EDN: QBKURD
22. *Сливчикова Ю.В.* Прагмалингвистические характеристики вестиментарного описания в газетно-публицистическом дискурсе (на материале испанского языка) : дис. … канд. филол. наук. М., 2016. EDN: ZQIEQF

References

1. Krasnykh, V.V. (2003). *“Ours” among “Strangers”: Myth or Reality?* Moscow: Gnosis. (In Russ.).
2. Alefirenko, N.F. (2010). *Linguoculturology. The Value-Semantic Space of Language*. Moscow: Flinta: Nauka. (In Russ.).
3. Vorobiev, V.V., Vinogradov, V.V., Sayakhova, L.G., & Khairullina, R.Kh. (2015). *Comparative Linguoculturology: Theory and Principles of Analysis of Linguistic Units*. Ufa: Bashkir State University publ. (In Russ.).
4. Maslova, V.A. (2004). *Linguocultural Studies*. Moscow: Academy. (In Russ.). EDN: QRFEZH
5. Obolenskaya, Yu.L. (2013). Iberian World: Limitrophes Civilizations on the Epoch of Globalization. *Russian Journal of Linguistics*, 4, 39–45. (In Russ.). EDN: RKUXIB
6. Olshansky, I.G. (2000). Linguistic Studies at the End of the Twentieth Century: Results, Trends, Prospects. In: *Linguistic Studies at the End of the Twentieth Century* (pp. 26–55). Moscow: INION. (In Russ.).
7. Khrolenko, A.T. (2005). *Fundamentals of Linguocultural Studies*. Moscow: Flinta-Nauka. (In Russ.). EDN: QRWUVD

8. Larionova, M.V. (2024). Latin American Civilizational Discourse: Toponyms as Patterns of Cultural Identity. *Language and Culture*, 67, 32–57. (In Russ.) https://doi.org/10.17223/19996195/67/2 EDN: FCNSRM
9. Larionova, M.V. (2025). One of Our Own Among Strangers, a Stranger Among Our Own: On Latin American Identity Through the Prism of the Ethnonyms Latinoamericano/Latin, Hispanic. *Cuadernos Iberoamericanos*, *13*(3), 39–63. (In Russ.). https://doi.org/10.46272/2409-3416-2025-13-3-39-63 EDN: BFWUFZ
10. Chesnokova, O.S. (2021). *Spanish Language of Mexico: a Linguistic Picture of the World*. Moscow: Lenand. (In Russ.). EDN: DOCBPD
11. Kofman, A.F. (2023). What's Great for a Gringo Is Death for a Latino. Features of Building a Latin American Picture of the World. *Cuadernos Iberoamericanos*, 1, 16–40. (In Russ.). https://doi.org/10.46272/2409-3416-2023-11-1-16-40 EDN: JBTLVO
12. Shemyakin, Ya.G. (2023). "Great Geographical Discoveries": the Meaning and Historical Content of the Term in the Light f the Experience of Interaction of Civilizations in the New World. *Cuadernos Iberoamericanos*, 1, 41–67. (In Russ.). https://doi.org/10.46272/2409-3416-2023-11-1-41-67 EDN: FWXAKH
13. Correa Henao, J.D. (2020). Panamericanismo Versus Latinoamericanismo: Tensión Geopolítica y Civilizacional. *Analecta Política*, *10*(19), 56–76. https://doi.org/10.18566/apolit.v10n19.a03 EDN: NYRFLR
14. Ramos, V.H. (2012). La Identidad Latinoamericana: Proceso Contradictorio de su Construcción- Deconstrucción-Reconfiguración Dentro de Contextos Globales. *Universitas Humanística*, 73, 15–58.
15. Larionova, M.V., & Mosina, S.K. (2024). Semiotics of Visual Metaphor in Mexican Electoral Discourse: Case-Study of the 2024 Presidential Election Campaign. *RUDN Journal of Language Studies, Semiotics and Semantics*, *15*(3), 1021–1039. (In Russ.). https://doi.org/10.22363/2313-2299-2024-15-3-1021-1039 EDN: KFYAXU
16. Karasik, V.I. (2004). *Language Circle: Personality, Concepts, Discourse*. Moscow: Gnosis. (In Russ.).
17. Humboldt, W. (1985). *Language and Philosophy of Culture*. Moscow: Progress. (In Russ.).
18. Lotman, Yu.M. (2000). *Semiosphere*. Saint Petersburg: Art–SPB. (In Russ.).
19. Tonqvist, G. (1993). *Aspects of Color. What They Mean and How they Can be Used. The Problem of Color in Psychology*. Moscow: Akademkniga. (In Russ.)
20. Lipskaya, V.M. (2012). Symbolism of Color in Culture and Its Role in Costume. *Theory and Practice of Service: Economics, Social Sphere, Technology*, 2(12), 67–73. (In Russ.).
21. Leontiev, A.N. (1981). *Problems of Psyche Development: monograph*. Moscow: Moscow University publ. (In Russ.). EDN: QBKURD
22. Slivchikova, Yu.V. (2016). *Pragmalinguistic Characteristics of Vestimentary Description in Newspaper and Journalistic Discourse (Based on the Spanish Language)* [PhD Thesis]. Moscow. (In Russ.). EDN: ZQIEQF

Сведения об авторах:

Ларионова Марина Владимировна, доктор филологических наук, профессор, профессор кафедры испанского языка, Московский государственный институт международных отношений (университет) Министерства иностранных дел Российской Федерации (119454, Российская Федерация, г. Москва, пр-т Вернадского, д. 76); *сфера научных интересов*: семиотические, когнитивные и дискурсивные исследования; *e-mail*: larionova.m@list.ru ORCID: 0000-0001-6466-7363, Scopus ID: 57207834882, SPIN-код: 365254, Author ID: 5303-1160, Researcher ID: E-3214-2016.

Мосина София Константиновна, кандидат филологических наук, преподаватель кафедры испанского языка, Московский государственный институт международных отношений (университет) Министерства иностранных дел Российской Федерации (119454, Российская Федерация, г. Москва, пр-т Вернадского, д. 76); *сфера научных интересов*: испанский политический дискурс; *e-mail*: sophiemosina@yandex.ru
ORCID: 0000-0002-0569-9717, SPIN-код: 4201-8787, Scopus ID: 58886125600, Author ID: 927655.

Information about the authors:

Marina V. Larionova, Dr. Sc. (Philology), Professor, Department of Spanish Language, Moscow State Institute of International Relations (University) of the Ministry of Foreign Affairs of the Russian Federation (76 Vernadsky ave., Moscow, Russian Federation, 119454); *Research interests*: semiotics, cognitive and discourse studies; *e-mail*: larionova.m@list.ru
ORCID: 0000-0001-6466-7363, Scopus ID: 57207834882, SPIN-code: 365254, RSCI Author ID: 5303-1160, Researcher ID: E-3214-2016.
Sofia K. Mosina, PhD in Philology, lecturer at the Department of Spanish Language, Moscow State Institute of International Relations (University) of the Ministry of Foreign Affairs of the Russian Federation (76 Vernadsky ave., Moscow, Russian Federation, 119454); *Research interests*: Spanish political discourse; *e-mail*: sophiemosina@yandex.ru
ORCID: 0000-0002-0569-9717, Scopus ID: 58886125600, SPIN-code: 4201-8787, Author ID: 927655.

RUDN Journal of Language Studies, Semiotics and Semantics
2313-2299 (Print), ISSN 2411-1236 (Online)
Вестник Российского университета дружбы народов.
Серия: ТЕОРИЯ ЯЗЫКА. СЕМИОТИКА. СЕМАНТИКА

2026 Vol. 17 No. 2 668–689
http://journals.rudn.ru/semiotics-semantics

DOI: 10.22363/2313-2299-2026-17-2-668-689
EDN: MAVBOP

Научная статья / Research Article

Эволюция поликодового текста в российском медиадискурсе: семантика, семиотика, институциональность

С.С. Хромов ✉, **А.А. Распопов**

Московский политехнический университет, *Москва, Российская Федерация*
✉chelovek653@mail.ru

Аннотация. Цель исследования — уточнение понятийных границ медиатекста как разновидности поликодового текста и медиадискурса в условиях цифровой конвергенции с учётом трансформации современной медиасреды, медиапотребления и медиапроизводства, а также обоснование вертикального видеотекста как новой формы функционирования медиатекста в платформенной среде. В статье рассматриваются понятийные границы медиатекста как вида поликодового текста и медиадискурса в условиях цифровой конвергенции. Методы — функционально-стилистический, дискурс- и контент-анализ, а также семиотический анализ взаимодействия вербального, иконического и мелодического кодов; эмпирическую базу составляют 200 медиатекстов российских СМИ, бренд-медиа и гражданской журналистики, представленных в форматах вертикальных видео, карточек и видеосообщений («кружков») в мессенджере Telegram. Результаты — уточнено содержание категории «медиатекст» как поликодового образования, формирующегося в условиях экранной культуры, клипового мышления, поточного сознания, децентрализации СМИ и «взрыва авторства»; показано, что медиатекст утрачивает институциональность и функционирует в едином потоке журналистских, пользовательских, рекламных и корпоративных высказываний; формат обоснован как наджанровая единица медиапроизводства, задающая визуально-коммуникативную структуру медиатекста и режим его восприятия. Выводы — современный медиатекст целесообразно рассматривать как разновидность поликодового текста, функционирующего в цифровом медиадискурсе; вертикальный видеотекст определяется как одна из наиболее показательных форм его семиотической и прагматической трансформации в условиях платформенной логики распространения контента.

Ключевые слова: медиатекст, новые медиа, вертикальное видео

Вклад авторов: вклад авторов равнозначен на всех этапах исследования и подготовки текста статьи.

Конфликт интересов: авторы заявили об отсутствии конфликтов интересов.

История статьи: дата поступления: 21.11.2025; дата приема в печать: 5.12.2025.

© Хромов С.С., Распопов А.А., 2026
This work is licensed under a Creative Commons Attribution-NonCommercial 4.0 International License
https://creativecommons.org/licenses/by-nc/4.0/legalcode

Для цитирования: *Хромов С.С., Распопов А.А.* Эволюция поликодового текста в российском медиадискурсе: семантика, семиотика, институциональность // Вестник Российского университета дружбы народов. Серия: Теория языка. Семиотика. Семантика. 2026. Т. 17. № 2. С. 668–689. https://doi.org/10.22363/2313-2299-2026-17-2-668-689 EDN MAVBOP

Evolution of the Polycode Text in Russian Media Discourse: Semantics, Semiotics, and Institutional Dimensions

Sergey S. Khromov ✉, Artyom A. Raspopov

Moscow Polytechnic University, *Moscow, Russian Federation*
✉chelovek653@mail.ru

Abstract. The aim of the study is to clarify the conceptual boundaries of the media text as a variety of polycode text and media discourse under conditions of digital convergence, taking into account the transformation of the contemporary media environment, media consumption, and media production, as well as to substantiate the vertical video text as a new form of media text functioning within a platform-based environment. The article examines the conceptual boundaries of the media text as a type of polycode text and media discourse in the context of digital convergence. The methods employed include functional-stylistic analysis, discourse analysis, and content analysis, as well as a semiotic analysis of the interaction between verbal, iconic, and melodic codes; the empirical corpus consists of 200 media texts from Russian mass media, brand media, and civic journalism, presented in the formats of vertical videos, cards, and video messages (“circles”) in the Telegram messenger. The results demonstrate that the content of the category “media text” is specified as a polycode formation shaped by screen culture, clip thinking, stream consciousness, the decentralization of mass media, and the “explosion of authorship”; it is shown that the media text loses its institutional anchoring and functions within a unified flow of journalistic, user-generated, advertising, and corporate utterances; the format is substantiated as a supragenre unit of media production that determines the visual–communicative structure of the media text and the mode of its perception. The conclusions suggest that the contemporary media text should be regarded as a type of polycode text functioning within digital media discourse; the vertical video text is defined as one of the most indicative forms of its semiotic and pragmatic transformation under the platform logic of content distribution.

Keywords: media text, new media, vertical video

Authors’ contribution: The authors contributed equally to this article and preparation of the article text**.**

Conflict of interests: The authors stated that there were no conflicts of interest.

Article history: Received: 22.11.2025; accepted: 5.12.2025.

For citation: Khromov, S.S., & Raspopov, A.A. (2026). Evolution of the Polycode Text in Russian Media Discourse: Semantics, Semiotics, and Institutional Dimensions. *RUDN Journal of Language Studies, Semiotics and Semantics, 17*(2), 668–689. (In Russ.). https://doi.org/10.22363/2313-2299-2026-17-2-668-689 EDN MAVBOP

Введение

Современная медиасреда характеризуется радикальной трансформацией форм и способов распространения информации, что обусловлено переходом

в цифровую, алгоритмизированную и мультимодальную коммуникационную среду. В этих условиях понятие медиатекста требует уточнения и расширения, поскольку традиционные подходы к пониманию медиатекста уже не охватывают всё разнообразие современных практик. Новые форматы новых медиа, становясь ключевой средой функционирования медиатекстов, определяют их семиотическую организацию и способы интерпретации.

Фундаментальным свойством медиатекста в новых медиа в условиях доминирования аудиовизуального контента становится поликодовость. Кроме того, технологии Интернета радикально изменили парадигму медиакоммуникации: из вертикальной модели она превратилась в горизонтальную [1. С. 261]. Новые форматы новых медиа — короткие вертикальные видео, сторис, серии фотографий, видеосообщения в мессенджерах и др. — создаются, распространяются и потребляются в цифровых экосистемах, где журналистские, рекламные и личные высказывания сосуществуют в едином потоке. Это приводит к размыванию границ между институциональной и повседневной коммуникацией, к возникновению гибридных поликодовых медиатекстов, в которых смешиваются черты новостного, публицистического и рекламного дискурсов, трансформируются механизмы когезии, семиотической и лингвистической организации.

Изучение языковых процессов, которые изменяются под влиянием новейших медиатехнологий, стирающих границы между теорией и практикой взаимодействия языка и культуры, как подчеркивает В.В. Воробьев, требует «междисциплинарного мышления», необходимого «для понимания усложненной современной культуры с ее системой аксиологических и смысловых связей», и «постоянного поиска новых методов и методик, заимствуемых из различных дисциплин» [2. С. 12].

Теоретическое осмысление медиатекста опирается, с одной стороны, на отечественные научные традиции, связанные с изучением публицистического стиля, лингвистики текста и языка СМИ. Медиатекст закрепился как одна из ключевых категорий в ряде относительно новых исследовательских направлений — медиалингвистики, медиастилистики, медиакультуры, медиаобразования. Медиатекст как особый тип речевого продукта неразрывно связан с характеристиками интернет-коммуникации, которые существенно отличают ее от традиционных форм общения[1]. Среди базовых признаков исследователи выделяют электронную форму существования текста, гипертекстуальность, мультимедийность (поликодовость), интерактивность, синхронность/асинхронность общения, вариативность количества и эксплицированности коммуникантов [3]. Каждая из этих характеристик интернет-текстов получила отдельное осмысление в научных работах: например, поликодовость подробно рассматривалась в работах А.Г. Сонина [4], Е.Е. Анисимовой

[1] *Дускаева Л.Р., Протопопова О.В.* Язык сети Интернет // Стилистический энциклопедический словарь русского языка / под ред. М.Н. Кожиной. М. : Флинта, 2003. С. 648–651.

[5]; устно-письменная форма общения анализировалась в трудах Г.Н. Трофимовой и В.В. Барабаша [6] и М.Ю. Казак [7]; интерактивность рассматривалась в трудах В.В. Богуславской и И.В. Богуславского [8] и М.Г. Маклюэна [9]; соотношение традиционных и сетевых жанров показано, например, в работе С.С. Ярцевой [10], посвященной эволюции газетной колонки; особого внимания потребовало также явление электронных сетевых СМИ, исследованное, в частности, А.А. Калмыковым [11].

Проблема определения границ медиатекста и медиадискурса остаётся одной из ключевых в современной медиалингвистике. Большинство существующих подходов опираются на понимание медиатекста как сущности, объединяющей журналистские, рекламные и пиар-тексты [12–14]. При этом ряд исследователей трактуют медиатекст уже, отмечая, что он предстает родовым понятием ряда предшествующих терминов — «журналистский текст», «текст СМИ» и др. [15–18]. Цель настоящего исследования — переосмыслить понятийные границы медиатекста и медиадискурса в контексте цифровой трансформации и медиаконвергенции, в том числе описав новые разновидности медиатекстов, отражающие семиотическую осложненность современной медиакоммуникации.

Материал и методы исследования. Эмпирической базой исследования стал корпус из 200 медиатекстов российских СМИ, бренд-медиа, а также представителей гражданской журналистики, представленных в новейших форматах новых медиа: вертикальных видео, карточках, видеосообщений в Telegram. В корпус, сформированный методом частичной выборки, вошли материалы информационных агентств «РИА Новости» и ТАСС, газет «Аргументы и факты», «Известия», «Комсомольская правда», «Новые известия», интернет-издания «Лента.ру», деловых изданий Russian Business, «Ведомости» и «Коммерсант», телеканалов РЕН ТВ, 360 и др. Корпус сформирован на основании двух критериев: во-первых, регулярность публикации контента в новейших форматах, что свидетельствует о стратегической активности редакций на цифровых платформах; во-вторых, репрезентативное разнообразие типов и направленности медиа: в выборку вошли государственные и частные СМИ, информационные агентства, деловые и общественно-политические издания, телеканалы, бренд-медиа и рядовые пользователи, чьи медиатексты конкурируют за внимание в конвергентной среде. В ходе работы была использована совокупность методов, обеспечивающих комплексный анализ представленных медиатекстов: функционально-стилистический анализ, включающий исследование композиционных параметров медиатекста, а также его графико-стилистических особенностей; контент-анализ, заключающийся в анализе тематических и модусных паттернов; семиотический анализ применён для описания взаимодействия вербального, иконического, мелодического кодов в структуре вертикального медиатекста; дискурс-анализ.

Концептуальные подходы к изучению медиатекста и медиадискурса

Подходы отечественных исследователей к пониманию медиатекста можно условно разделить на три группы, отличающиеся степенью широты трактовки:

1) **медиатекст приравнивается к журналистскому тексту**: В.В. Завацкая: «тексты массовой коммуникации, которые распространяют СМИ» [18. С. 93]; Н.И. Клушина: «совокупность текстов, созданных журналистами и транслируемых на массовую аудиторию через различные медийные каналы (газеты, журналы, радио, телевидение, Интернет)» [16. С. 71]; Г.Я. Солганик: «Медиатекст — разновидность текста, принадлежащая массовой информации, характеризующаяся особым типом автора (принципиальное совпадение производителя речи и её субъекта), специфической текстовой модальностью (открытая речь, многообразное проявление авторского *я*), рассчитанная на массовую аудиторию» [15. С. 15];

2) **в медиатекст включаются рекламные тексты и PR-тексты**: Н.А. Кузьмина: динамическая сложная единица высшего порядка, «посредством которой осуществляется речевое общение в сфере массовых коммуникаций» [12. С. 13]; Я.О. Якуба также выделяет три типа медиатекстов: журналистский, пиар, и рекламный текст [14]; М.Ю. Казак подчеркивает, что термин «медиатекст» закрепился именно за текстами, функционирующими в сфере массовой коммуникации. При этом содержание понятия нуждается в уточнении: «все, что медиатизируется, вовлекается в сферу СМИ, вполне уживается в рамках этого понятия». В то же время «объяснительная сила» данного термина проявляется, когда медиатекст рассматривается как совокупный продукт журналистики, PR и рекламы [7. С. 65]. И. В. Челышева: медиатекст — «текст, созданный средствами и на материале медиа» [19]; Л.Р. Дускаева: особая форма речевой деятельности, отражающая профессионально закреплённые правила речевого поведения и устойчивые стереотипы текстовой организации в медийной среде [13. С. 11]; Г.В. Бобровская: в структуре медиадискурса центральное место занимают тексты журналистики, тогда как PR- и рекламные тексты располагаются на дальней периферии [20];

3) **медиатексты как контент, создаваемый и публикуемый пользователями Интернета в рамках личной коммуникации**: И.Б. Игнатова, Е.С. Зубаркина, А.В. Землянский: «журналистика» не синонимична термину «медиа», такое отождествление представляет собой метонимию — использование частного и легко воспринимаемого термина для обозначения более сложного и многослойного явления [21. С. 19]; А.Г. Качкаева и С.А. Шомова: «медиа» следует понимать как множественное число от «медиум», т.е. средство или посредник в коммуникации [22. С. 27]; М. Маклюэн: медиа — это социальные и культурные практики, сопровождающие человечество на разных этапах цивилизационного развития [9. С. 71]; Г.Н. Трофимова и В.В. Барабаш: медиатексты — «многочисленные блоги, твиты, посты в социальных сетях, рассчитанные на массовую аудиторию, <…> данном сегменте так

называемой гражданской журналистики медиатексты имеют более просторечный характер, приближенный к обиходно-бытовой коммуникации» [6. С. 158].

Т.Г. Добросклонская делит медиатексты на три группы:

1. **Медиатексты первого порядка** — это тексты, создаваемые профессиональными журналистами и распространяемые в онлайн-режиме через официальные медиаплатформы. Они формируют базовый информационный контент цифровой среды и выступают в качестве структурообразующего ядра медиапространства. К ним относятся материалы, публикуемые крупными информационными агентствами, национальными и региональными СМИ, обладающими официальным статусом.

2. **Медиатексты второго порядка** — это материалы, производимые и распространяемые так называемыми индивидуальными медиапроизводителями: блогерами, влогерами, независимыми обозревателями. Эти тексты зачастую стилистически ориентированы на профессиональные образцы, но создаются вне институциональных рамок традиционных СМИ.

3. **Медиатексты третьего порядка** охватывают контент, создаваемый и публикуемый любыми пользователями интернета в рамках личной коммуникации. К этой категории относятся посты в социальных сетях, сторис, видеоролики на YouTube и других платформах, не претендующих на статус журналистских материалов, но активно вовлекающих аудиторию в медийное взаимодействие [23. С. 43—44].

Т.Г. Добросклонская отмечает постепенное снижение монопольного влияния традиционных СМИ и усиление позиций альтернативных источников информации. По справедливому замечанию исследователя, медиатексты второго и третьего порядка всё чаще воспринимаются аудиторией как более подлинные, а в некоторых случаях — более надёжные, чем официальные сообщения. Стремление профессиональной журналистики к объективности и сдержанности, декларируемое как профессиональный стандарт, нередко уступает эмоциональной вовлечённости и динамике пользовательского контента. Данную трансформацию фиксирует и Л.Р. Дускаева, отмечая, что сегодня «поставщиками новости становятся не только СМИ, но и аудитория, последнее стало возможным благодаря блогам, твиттерам, социальным сетям и т.п.». По справедливому замечанию исследователя, «формирование новости становится еще более диалогичным» [24. С. 179].

Классификация медиатекстов, приводимая Т.Г. Добросклонской, коррелирует с типологией современных публикаторов, предложенной исследователями в коллективной монографии «Мультимедийная журналистика» под редакцией А.Г. Качкаевой и С.А. Шомовой, что указывает на расширение исследовательского понимания медиатекста: он рассматривается не только как продукт профессиональной журналистики, но и как результат деятельности различных акторов цифровой среды. Классификация публикаторов, приводимая в монографии, включает: человека как медиа; знаменитость как медиа

(публичные фигуры, способные формировать и управлять собственной медийностью без посредничества СМИ); бренд как медиа (корпорации, конкурируя за внимание аудитории, перенимают у традиционных СМИ журналистские технологии и нанимают профессионалов, тем самым формируя корпоративную журналистику); СМИ как «первородные» медиа. К последнему типу публикаторов исследователи относят классические институции, обладающие опытом, инфраструктурой и профессиональной культурой. Отмечая, что их монополия утрачена, исследователи пишут, что медиатексты, маркированные как профессионально-журналистские, все еще вызывают больше доверия у аудитории [22. С. 53–56].

Эволюция термина «медиатекст» в контексте цифровой коммуникации

Интерпретации медиатекста, выходящие за рамки его понимания как исключительно журналистского текста, соотносятся с процессами трансформации самой журналистики под влиянием конвергенции. Конвергенция устраняет границы между видами СМИ, жанрами, профессиональными журналистами и активными участниками сетевой коммуникации. А.А. Калмыков определяет конвергенцию как сближение, происходящее одновременно в нескольких плоскостях: на уровне технических устройств, используемых журналистами для сбора и обработки материала; на уровне профессионализма, поскольку новые цифровые инструменты требуют дополнительных компетенций и задают конвергентные форматы предоставления контента; а также на уровне компонентов системы СМИ (интернет-ТВ, онлайн-газета, интернет-радио) [11]. Е.Л. Вартанова отмечает: «Очевидно, что в ходе трансформаций последних десятилетий, прежде всего перехода к рынку и технологической революции, само понятие СМИ изменилось и расширилось. К традиционным — газете, журналам, ТВ и радио — все чаще стали причислять телефильмы, кино, книги. При этом главное изменение в медиасистеме связано с вхождением в нее новых медиа и Интернета, их современной инфокоммуникационной среды, действующей на многочисленных технологических платформах проводных и беспроводных компьютеров, мобильных телефонов, планшетах» [25. С. 3].

К похожим выводам приходит В.В. Богуславская и И.В. Богуславский, отмечая, что «цифровизация стирает различия между СМИ и типами медиаконтента»: инфокоммуникационная среда «дополняет содержание традиционных СМИ контент-продуктами новых медиа» [8. С. 51]. По наблюдению исследователя, современный медиатекст перестаёт «распознаваться» в традиционном смысле и всё чаще «формируется» самим читателем-пользователем. Массовая коммуникация принимает характер «кастомизированной» и «нишевой», ориентированной на индивидуальные интересы аудитории. В этой связи медиатекст оказывается гибридным образованием, объединяющим в себе черты журналистского текста и пользовательского медиаконтента.

Для обозначения роли пользователя в современной медиасреде исследователи вводят понятия «человека коммуникационного» (А.Г. Качкаева), «просьюмера» (впервые термин использован Э. Тоффлером), produsage (A. Bruns), «авторствующей публики» (J. Rosen), «активного (со)производителя» (Е.Л. Вартанова) и другие, суть которых сводится к тому, что сегодня, как отмечает Г. Дженкинс, «участники новой медийной культуры не просто пассивно потребляют медиапродукты, но становятся их соавторами» [26. P. 3]. Действительно, с развитием Интернета число потенциальных авторов выросло кратно: пользователи получили возможность публикации собственных высказываний без посредничества и формального допуска к медиаинфраструктуре. Это явление получило название «взрыва авторства».

Очевидно, что изменение характера медиапотребления и медиапроизводства, связанное с феноменами экранной культуры, клипового мышления, поточного сознания, децентрализации СМИ и «взрыва авторства», изменило суть медиатекстов. Так, границы между типами медиатекстов сегодня являются подвижными. В условиях функционирования глобального медиапространства они постоянно размываются под воздействием объективных факторов, связанных с динамикой информационного обмена и цифровой инфраструктурой. Так, то, что Т.Г. Добросклонская называет медиатекстом третьего порядка — индивидуальное сообщение, изначально не претендующее на массовое внимание, — может приобрести черты медиатекста первого порядка (профессионального журналистского текста) в случае широкого резонанса. В то же время значительно чаще наблюдается обратное явление: медиатексты первого порядка становятся катализатором для создания медиатекстов второго и третьего порядка — блогов, комментариев, мемов и пользовательских реплик в социальных сетях.

Как было выяснено в ходе исследования, трансформация сущности медиатекстов в СМИ происходит под влиянием трех основных процессов: адаптации контента СМИ в социальных сетях и Telegram-каналах под требования платформ; внедрения новых форматов и механик на сайты СМИ; слияния журналистского, рекламного и пользовательского текста в бренд-медиа (речь идет о тематических изданиях, созданных компанией, которые при этом могут существовать независимо от ее имиджа и целевой аудитории).

Действительно, показательным примером размытия границ медиатекста становится именно феномен бренд-медиа: с одной стороны, они копируют журналистские практики, создавая статьи в традиционных журналистских жанрах (например, «ПСЖР» — бренд-медиа Aviasales о путешествиях[2] (рис. 1); «Т–Ж» — бренд-медиа «Т-Банка» о финансах[3], Flacon — журнал о

[2] ПСЖР — журнал о путешествиях от Aviasales Режим доступа: https://www.aviasales.ru/faq/zhurnal-psgr (дата обращения: 05.10.2025).

[3] Т–Ж: бренд-медиа Т-Банка Режим доступа: https://t-j.ru (дата обращения: 05.10.2025).

красоте торговой сети «Золотое яблоко»[4]), с другой — остаются инструментами продвижения собственных продуктов.

Обратимся, например, к материалу «860 километров за 2 недели: как я с семьёй объехала весь Кольский полуостров на автодоме», опубликованному в «ПСЖР» в июле 2025 г.[5]. Автор медиатекста — профессиональный тележурналист Ксения Нестерова, работающая на ГТРК «Мурман»[6]. Материал содержит черты классического репортажа: здесь есть точная фактура (географические указания, данные о расстояниях и времени пути, конкретные объекты инфраструктуры, детали культурного, исторического и природного контекста: *«Добирались до неё мы почти семь часов, по дороге останавливаясь, чтобы насладиться красотами и согреться кофе (на момент поездки, 15 июля, было +16 °С)»; «В маршрут вошли восемь основных точек от села Варзуга в Мурманской области до Териберки включительно. Мы посетили посёлки Кузомень, Умба, Колвица, а также города Кандалакша, Кировск и Мурманск — это почти 24 часа в пути, если ехать без перерывов»; «То ли из-за того, что её окружали неповторимые северные ландшафты, то ли из-за своего необычной архитектуры: эта церковь была построена без единого гвоздя ещё в XVII веке»)*. Медиатекст демонстрирует использование типичных для журналистских интернет-публикаций инструментов мультимедийной связности — прежде всего гиперссылок, встроенных в топонимы и объекты маршрута, что обеспечивает возможность перехода к внешним источникам (например, картам или справочным материалам) и повышает интерактивность восприятия. Кроме того, автор создает эффект присутствия (феномен, при котором человек, читающий репортаж, ощущает себя участником или очевидцем описываемого события). Автор целенаправленно формирует контекст — в журналистском смысле этого термина, то есть совокупность сопутствующих данных, позволяющих читателю соотнести личный опыт с пространственными и временными координатами повествования (*«Суммарное расстояние, которое нам удалось проехать за всё путешествие, превысило 860 километров — это как от Москвы до Нижнего Новгорода и обратно»*).

В то же время данный материал сложно назвать медиатекстом первого порядка, поскольку он содержит черты медиатекстов, распространяемых в блогах: личные и эмоциональные описания (*«Когда мы наконец доехали… на ули-це ещё светило солнце… появилась радуга — это зрелище было лучше любой достопримечательности»*), представление автора, оформленное как восклицательное предложение, которое открывает текст: *«Меня зовут Ксюша, и я езжу в малоизвестные города России: за последние годы*

[4] Flacon Режим доступа: https://flacon-magazine.com/ (дата обращения: 05.10.2025).

[5] *Нестерова К.* 860 километров за 2 недели: как я с семьёй объехала весь Кольский полуостров на автодоме // ПСЖР — журнал Aviasales о путешествиях. 31.07.2025. Режим доступа: https://www.aviasales.ru/psgr/article/na-avtodome-po-kolskomu (дата обращения: 05.10.2025).

[6] *Нестерова К.* Профиль журналиста на сайте ГТРК «Мурман» Режим доступа: https://murman.tv/author-n-38--kseniya-nesterova (дата обращения: 05.10.2025).

посетила более ста!». Наконец, данный материал дополнен продающим элементом: встроенной плашкой о возможности купить дешевый билет до Мурманска через телеграм-бот Aviasales. В результате материал представляет собой симбиоз журналистского, пользовательского, рекламного медиатекста.

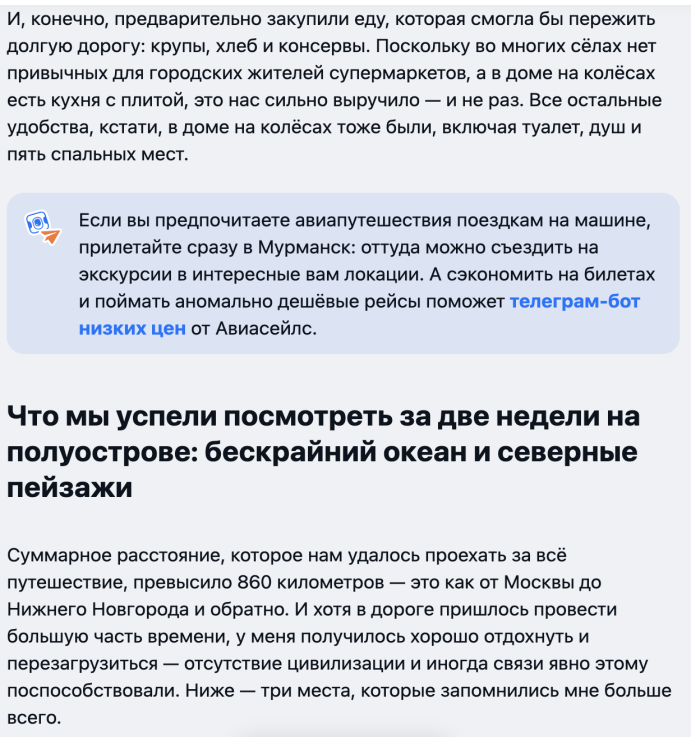


Рис. 1. Встраивание рекламного блока в журналистский материал на сайте ПСЖР (бренд-медиа Aviasales): интеграция телеграм-бота для покупки авиабилетов внутри авторского тревел-репортажа

Источник: составлено С.С. Хромовым, А.А. Распоповым.

Figure 1. Embedding of an advertising block in the journalistic material on the PJWR website (brand Aviasales): integration of a telegram bot for purchasing airline tickets within the author's travel report report

Source: compiled by Sergey S. Khromov & Artyom A. Raspopov.

Подобная гибридность характерна и для социальных сетей, где бренды копируют подходы блогеров: так, бренд «Русский продукт» продвигает овсянку «Геркулес» в запрещённой в России сети Instagram*, создавая постироничные вертикальные видео о пользе каши, встраивая рекламное содержание в популярные на платформе тренды, деконструируя их[7] (рис. 2).

[7] Геркулес «Русский продукт». Официальная страница бренда в запрещенной на территории РФ социальной сети Instagram. Режим доступа: https://www.instagram.com/moygerkules?igsh=MXhnN3E4cWRqYWUwag== (дата обращения: 05.10.2025).

* Instagram – организация, деятельность которой признана нежелательной/экстремистской и запрещена на территории Российской Федерации.

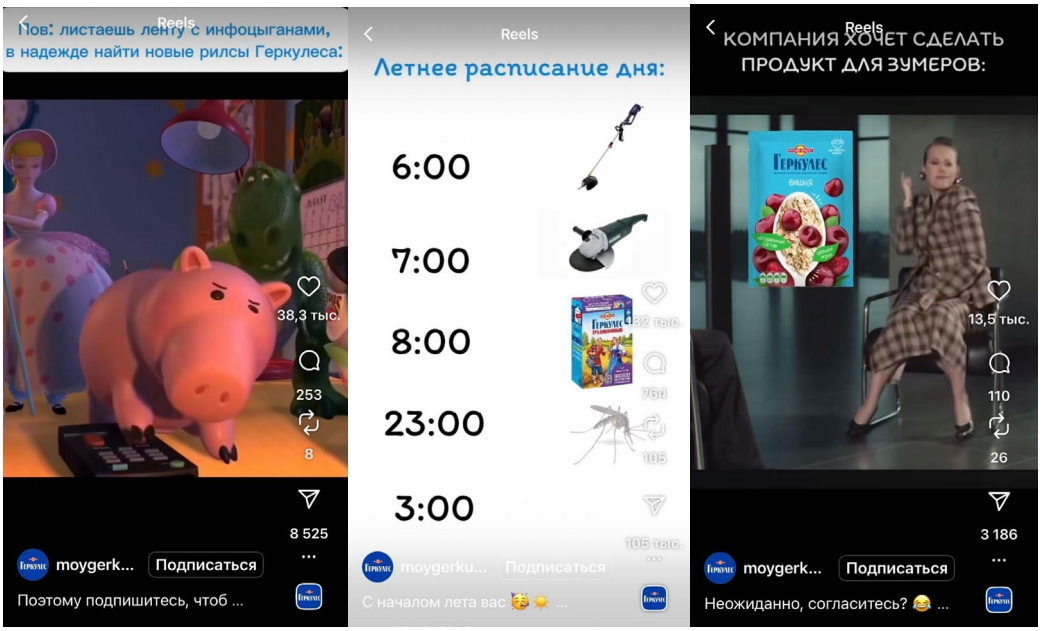


Рис. 2. Скриншоты из официального аккаунта бренда «Геркулес» (компания «Русский продукт») в запрещенной в России социальной сети Instagram*: примеры постироничных вертикальных видео, встроенных в актуальные тренды платформы

Источник: составлено С.С.Хромовым, А.А. Распоповым.

Figure 2. Screenshots from the official account of the brand "Hercules" ("Russian product company") on the social network Instagram*, which is banned in Russia: examples of post-ironic vertical videos embedded in current platform trends.

Source: compiled by Sergey S. Khromov & Artyom A. Raspopov.

Платформенная дистрибуция контента меняет логику функционирования и медиатекстов СМИ, распространяемых ими в официальных каналах в мессенджерах и социальных сетях.

Ярким примером смешения медиатекстов служит ситуация, когда пост в социальной сети или публикация в Telegram-канале получает широкое распространение благодаря «перепечатке» крупного СМИ. Так, многие российские издания регулярно создают подборки «лучших шуток в соцсети X» или «трендов TikTok», создавая свои материалы на основе постов пользователей (рис. 3).

Стоит также отметить, что СМИ все активнее копируют подходы блогеров к созданию материалов, создавая мемы, игры или тесты по мотивам своих текстов. Размывание институциональных границ медиапространства вынуждает СМИ адаптироваться к требованиям аудитории платформ. В этой среде

* Instagram – организация, деятельность которой признана нежелательной/экстремистской и запрещена на территории Российской Федерации.
Instagram – organizations whose activities have been officially designated as undesirable or extremist within the territory of the Russian Federation.

пост человека без значительной базы подписчиков в формате «гражданской журналистики» благодаря алгоритмам нередко набирает больший охват и отклик, чем сообщение СМИ о той же проблеме. Например, частая ситуация в запрещенной в России социальной сети Instagram* демонстрирует смещение баланса в медиапространстве: вертикальное видео или пост в формате «карусели», представляющей собой серию изображений, объединенных общей темой, способны набрать охват, превышающий миллион просмотров, даже у пользователя с аудиторией в несколько десятков подписчиков, если тема и подача оказываются интересны аудитории. В таких случаях публикации институциональных СМИ нередко проигрывают по показателям вовлеченности и охвата, что иллюстрирует децентрализацию производства массовой информации в цифровой среде (рис. 4).

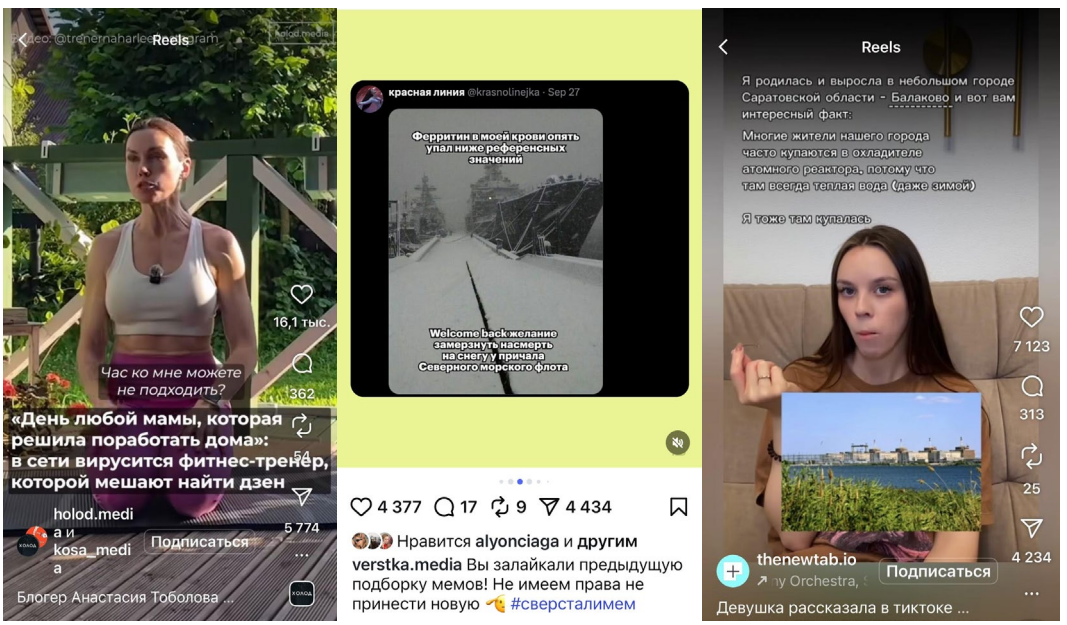


Рис. 3. Медиапрактика «обзора» пользовательского контента: публикации российских медиа, превращающих посты пользователей в самостоятельные материалы (скриншоты из Instagram*)

Источник: составлено С.С.Хромовым, А.А. Распоповым.

Figure 3. Media practice “review” user content: publications of Russian media, turning users’ posts into independent materials (screenshots from the social network Instagram*)

/ *Source:* compiled by Sergey S. Khromov & Artyom A. Raspopov.

* Instagram – организация, деятельность которой признана нежелательной/экстремистской и запрещена на территории Российской Федерации.

Instagram – organizations whose activities have been officially designated as undesirable or extremist within the territory of the Russian Federation.

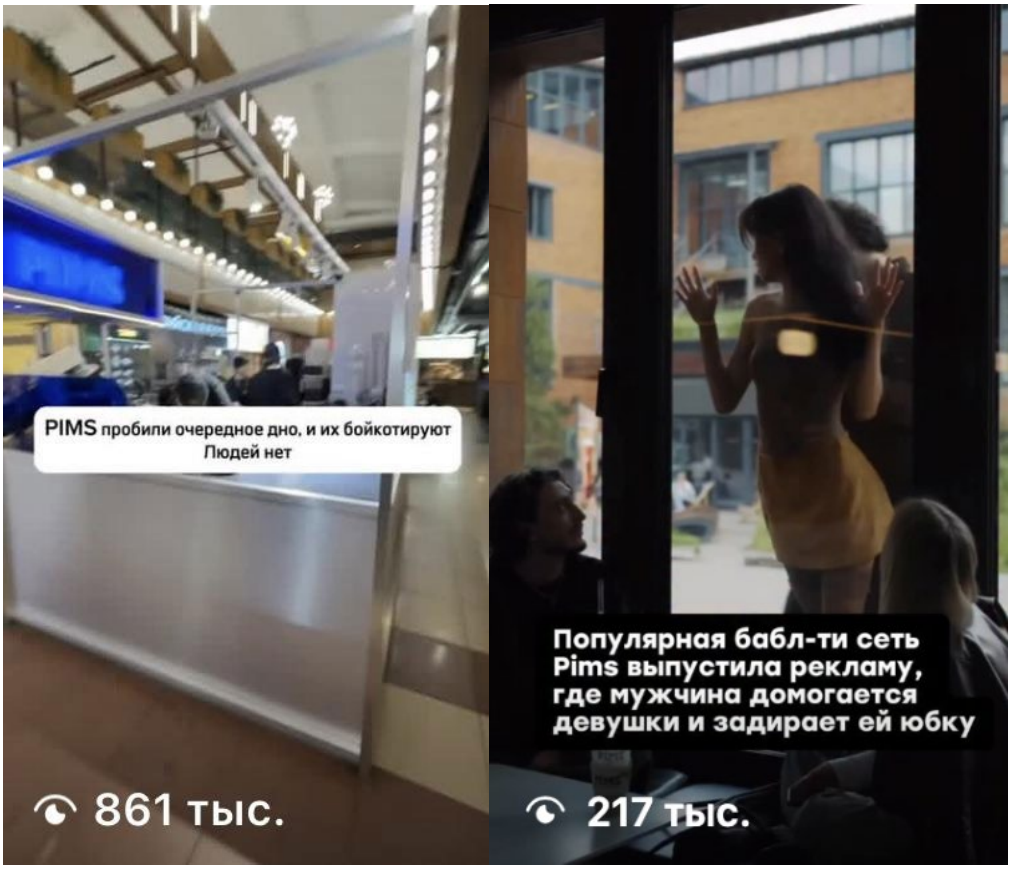


Рис. 4. Иллюстрация децентрализации медиапространства: пользовательский ролик о скандале вокруг сети кофеен Pims (861 тыс. просмотров) набрал больший охват, чем профессиональный сюжет медиа «Коса» на ту же тему (217 тыс. просмотров)
Источник: составлено С.С. Хромовым, А.А. Распоповым.
Figure 4. Illustration of media decentralization: A user's video about the scandal around the coffee chain Pims (861 thousand views) gained more coverage than the professional story of "Kosa" on the same topic (217 thousand views)
Source: compiled by Sergey S. Khromov & Artyom A. Raspopov.

Наиболее репрезентативным форматом СМИ в платформенной медиасреде становятся короткие вертикальные видео. Они отвечают характеристикам медиапотребления в цифровую эпоху, ориентированного в первую очередь на визуальные форматы, что обусловлено в том числе когнитивными изменениями: снижением способности к длительной концентрации и запросом на быстрое получение дофаминовых стимулов. Как отмечал А.А. Амзин еще в 2016 г., у поколения школьников наблюдается явное предпочтение визуального контента, в том числе из-за того, что видеопотребление занимает время, несоразмерное с прочтением текстового эквивалента [27. С. 149]. А.В. Градюшко также подчеркивает, что «визуальный контент становится доминирующим в мобильной среде» [28. С. 68]. Особое место в этой визуальной трансформации занимает короткое вертикальное видео — формат, выросший

из повседневной мобильной практики и приспособленный под экраны смартфонов. Крайняя востребованность формата вертикального видео подтверждается статистикой российского медиапотребления. По данным опроса Anketolog.ru и Sixth Sense (февраль 2024), 49 % интернет-пользователей в России смотрят в соцсетях короткие вертикальные видео[1].

Важнейшей характеристикой вертикальных видео является алгоритмизированная дистрибуция в условиях дефицита внимания — контент распространяется в бесконечной ленте рекомендаций, формируемой алгоритмами на основе поведенческих метрик пользователя. Пользователь обыкновенно не ищет материал целенаправленно и не погружается в контекст, а просматривает его «на бегу» в потоке не связанных сообщений. Функциональные следствия этого фактора: малый хронометраж (например, в TikTok 86 % вертикальных видео имеют длительность меньше одной минуты, а ролики до 15 секунд получают наивысший уровень вовлеченности); высокая визуальная и эмоциональная насыщенность; снижение роли дискуссии (пользователь может лайкнуть или репостнуть ролик, но не вовлекается в полноценное обсуждение, а комментарии носят эпизодический характер); алгоритмическая пессимизация социально-политических тем; размывание институциональных границ (в ленте контент СМИ конкурирует с блогерским и пользовательским контентом)[2]; интеграция мультимодальных элементов [29].

Такая среда определяет семиотические и композиционные свойства медиатекстов. В частности, вертикальные видео представляют собой поликодовые образования, в структуре которых можно выделить три компонента: вербальный, иконический и мелодический, которые находятся в различных отношениях друг с другом. Из 64 коротких вертикальных видео российских СМИ, проанализированных в ходе этой части исследования, 33 видео представляют собой ролики с доминирующим иконическим компонентом, 17 — с оппозиционной корреляцией (особенностью таких видео является то, что в них вследствие противоречия между содержанием, отраженным вербальным и невербальным компонентами, возникает комический эффект), 14 — с доминирующим вербальным компонентом. Высокая доля оппозиционной корреляции свидетельствует о меметизации дискурса и внедрении в профессиональные медиатексты приёмов пользовательской культуры. При этом для описания ряда таких роликов нами предлагается ввести понятие «синергетического знака», определяемого как знак, состоящий из компонентов разных кодов, лишь тесное взаимодействие которых рождает глубинный, непосредственно не наблюдаемый смысл образования. Важно также отметить, что медиатексты российских СМИ, распространяемые в вертикальных видео, копируют не только формальные, но и содержательные особенности

[1] 49 % россиян «залипают» в короткие вертикальные видео. Эксклюзивные данные // Cossa.ru. Режим доступа: https://www.cossa.ru/news/328823/ (дата обращения: 23.06.2025).

[2] *Guan S.* 80 Vital TikTok Statistics for 2025 // Thunderbit. 05.06.2025. Режим доступа: https://thunderbit.com/blog/tiktok-stats (дата обращения: 06.10.2025).

формата: так, среди роликов российских СМИ доминирует развлекательный контент про животных.

Проведя комплексное исследование вертикального видеотекста, предлагаем следующее определение данного формата. Вертикальный видеотекст — это мультимодальный медиатекст, созданный с ориентацией на вертикальное (портретное) восприятие экрана, обладающий синкретичными признаками информационного, публицистического, рекламного и развлекательного дискурсов. Он характеризуется поликодовостью, фрагментарной структурой, экспрессивной подачей и монтажной логикой синтаксической организации. Вертикальный видеотекст ориентирован на цифровую среду с коротким временем удержания внимания, поэтому отличается сжатием информации, квазидиалогичностью, высокой прагматической нагруженностью и использованием вербальных, визуальных, мелодических кодов в их синергетическом взаимодействии.

Другим примером формата, в котором отражается смешение дискурсов в социальных сетях, является формат карточек (рис. 5 и 6) — постов из нескольких слайдов, посвященных одной теме, сочетающих текст, фото, инфографику и цитаты, и зачастую снабженных музыкальным треком в качестве подложки. В этом формате СМИ представляют репортажи, интервью, новостные заметки и аналитические статьи, но пользователем все они воспринимаются как единый тип медиатекста — «карточки». Жанровая принадлежность оказывается размыта и отходит на второй план по отношению к форматной. Исходя из этого, в современных условиях формат можно рассматривать как наджанровую единицу медиапроизводства, то есть как устойчивую модель организации, композиции и дистрибуции медиатекста, задающую его визуально-коммуникативную структуру, режим потребления и характер взаимодействия с аудиторией. В отличие от жанра, фиксирующего содержательные и стилистические параметры высказывания, формат задаёт технологические и организационные рамки медиатекста, в которых формируются его языковые, структурные и прагматические характеристики.

Ещё одним показателем расширения поля понятия медиатекста можно считать использование формата видеосообщений («кружков») в мессенджере Telegram. Первоначально этот формат принадлежал сфере личной коммуникации: он служил инструментом неформального обмена сообщениями между пользователями мессенджера, обладая характерными признаками бытовой спонтанности, ограниченной длительности и вертикальной ориентации кадра. Однако в последние годы данный формат активно осваивается медиапроектами. Показателен пример медиа FOMOTEKA, функционирующего по модели городского СМИ[3]. В своих публикациях редакция использует «кружки» как инструмент мгновенного визуального включения аудитории в пространство события: короткие видеосообщения создают эффект

[3] FOMOTEKA. Медиа о мемо- и мимолетной жизни. Режим доступа: https://t.me/fomoteka (дата обращения: 10.11.2025).

присутствия, характерный для репортажа. При этом текстовая часть выполняет аналитическую функцию, а визуальный элемент — эмоционально-иллюстративную. Интеграция такого «пользовательского» формата в медиапрактику изданий демонстрирует смещение границ между личной и профессиональной коммуникацией и подтверждает тенденцию к расширению понятия медиатекста.

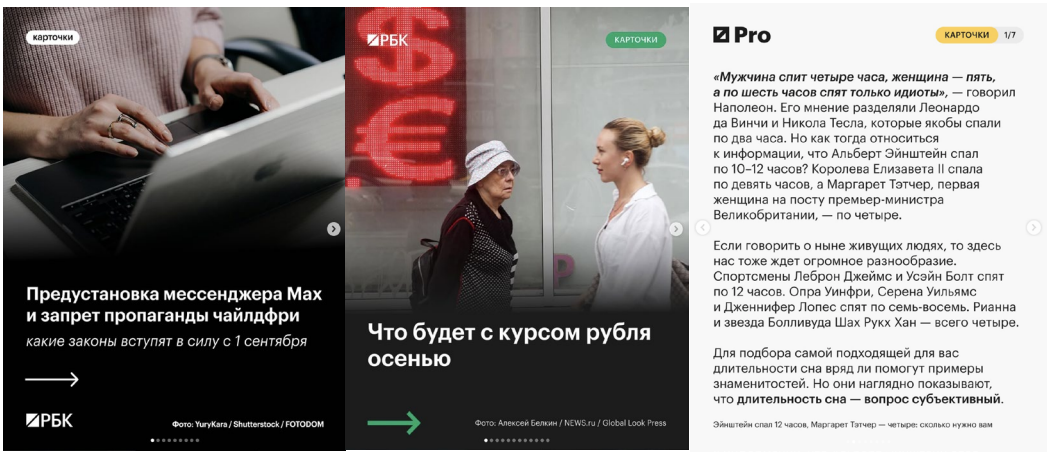


Рис. 5. Примеры карточек из Instagram*-аккаунта издания РБК: новостной дайджест (*слева*), проблемно-аналитический материал (*в центре*) и научпоп-заметка (*справа*)[1]
Источник: составлено С.С. Хромовым, А.А. Распоповым.
Figure 5. Examples of cards from the Instagram* account of the RBC edition: news digest (*left*), problem-analytical material (*center*) and science post (*right*).
Source: compiled by Sergey S. Khromov & Artyom A. Raspopov.

Аргументом в защиту «форматоцентричного» подхода к пониманию медиатекстов в цифровой среде может служить факт интеграции вертикальных видео не только в ведущие соцсети (запрещенная на территории РФ социальная сеть Instagram (Reels, 2020), VK (Клипы, 2020)), но и видеохостинги (YouTube (Shorts, 2021), Rutube), блог-платформу «Дзен», маркетплейсы (так, летом 2025 г. Wildberries запустил свою видеоплатформу с вертикальными видео Wibes), игровые онлайн-платформы (осенью 2025 года стало известно, что Roblox внедряет ленту вертикальных видео в приложение), банковские приложения и собственно сайты СМИ. Так, проект «ВЕДы», запущенный 27 октября 2025 г., представляет собой молодежную версию делового издания «Ведомости», доступ к которой осуществляется через основной сайт газеты: после появления всплывающего окна с вопросом «Вам больше 30 или меньше 30?» пользователи, выбирающие вариант «меньше 30», перенаправляются на

* Instagram – организация, деятельность которой признана нежелательной/экстремистской и запрещена на территории Российской Федерации.
Instagram – organizations whose activities have been officially designated as undesirable or extremist within the territory of the Russian Federation.

отдельную платформу[2]. Интерфейс ресурса имитирует визуальную логику мобильных приложений и сервисов коротких видео: вертикальная ориентация контента, минималистичное черно-белое оформление, крупная верстка и навигация с помощью иконок-эмодзи между рубриками вертикальных видео, обозначенных английскими словами (Tech, Work, Trendz, Politics и др.). Основу ресурса составляют минутные вертикальные ролики, а также текстовые материалы, стилистически приближенные к разговорной интернет-коммуникации. Тематический диапазон включает социально-экономические сюжеты, отражающие повседневный опыт поколения Z (финансовая грамотность, рынок труда, технологические тренды, медиапотребление и др.). Данный пример наглядно демонстрирует трансформацию институциональных СМИ — переход от традиционной газетной модели к мультимедийному, вертикально-ориентированному способу взаимодействия с аудиторией.

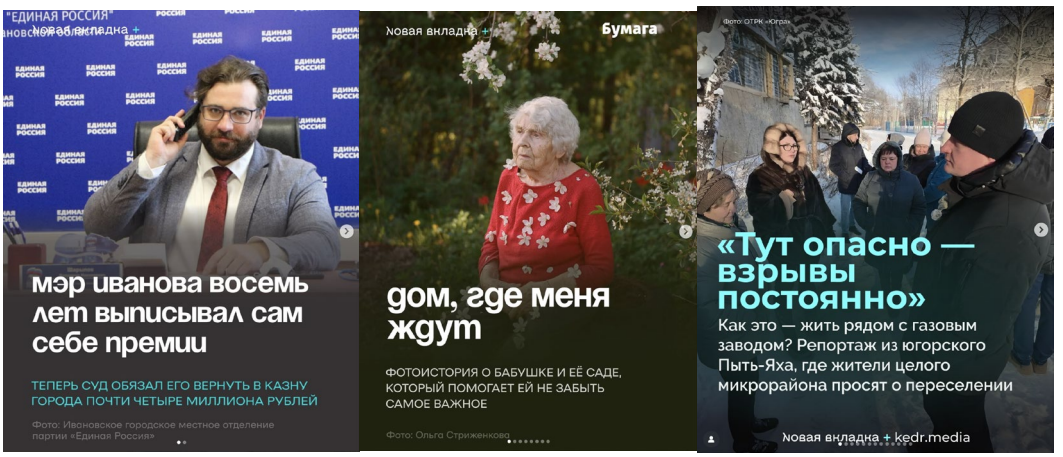


Рис. 6. Примеры публикаций в формате карточек издания «Новая вкладка» (скриншоты соцсети Instagram*): новостная заметка (*слева*), фоторепортаж (*в центре*) и репортаж (*справа*)

Источник: составлено С.С.Хромовым, А.А. Распоповым.

Figure 6. Examples of publications in the format of cards of the edition "New tab" (screenshots from the account of the social network Instagram*): news article (*left*), photo report (*center*) and report (*right*)

Source: compiled by Sergey S. Khromov & Artyom A. Raspopov.

Заключение

Таким образом, медиатекст как основную форму существования языка в цифровую эпоху можно определить как текст, формирующийся

[2] ВЕДы — деловое издание газеты «Ведомости» для зумеров Режим доступа: https://ved.vedomosti.ru/ (дата обращения: 28.10.2025).

* Instagram – организация, деятельность которой признана нежелательной/экстремистской и запрещена на территории Российской Федерации.

Instagram – organizations whose activities have been officially designated as undesirable or extremist within the territory of the Russian Federation.

в медиасфере в условиях мультимодальной и интерактивной коммуникации и представляющий собой гибридную форму текста на стыке журналистских и пользовательских практик. Основные признаки медиатекста — медийность, семиотическая осложненность, массовость, открытость. Именно поликодовая природа обеспечивает медиатексту его выразительные и коммуникативные возможности, особенно ярко проявляясь в медиапрактиках. Особой разновидностью медиатекста является вертикальный видеотекст, который предстает наиболее репрезентативным форматом платформенной медиасреды. Его можно определить как мультимодальный медиатекст, созданный с ориентацией на вертикальное (портретное) восприятие экрана, обладающий синкретичными признаками информационного, публицистического, рекламного и развлекательного дискурсов, характеризующийся фрагментарной структурой, экспрессивной подачей, монтажной логикой синтаксической организации, а также квазидиалогичностью, высокой прагматической нагруженностью и использованием вербальных, визуальных, мелодических кодов в их синергетическом взаимодействии.

Современный медиатекст — более широкое понятие, чем «журналистский», «рекламный» или «пиар-текст»; он охватывает их, но не исчерпывается ими, включая весь спектр публичных высказываний, циркулирующих в институциональных СМИ, блогах и на гибридных цифровых платформах. Актуальной средой функционирования медиатекстов выступает прежде всего интернет, где медиатексты образуют массив сообщений, с которыми пользователь сталкивается при обращении к социальной сети или цифровой платформе. Выполняя функции трансляции информации о событиях и актуализации культурно специфических знаний, медиатекст конструирует социокультурно значимые феномены, обладает высоким перлокутивным потенциалом и становится одним из ключевых объектов анализа в современной гуманитарной науке.

Его структура отражает особенности медиадискурса — контекстуально обусловленного, институционально маркированного и технологически опосредованного — который предлагается определять как комплекс текстов и практик, публично распространяемых через медийные каналы и платформы, вне зависимости от того, кем они созданы. Такой подход позволяет учитывать гибридную структуру современной медиасреды, в которой журналистские, пользовательские (т.н. персональный журнализм) и рекламные медиатексты образуют единую систему массовой коммуникации (алгоритмы социальных медиа стирают границы между медиатекстами), влияющую на общественные процессы и формирование смыслов.

Список литературы

1. *Попова Т.И.* Актуальные направления исследования медийного интернет-пространства // Медиалингвистика. 2018. Т. 5. № 3. С. 258–272. EDN: YWWNXV
2. *Воробьев В.В.* Лингвокультурология в условиях вызовов XXI в.: новые тренды в гуманитарном образовании // Русский язык в поликультурном мире: IX Международный

симпозиум (8–12 июня 2025 г.) : сб. науч. статей. / отв. ред. Т.В. Аржанцева, Л.А. Орехова. Симферополь : Издательский дом КФУ, 2025. С. 10–14. EDN: NZLROR
3. *Щипицина Л.Ю.* Компьютерно-опосредованная коммуникация. Лингвистический аспект анализа : монография. М. : КРАСАНД, 2010.
4. *Сонин А.Г.* Моделирование механизмов понимания поликодовых текстов: автореф. дис. ... д-ра филол. наук. М. : МГЛУ, 2006. EDN: NJVLAZ
5. *Анисимова Е.Е.* Лингвистика текста и межкультурная коммуникация (на материале креолизованных текстов). М. : Тезарус, 2013.
6. *Трофимова Г.Н., Барабаш В.В.* Языковое послевкусие интернет-эпохи в России: эффект бумеранга (актуальные процессы в русскоязычной цифровой медиакоммуникации): монография. М. : РУДН, 2020. EDN: GCYADT
7. *Казак М.Ю.* Современные медиатексты: проблемы идентификации, делимитации, типологии // Медиалингвистика. 2014. № 1(4). С. 65–76. EDN: SYMXGV
8. *Богуславская В.В., Богуславский И.В.* Медиатекст и хэштеги: цифровая трансформация СМИ // Гуманитарный вектор. 2017. Т. 12. № 5. С. 51–58. https://doi.org/10.21209/1996-7853-2017-12-5-51-58 EDN: ZTNELH
9. *Маклюэн М.Г.* Понимание медиа: внешние расширения человека. М. : Кучково поле, 2019.
10. *Ярцева С.С.* Колумнистика: история возникновения и перспективы развития : автореф. дис. ... канд. филол. наук. Воронеж : Воронеж. гос. ун-т, 2011. EDN: ZOJGUZ
11. *Калмыков А.А.* Конвергенция — возможность универсального журнализма в рамках профессиональной идентичности // Оптимальные коммуникации. 2011. Режим доступа: http://jarki.ru/wpress/2011/04/28/2198/ (дата обращения: 04.10.2025).
12. *Кузьмина Н.А.* Современный медиатекст. Омск : Полиграфический центр «Татьяна», 2011. EDN: QXGBOT
13. *Дускаева Л.Р.* Интенциональность медиатекста как детерминант речевого поведения в медиасфере // Медиатекст как полиинтенциональная система : сб. статей / отв. ред. Л.Р. Дускаева, Н.С. Цветова. СПб. : Санкт-Петербургский государственный университет, 2012. С. 10–16.
14. *Якуба Я.О.* Изучение медиатекста в семиотике: методологические аспекты // Бизнес. Общество. Власть. 2013. № 14. С. 55–63. EDN: STSIUB
15. *Солганик Г.Я.* К определению понятий «текст» и «медиатекст» // Вестник Московского университета. Серия 10: Журналистика. 2005. № 2. С. 7–15. EDN: JXDHGL
16. *Клушина Н.И.* Медиастилистика // Экология языка и коммуникативная практика. 2014. № 2. С. 69–77. EDN: SVYYBV
17. *Борисова С.С.* Реализация персуазивных стратегий в аналитическом жанре «передовая статья» (Leitartikel) как ресурс информационного доминирования // Научный вестник Воронежского государственного архитектурно-строительного университета. Серия: Современные лингвистические и методико-дидактические исследования. 2015. № 1 (25). С. 102–114. EDN: TPUVVX
18. *Завацкая В.В.* К определению понятия медиатекст // Лингвистика, лингводидактика, лингвокультурология: актуальные вопросы и перспективы развития : материалы IV Международной научно-практической конференции (Минск, 19–20 марта 2020 г.) / под ред. О. Г. Прохоренко, О. В. Дубровиной, А. И. Головня и др. Минск : Белорусский государственный университет, 2020. С. 92–95. EDN: XWEMHN
19. *Челышева И.В.* Методы работы с медиатекстом: философские аспекты проблемы // В.В. Савчук. Медиафилософия. СПб. : Издательство Санкт-Петербургского философского общества, 2008. С. 325–329. EDN: WAZGAL
20. *Бобровская Г.В.* Когнитивно-элокутивный потенциал газетного дискурса : монография. Волгоград : Перемена, 2011. EDN: SUHLKN
21. *Игнатова И.Б., Зубаркина Е.С., Землянский А.В.* Новые медиа: теория и практика. М. : МПГУ, 2022. https://doi.org/10.31862/9785426310957 EDN: TSSGTC

22. *Качкаева А.Г., Шомова С.А*. Мультимедийная журналистика. М. : НИУ ВШЭ, 2017. https://doi.org/10.17323/978-5-7598-1189-3 EDN: XYMVJJ
23. *Добросклонская Т.Г.* Медиалингвистика: теория, методы, направления. М. : Добросвет, 2020. https://doi.org/10.31453/kdu.ru.91304.0107 ISBN: 978-5-7913-1128-3 EDN: CQKWES
24. *Дускаева Л.Р.* Речевая структура новостных текстов сетевых изданий // Вестник Пермского университета. Российская и зарубежная филология. 2014. № 1(25). С. 179–183. EDN: QEJSHA
25. *Вартанова Е.Л.* Язык российских СМИ как индикатор социальных перемен // Г.Я. Солганик. Язык СМИ и политика. М. : Изд-во МГУ, 2012. С. 3–7.
26. *Jenkins H*. Convergence Culture: Where Old and New Media Collide. N.Y.; L. : New York University Press, 2006.
27. *Амзин А.А.* Новые и экспериментальные форматы // С.Л. Балмаева, М.Ю. Лукина. Как новые медиа изменили журналистику: 2012—2016. М. : НИУ ВШЭ, 2016. С. 140–151.
28. *Градюшко А.В*. Ключевые тенденции развития зарубежных интернет-СМИ в условиях дигитализации медиасреды // Журналистика–2017: состояние, проблемы и перспективы : материалы 19-й Международной научно-практической конференции, Минск, 9–10 ноября 2017 г. Вып. 19. Минск : Белорусский государственный университет, 2017. С. 66–71.
29. *Lang K.* Longer TikToks Get More Views – Data Shows Best TikTok Video Length [Электронный ресурс] // Buffer, 17.03.2025. Режим доступа: https://buffer.com/resources/longer-tiktoks-get-more-views-data/ (дата обращения: 06.10.2025).

References

1. Popova, T.I. (2018). The Actual Trends of Media Internet Space Researches. *Media Linguistics*, *5*(3), 258–272. https://doi.org/10.21638/spbu22.2018.301 (In Russ.). (In Russ.). EDN: YWWNXV
2. Vorobyev, V.V. (2025). Linguvoculturology in the Context of Challenges of the XXIst Century: New Trends in Humanitarian Education. In: *Russian Language in a Multicultural World: IXth International Symposium (June 8–12, 2025): Research articles. T.V. Arzhantseva, L.A. Orekhov* (eds.) (pp. 10–14). Simferopol: KFU publ. (In Russ.). EDN: NZLROR
3. Schipizina, L.U. (2010). *Computer-Mediated Communication. Linguistic Aspect of Analysis*: Monograph. Moscow: KRASAND. (In Russ.).
4. Sonin, A.G. (2006). *Modeling Mechanisms of Understanding Polycoded Texts* [PhD Thesis]. Moscow: Moscow State Linguistic University publ. (In Russ.). EDN: NJVLAZ
5. Anisimova, E.E. (2013). *Text Linguistics and Intercultural Communication (On the Material of Creolized Texts)*. Moscow: Tesarus. (In Russ.).
6. Trofimova, G.N., & Barabash, V.V. (2020). *Linguistic Postcusis of the Internet Age in Russia: the Effect of Boomerang (Current Processes in Russian Digital Media Communication):* Monograph. Moscow: RUDN publ. (In Russ.). EDN: GCYADT
7. Kazak, M.Yu. (2014). Modern Media Texts: Issues of Identification, Delimitation and Typology. *Media Linguistics*, *1*(4), 65–76. (In Russ.). EDN: SYMXGV
8. Boguslavskaya, V.V., & Boguslavskiy, I.V. (2017). Media Text and Hashtags: Digital Media Transformation. *Humanitarian Vector*, *12*(5), 51–58. (In Russ.). https://doi.org/10.21209/1996-7853-2017-12-5-51-58 EDN: ZTNELH
9. McLuhan, M.G. (2019). *Understanding Media: External Human Expansions*. Moscow: Kuchkovo polye. (In Russ.).
10. Yartseva, S.S. (2011). *Columnistics: History of Emergence and Prospects for Development*: [PhD Thesis]. Voronezh: Voronezh State University. (In Russ.). EDN: ZOJGUZ
11. Kalmykov, A.A. (2011). Convergence as The Possibility of Universal Journalism in The Framework of Professional Identity. In: *Optimal Communications (OK): Epistemic Resource of the Academy of Media Industry and the Department of Theory and Practice of Social*

Connectivity of the Russian State University for the Humanities. URL: http://jarki.ru/wpress/2011/04/28/2198/ (accessed: 12.12.2025). (In Russ.).
12. Kuzmina, N.A. (2011). *Modern Media Text*. Omsk: Tatyana. (In Russ.). EDN: QXGBOT
13. Duskayeva, L.R. (2012). The Intensity of Media Text as a Determinant of Speech Behavior in Media. In: *Media Text as a Polyintensive System: Collection of Articles*, L.R. Duskaeva, N.S. Tsvetova (eds.) (pp. 10–16). Saint Petersburg: Saint Petersburg State University. (In Russ.).
14. Jakub, J.O. (2013). The Study of Media Text in Semiotics: Methodological Aspects. *Business. Society. Power*,14, 55–63. (In Russ.). EDN: STSIUB
15. Solganik, G.Ya. (2005). How To Attribute Text and Media Text. *Vestnik Moskovskogo universiteta. Seriya 10. Zhurnalistika*, 2, 7–16. (In Russ.). EDN: JXDHGL
16. Klushina, N.I. (2014). Mediastylistics. *Ecology of Language and Communicative Practice*, 2(3), 69–77. (In Russ.). EDN: SVYYBV
17. Borisova, S.S. (2015). Implementation of Persuasive Strategies in Analytical Genre of Editorial as Resource of Information Dominance. *Modern Linguistic and Methodical-and-Didactic Researches*, 1 (25), 102–114. (In Russ.). EDN: TPUVVX
18. Zavatskaya, V.V. (2020). To the Definition of Media-Text. In: *Linguistics, Linguodidacticss, Llinguoculturology: Current Issues and Prospects for Development: Materials of the IV International Scientific Practical Conference (Minsk, 19–20 March 2020)*, O.G. Prokhorenko, O.V. Dubrovinia, A.I. Golovna et al. (eds.) (pp. 92–95). Minsk: Belarusian State University. (In Russ.). EDN: XWEMHN
19. Chelysheva, I.V. (2008). Methods of Working with Media Text: Philosophical Aspects of the Problem. In: *Mediaphilosophy.,* V. V. Savchuk (Ed.) (pp. 325–329). Saint Petersburg: Philosophical Society. EDN: WAZGAL
20. Bobrovskaya, G.V. (2011). *Cognitive-Educational Potential of Newspaper Discourse*. Volgograd: Peremena. (In Russ.). EDN: SUHLKN
21. Ignatova, I.B., Zubarkina, E.S., & Zemlyansky, A.V. (2022). *New Media: Theory and Practice.* Moscow: Moscow State Pedagogical University. (In Russ.). https://doi.org/10.31862/9785426310957 EDN: TSSGTC
22. Kakchkaeva, A.G., & Shomova, S.A. (2017). *Multimedia Journalism*. Moscow: Higher School of Economics. (In Russ.). https://doi.org/10.17323/978-5-7598-1189-3 EDN: XYMVJJ
23. Dobrosklonskaya, T.G. (2020). *Medialingvistica: Theory, Methods, Directions*. Moscow: Dobrosvet. (In Russ.). https://doi.org/10.31453/kdu.ru.91304.0107 EDN: CQKWES
24. Duskaeva, L.R. (2014). Speech Structure of News Nexts in Networks Editions. *Perm University Herald. Russian and Foreign Philology*, 1(25), 179–183. (In Russ.). EDN: QEJSHA
25. Vartanova, E.L. (2012). The Language of Russian Media as an Indicator of Social Change. In: G.Ya. Solganik (Ed.). *The Language of Media and Politics* (pp. 3–7). Moscow: Moscow State University Publ. (In Russ.).
26. Jenkins, H. (2006). *Convergence Culture: Where Old and New Media Collide*. New-York, London: New York University Press.
27. Amzin, A.A. (2016). New and Experimental Formats. In: S.L. Balmaeva, M.Y. Lukin (eds.) *How New Media Have Changed Journalism: 2012–2016* (pp. 140–151). Moscow: National Research University Higher School of Economics. (In Russ.).
28. Gradyuszko, A.V. (2017). Key Trends in the Development of Foreign Internet Media in the Context of Digitalization of the Media. In: *Journalism-2017: State, Problems and Prospects : Materials of the 19th International Scientific and Practical Conference, Minsk, 9–10 November 2017* (pp. 66–71). Minsk: Belarusian State University. (In Russ.).
29. Lang, K. *Longer TikToks Get More Views – Data Shows Best TikTok Video Buffer*, 17.03.2025. URL: https://buffer.com/resources/longer-tiktoks-get-more-views-data/ (accessed date: 06.10.2025).

Сведения об авторах:
Хромов Сергей Сергеевич, доктор филологических наук, профессор профессор кафедры русского языка и истории литературы факультета издательского дела и журналистики, Московский политехнический университет (107023, Российская Федерация, г. Москва, ул. Большая Семеновская, д. 38); *сфера научных интересов*: поликодовый текст, медиадискурс, семантика, семиотика; *e-mail*: chelovek653@mail.ru
ORCID: 0000-0003-4977-2960, SPIN-код: 021-9409, AuthorID: 468188.
Распопов Артем Андреевич, кандидат филологических наук ассистент кафедры русского языка и истории литературы факультета издательского дела и журналистики, Московский политехнический университет (107023, Российская Федерация, г. Москва, ул. Большая Семёновская, д. 38); *сфера научных интересов*: поликодовый текст, медиадискурс, семантика, семиотика; *e-mail*: raspopov.tema@gmail.com
ORCID: 0009-0002-6077-5839, SPIN-код: 1790-1401, AuthorID: 1336307.

Information about the authors:
Sergey S. Khromov, Dr.Sc. (Philology), Professor, Professor of the Department of Russian Language and Literary History, Moscow Polytechnic University (38 Bol′shaya Semyonovskaya Str., Moscow, Russian Federation; 117198); *Research interests*: polycode text, media discourse, semantics, semiotics; *e-mail*: chelovek653@mail.ru
ORCID: 0000-0003-4977-2960, SPIN-code: 021-9409, AuthorID: 468188.
Artyom A. Raspopov, PhD in Philology, Assistant of the Department of Russian Language and Literary History, Moscow Polytechnic University (38 Bol′shaya Semyonovskaya Str., Moscow, Russian Federation; 107023); *Research interests*: polycode text, media discourse, semantics, semiotics; *e-mail*: raspopov.tema@gmail.com
ORCID: 0009-0002-6077-5839, SPIN-code: 1790-1401, AuthorID: 1336307.

RUDN Journal of Language Studies, Semiotics and Semantics
2313-2299 (Print), ISSN 2411-1236 (Online)
Вестник Российского университета дружбы народов.
Серия: ТЕОРИЯ ЯЗЫКА. СЕМИОТИКА. СЕМАНТИКА

2026 Vol. 17 No. 2 690–711
http://journals.rudn.ru/semiotics-semantics

DOI: 10.22363/2313-2299-2026-17-2-690-711
EDN: MAVNTQ

Research Article / Научная статья

Professional and Popular Science: Mapping Linguistic Overlap and Divergence

Keith Topping[1], Liudmila A. Egorova[2]✉, Natalia V. Nikashina[2], Ekaterina V. Nagornova[2]

[1]University of Dundee, *Dundee, Scotland, UK*
[2] RUDN University, *Moscow, Russian Federation*
✉egorova-la@rudn.ru

Abstract. The authors introduce a novel application of the Linguistic Inquiry and Word Count (LIWC) methodology to quantitatively differentiate the linguistic profiles of professional and popular science articles, establishing them as belonging to distinct functional styles. The results demonstrate that popular science articles are characterized by affective language, future-time orientation, and narrative devices to enhance accessibility and engagement for non-specialists. In contrast, professional articles exhibit terminological density, formal punctuation, and structural rigidity, reflecting their precision-oriented communication for experts. Statistically significant differences between professional and popular science articles show that the popular science functional style is being formed as an independent entity with its own linguistic system blending journalistic engagement with scientific discourse, rather than existing merely as a simplified derivative of professional science style.

Keywords: scientific functional style, science discourse, popular science text, Linguistic Inquiry and Word Count (LIWC)

Authors' contribution: The authors contributed equally to this article and preparation of the article text.

Conflict of interests: The authors stated that there were no conflicts of interest.

Article history: Received: 20.01.2026; Accepted: 10.02.2026.

For citation: Topping, K., Egorova, L.A., Nikashina, N.V., & Nagornova, E.V. (2026). Professional and Popular Science: Mapping Linguistic Overlap and Divergence. *RUDN Journal of Language Studies, Semiotics and Semantics, 17*(2), 690–711. https://doi.org/10.22363/2313-2299-2026-17-2-690-711 EDN MAVNTQ

© Topping K., Egorova L.A., Nikashina N.V., Nagornova E.V., 2026
This work is licensed under a Creative Commons Attribution-NonCommercial 4.0 International License
https://creativecommons.org/licenses/by-nc/4.0/legalcode

Язык академической науки и научно-популярной литературы: сходства и различия

К. Топпинг[1], Л.А. Егорова[2] ✉, Н.В. Никашина[2], Е.В. Нагорнова[2]

[1]Университет Данди, *Данди, Шотландия, Великобритания*
[2] Российский университет дружбы народов, *Москва, Российская Федерация*
✉egorova-la@rudn.ru

Аннотация. Впервые методология ЛИВК (лингвистический анализ и подсчёт слов) применяется для количественного выявления различий между языковыми профилями академических и научно-популярных статей и их отнесения к разным функциональным стилям. Результаты показывают, что научно-популярные статьи отличаются эмоциональной окрашенностью, ориентацией на будущее и использованием нарративных стратегий, что делает их более понятными и интересными для широкой аудитории. В то же время академические статьи демонстрируют высокую терминологическую плотность, формальную пунктуацию и структурную жёсткость, что соответствует установке на точность и адресность в коммуникации со специалистами. Статистически значимые расхождения между двумя типами статей подтверждают, что научно-популярный стиль сложился как самостоятельный лингвистический феномен, синтезирующий элементы научного и медийного дискурсов, а не является упрощенным вариантом академического научного стиля.

Ключевые слова: научный функциональный стиль, научный дискурс, научно-популярный текст, лингвистический анализ и подсчет слов (LIWC)

Вклад авторов: равнозначен на все этапах исследования и подготовки статьи к публикации.

Конфликт интересов: авторы заявили об отсутствии конфликтов интересов.

История статьи: дата поступления: 20.01.2026; дата приема в печать: 10.02.2026.

Для цитирования: *Topping K., Egorova L.A., Nikashina N.V., Nagornova E.V.* Professional and Popular Science: Mapping Linguistic Overlap and Divergence // Вестник Российского университета дружбы народов. Серия: Теория языка. Семиотика. Семантика. 2026. Т. 17. № 2. С. 690–711. https://doi.org/10.22363/2313-2299-2026-17-2-690-711 EDN MAVNTQ

Introduction

Globalization of science leads to the changes in the organization of science communication, defined by Burns et al. [1. P. 183]. as “the use of appropriate skills, media, activities and dialogue to produce one or more of the following personal responses to science (the AEIOU vowel analogy): Awareness, Enjoyment, Interest, Opinion-forming, and Understanding.” Thus, the main trends of professional communication development noted by Gunnarsson [2] — multi-linguality, multiculturality, flexibility, and diversity — become the main features of science communication as well. The research community became aware that scientific progress was impossible without the global involvement of many researchers in the cycle of creating, exchanging, processing and evaluating large quantities of information. These global processes led to the development of new ways of disseminating knowledge, such as digital libraries, open access journals and social networking sites [3]. As science discourse became more international, less individualized, and more uniform and structured, it formed several subsystems

within it, fulfilling different functions besides research development: for example, informational, popularizing, educational, political, advertising, and social [4; 5]. On the periphery of science discourse, new subtypes were formed, e.g., popular, didactic, technical, and so forth, distinguished mainly by the audience they were aimed at and the functions they performed.

Online communication changes the processing of information. Publishing new articles and research data becomes easier and faster. Researchers can connect their new works with their previous ones, include references to preceding texts and data, and spread new data among wider audiences. New media allow showing and hiding essential and additional text components, to single out structural elements fulfilling different functions. Means for supporting the communication between members of the scientific community also acquire new features [7]. New tools appear that allow making comments, connecting new publications with similar ones, and uniting all publications on the same topic into clusters, ranging the publications according to number of citations, readings, and comments, all indicating the interest of the scientific community [8–11]. What is more, multimodal texts using various media prevail over purely verbal ones, as their non-verbal elements attract the attention of readers [11] and possibly increase the rating of the publication.

Therefore, the form of texts changes accordingly. New components fulfil specific functions, and the inner structure of the text becomes more oriented to dialogue and promotion by automatic tools. The discourse loses its homogeneous character, and its marginal subtypes separate and develop into new types with their various audiences, functions, genres and language peculiarities. Investigating emerging types of discourse can be carried out in several ways, e.g., via language means, genres, pragmatics and polymodality, depending on the aims of the research [12; 13]. According to Chernyavskaya [14], the key issue arising when studying popular science discourse is the role of the popularization type of narration in the system of functional styles, as popular science texts are quite flexible and combine features of different styles (academic, belles-lettres, and journalistic).

On the one hand, language is becoming more homogeneous and predictable, but on the other, new text elements appear to transform the text into cyber text with its complex nature. Text analysis is becoming impossible without a multidimensional approach, taking into account the characteristics of the members involved in the communication, and the aims and media used. What is more, large amounts of data require the application of computerized processing tools, that could help in revealing the text functions, define the genre of the text, and show stylistic errors to academic language learners. Applying such tools may identify the textual features relevant for distinguishing different genres and discourses.

Literature Review

The variety of language associated with a certain social setting, time, and participants is the focus of attention of modern linguistic research. Different terms are used depending on the approach to the study. For example, Biber and Conrad

referred to the notion of a register as the combination of "an analysis of linguistic characteristics that are common in a text variety with analysis of the situation of use of the variety" [15. P. 2]. They noted that "some linguistic features are common in a register because they are functionally adapted to the communicative purposes and situational contexts of texts from that register". The linguistic features distinguishing features of different registers are universal and found in different languages, e.g., Haider and Palmer [16] showed that advertising, novels, academic texts, and reports in German possess a set of linguistic features that were statistically different for genres belonging to different registers.

The term 'functional style', specific for the Czech and Russian schools of Functional Stylistics, denotes historically formed and socially organized speech variety that possesses the quality of systematicity, characterized by specific principles of selection and combination of language. This variety is correlated with a particular sphere of communication and form of cognition[1]. Quite often distinguishing the terms 'style', 'functional style' and 'register' arouses confusion [17; 18]. The term 'Functional Style' is used in this paper to emphasise the dependence of linguistic features on the functions of the chosen language variety and its connection with spheres of communication. A detailed analysis of functional styles in English was given by Galperin [19] and developed in subsequent research connected with various languages (see [20–24]), which prove a universal character of functional styles features.

The scientific functional style appears as a result of science communication and is characterized by objectivity, logical evidence and consistency, high level of abstractness, preciseness, clarity, formality, and terminological uniformity [20; 21; 25], thus opposing the belles-lettres style characterized by individuality, heterogeneity, particular selection of words revealing new meanings in the context helping to create imagery, interactions of meanings, variability, and uniqueness [19; 22].

A field approach to the notion of functional style [26] allows regarding a functional style not as a closed entity, but as a phenomenon with a nucleus and periphery. Genres possessing all the specifics of a particular functional style belong to its nucleus, while those which display new peculiarities due to new extra-linguistic factors are attributed to the periphery. Popular science genres first appeared on the periphery of the scientific functional style. Initially, the popular science functional style was regarded as a substyle of the scientific functional style [20; 25; 27], then as a separate functional style with its own aim, purposes, participants, specific interaction between the author and the recipient of message, and content of the message [28].

The last 40 years witnessed an unprecedented growth in mass communications based on new technologies, and social and political changes. The following factors

[1] Kozhina, M.N. (2003). Functional style. In: M.N. Kozhina (Ed.), *Stylistic Encyclopedic Dictionary of the Russian language*. Moscow: Flinta, Nauka. (In Russ.).

enhancing the transformation of professional and popular science functional styles can be singled out [1; 7; 10; 13; 29]:

— ideological crises in mass consciousness due to globalization on all levels of social life;

— global character of topics interesting for the audience at a certain period of time;

— a considerable increase in the number of publications due to electronic versions;

— the necessity to address people of different ages, nationalities, educational and social background to achieve the goals of changing the mass picture of the world;

— a new characteristic of author/ reader interaction — a dialogue, in which the readers have become active participants in communication, their activity promoting the rating of the publication together with the topic and the quality of the text.

These extra-linguistic factors caused changes in the structure of professional and popular science texts. They are no longer linear, though according to Clyne [30], the linear form is specific to low-contextual Anglo culture dominating natural science communication [31]. All elements of discourse must be verbalized and unambiguous, and formal aspects play a dominant role. There is a tendency for texts in natural sciences to be more influenced by English discourse patterns, i.e., to be more linear and more symmetrical, while in more content-oriented cultures "people are freer to 'digress', and 'digressions' (excursions) are utilized to broaden the topic of the text, narrowed by the Anglo confines of 'relevance'. For instance, historical, ideological and polemic academic issues can be more readily incorporated into a text not specifically dealing with these" [30. P. 73].

Now science texts are represented as a system of nodes and links created in virtual cyberspace. According to Egorova: The hypertext information model, proposed back in the 80s and based on the hypothesis that the processing and generation of ideas by the human brain is associative, is increasingly recognized as a structure for the effective representation and transfer of knowledge. A hypertext system containing a network of nodes (fragments, modules, frames) and the associated links create a multidimensional information space adequate to the deep structure of the ideas processing by the human brain [32. P. 82].

A hypertext goes through the stage of a hypermedia text to become a cybertext. As Aarseth puts it, the concept of a cybertext focuses on the processing of the text, the medium and 'the user' of the text. "During the cybertextual process, the user will have effectuated a semiotic sequence, and this selective movement is a work of physical construction that the various concepts of "reading" do not account for" [33. P. 1]. The recipient is active, his/her interaction with the text may be successful or not, and s/he is always at risk like a gambler. The role of the medium is changing as well. New forms of electronic communication change the representation of different structural elements according to their functions. Electronic means of

publishing change the visibility of the whole structure; the recipient may include various nodes in his/her individual path of the text processing. Perception is limited by the size of the screen, on the one hand, but is unlimited in terms of penetrating into distant nodes following the links from the text elements. The text becomes creolized, i.e., verbal and non-verbal elements are united and influence the results of the text evaluation [7; 34; 35].

Popular science texts tend to be more comprehensive, i.e., they include elements bearing social, advertising, expressive, or emotional character reflecting the functions they fulfil. They verbalize the values set by the addresser [36]. Therefore, their non-linear information structure requires an integral approach for the analysis.

Several attempts were made to design a model of the popular science text integrating segments fulfilling different functions. According to Petrushka, the popular science content is "a multicomponent object in its purpose, with oriental and auditorial typological features which organically combine the functions of the scientific, educational, popular information source and, as far as possible, among all other kinds and types of content mediums best represents the processes of convergence" [37. P. 4].

The model based on four sectors (cognitive, linguistic, social and cultural) suggested in Khomutova & Petrov [38] reflects the production, development, perception and functioning of the text on different levels and allows focusing attention on a particular aspect, without losing the integrity of a text as a complex object. In the cognitive sector, the text is analysed as a fragment of knowledge of a certain subject area. And in the social sector, as a fragment of social space. The cultural sector reveals cultural nation-specific values expressed in a particular text. The linguistic sector helps identify linguistic categories and language means conveying knowledge, cultural values and social actions. Communicative activity is a core that unites all four sectors into a single whole, in which all aspects are interdependent and interrelated [39]. So, professional/popular science cybertexts comprise components with different functions related to different sectors, obliging the reader to follow a specific path in processing them.

In suggesting that science and popular science texts might show significant differences in several sectors, it was necessary to find a tool that could help process texts, analysing parameters connected with various fields. Linguistic Inquiry Word Count (LIWC) originally was used in psychometric research and proved to be a valuable instrument due to the following features. First of all, the procedure for dictionary words selection is transparent and demonstrated reliable results in estimating different variables of texts [40–43]. The default systemic LIWC2015 Dictionary consists of approximately 6.400 words, word stems, and emoticons that tap basic emotional and cognitive dimensions often studied in social, health, and personality psychology: numbers, punctuation, short phrases, "netspeak" language from Roget's Thesaurus, standard English dictionaries and brain-storming sessions among judges. All items have undergone six stages of analysis and further selection:

Judge Rating Phase, Base Rate Analyses, Candidate Word List Generation, Psychometric Evaluation and Refinement Phase [41. P. 5–7] and have been distributed into different categories. The tool allows the analysis of punctuation, grammar, lexical, social, psychological, and emotive parameters (85 variables in total), which have demonstrated their reliability in various aspects of text analysis [44–46].

Summary variables were developed to investigate linguistic, social and psychological processes. They were: "Analytical thinking — a high number reflects formal, logical, and hierarchical thinking; lower numbers reflect more informal, personal, here and now and narrative thinking. Clout — a high number suggests that the author is speaking from the perspective of high expertise and is confident; low Clout numbers suggest a more tentative, humble, even anxious style. Authentic — higher numbers are associated with more honest, personal, and disclosing text; lower numbers suggest a more guarded, distanced form of discourse. Emotional tone — a high number is associated with a more positive, upbeat style; a low number reveals greater anxiety, sadness, or hostility. A number around 50 suggests either a lack of emotionality or different levels of ambivalence." [46. P. 21–22]. The summary variables facilitate textual analysis of different sectors.

Methodology

The genre of articles was chosen for the research, as articles represent the most frequent type of professional science and popular science texts, demonstrating the core features of each style, covering recent scientific events addressed to wide circles of scientists and general public interested in the subject.

First, 30 open access popular science articles written by science journalists and published in Scientific American and The Scientist were selected. Those magazines were selected for their interdisciplinary character, the quality of the presented research, the address wide circles of scientists and the principles they pursued. For example, Scientific American publishes work by journalists, scientists, scholars, policy makers and people with lived experience of scientific or social issues"[2]. The selected articles met the following criteria: 1) uniformity of the genre (feature article); 2) field of study connected with life sciences (medicine, health, biology, ecology); 3) topic interesting for broad audience.

Each popular science article was based on research, but we also chose open-access evidence-based professional science articles based on similar themes. This restriction on the choice of articles aimed at retaining the field of study and thematic alignment between the professional science and popular science articles. The factors in the selection of articles were as follows: 1) uniformity of the genre (popular or professional science article); 2) coverage of the topic discussed in both

[2] Scientific American. (2023, January 15). *About*. Springer Nature. URL: https://www.scientificamerican.com/page/about-scientific-american/ (accessed: 15.01.2023).

articles, although in the popular science article in the most general way; 3) length of the article (professional articles should be longer than popular articles).

Five popular science articles out of 30 did not have professional science articles in their reference list which met the above-mentioned criterion of length. So, we had to select two shorter professional science articles for each to keep the length proportion between professional and popular science articles. This procedure provided 30 popular science and 35 professional science articles, matched in the same fields of study.

The number of articles and number of words collected in two corpuses — popular science and professional science articles — allowed undertaking statistical analysis of the data and at the same time considering their compositional and thematic features individually (see Table 1).

Table 1. **Characteristics of Popular Science and Professional Science Corpuses**

	Popular Science Corpus	**Professional Science Corpus**	**Total**
Number of articles	30	35	65
Number of words	86.003	222.559	308.562
Average number of words per one article	2.867	6.359	
Years of Publication	2018–2020	2011–2019	

Source: compiled by Keith Topping, Liudmila A. Egorova, Natalia V. Nikashina, & Ekaterina V. Nagornova.

Compositional peculiarities of articles are usually studied in relation to their connection to different discourses [47], functions of elements [5; 37; 48] and perception [49]. We aimed to consider common structural elements of articles of each corpus to define their functions. On the basis of a conceptual model of medium of popular science content suggested by Petrushka [37], the following functions were regarded as essential for popular science articles: 1) Popularization of research; 2) Creating a network on the topic; 3) Navigating the article; 4) Informing about publication details.

Professional science article components fulfilled a wider set of functions, most of which were related to proving the validity and reliability of the research data and forming a cluster of researchers [48; 50–55]: 1) Informing about the area, background, methodology, results of the research; 2) Giving additional information about the data of the research; 3) Creating the science net on the topic; 4) Informing about the participants of the research; 5) Navigating the article; 6) Informing about publication details; 7) Promoting the author/journal/institution/foundation.

The next stage was connected with the textual compositional analysis of selected articles. The task was to single out the components of multi-modal texts fulfilling different functions common for all analysed popular or professional science articles, to identify the components possessing stylistic distinctive features and suitable for a comparative analysis by means of LIWC2015, Version 1.6.0. The hypothesis was that the principal functions that distinguished the professional and

popular science discourses would be somewhat different, systemically influencing textual features, i.e., in informing about the area, background, methodology, results of the research and popularization of research respectively. Thereafter, statistical analysis might reveal stylistic features specific for each corpus, reflecting the differential formation of popular science and professional science functional styles.

Results

For popular science articles published in online editions, the common elements were: 1) Route to the article from Home Page of the magazine; 2) Headline (with the Hyperlink to the main body of the article); 3) Summary; 4) Author; 5) Date of publication; 6) Information about the author with visuals; 7) Copyright information; 8) Icons of Social Networks; 9) Main Body; 10) Visuals with captions; 11) Hyperlinks to different parts of the article; 12) References (or hyperlinks in the main body) to research data; 13) Keywords (optional); 14) List of references; 15) References to related popular science articles.

The structure of professional science articles was more complicated, determined by the publication requirements of the specific journal though somewhat similar across different journals. It comprised: 1) Route to the article from Home Page of the journal; 2) Headline (with the Hyperlink to the main body of the article); 3) Author(s); 4) Correspondence author; 5) Affiliations of the authors; 6) Information about the author(s) (with hyperlinks to their profiles); 7) Access information; 8) Copyright information; 9) DOI (Digital Object Identifier); 10) Abstract /In Brief /Highlights; 11) Publication history (Dates of submission, reviewing, publication); 12) Keywords; 13) Icons of article tools (save, share, print, figures, request permissions, citation tools etc.); 14) Icons of Social Networks; 15) Main Body (Introduction, Methods, Procedures Results, Discussion, Conclusion); 16) Visuals with captions; 17) Hyperlinks to different parts of the article; 18) References (or hyperlinks) to research data, figures, tables, previous research; 19) Advertisement; 20) Footnotes/Notes; 21) Supplementary material; 22) Acknowledgment; 23) Author Contributions; 24) Declaration of interest; 25) List of References; 26) Appendices.

If specific distinctive features appeared in the latter components as they were specific for each article and fulfilled style-determining functions, we retained only the components connected with the functions ‘popularization of research’ and ‘informing about the area, background, methodology, results of the research’ for the next stage of analysis. Two corpuses of abridged professional science articles and popular science articles were compiled. Each article contained the following components: 1) Headline (with the Hyperlink to the main body of the article); 2) Summary; 3) Author; 4) Main Body; 5) Visuals with captions.

Tables 2 and 3 (below) show the functions of the major components of the popular and professional science articles under analysis. Textual analysis showed that some components contained standard textual or visual elements required by the publisher, and built a frame for the article (Route to the article from Home Page of

the magazine, Headline (with the hyperlink to the main body of the article), Date of publication, Publication history (dates of submission, reviewing, publication); Affiliations of the authors, Copyright information, References (or hyperlinks in the main body) to research data, References, Digital Object Identifier, Access information, Acknowledgment, Author contributions, Declaration of interest, Icons). Others included information peculiar for each article (Author(s), Headline, Summary, Main Body (Introduction, Methods, Procedures Results, Discussion, Conclusion, Visuals with captions).

Table 2. **The Basic Functions of Components of Popular Science Articles**

Component	Function			
	Popularisation of research	**Creating a network on the topic**	**Navigating the article**	**Informing about publication details**
Route to the article from Home Page of the magazine	–	–	+	+
Headline (with the Hyperlink to the main body of the article)	+	+	+	+
Summary	+	-	–	+
Author	+	+	–	+
Date of publication	–	–	–	+
Information about the author with visuals	–	+	–	+
Copyright information	–	–	–	+
Icons of Social Networks	–	+	–	–
Main Body	+	–	–	–
Visuals with captions	+	–	–	–
Hyperlinks to different parts of the article	–	–	+	–
References (or hyperlinks in the main body) to research data	–	+	–	–
Keywords (optional)	–	+	–	–
List of references	–	+	–	–
References to related popular science articles	–	+	–	–

+ = function present; – = function absent

Source: compiled by Keith Topping, Liudmila A. Egorova, Natalia V. Nikashina, & Ekaterina V. Nagornova.

Each abridged article in the corpuses was analysed as independent. All available variables (92 in total) were calculated. The Shapiro-Wilk test was used to check each variable was normally distributed. Then the two corpuses were compared using Levene's test (to check that variances were equal for both samples). Where equal variances and normal distribution were confirmed, an independent t-test was applied. Where equal variances were not confirmed but the distribution

was normal, Welch's t-test was applied. Where the distribution was not normal, the non-parametric Mann-Whitney test was used.

Table 3. **The Basic Functions of Components of Professional Science Articles**

Component	Function						
	Informing about the area, background, methodology, results of the research	**Giving additional information about the data of the research**	**Creating the science net on the topic**	**Informing about the participants of the research**	**Navi-gating the article**	**Informing about publication details**	**Promoting the Authors / Journal/ Institution / Foundation**
1	**2**	**3**	**4**	**5**	**6**	**7**	**8**
Route to the article from Home Page of the journal	–	–	–	–	–	+	–
Headline (with the Hyperlink to the main body of the article)	+	–	–	–	+	–	–
Author(s)	+	–	–	+	–	–	+
Correspondence author	–	–	–	–	–	+	-
Affiliations of the authors	–	–	–	–	–	+	+
Information about the author(s) (with hyperlinks to their profiles)	–	–	–	–	–	–	+
Access information	–	–	–	–	–	+	–
Copyright information	–	–	–	–	–	+	–
DOI (Digital Object Identifier)	–	–	–	–	–	+	–
Abstract /In Brief /Highlights	+	–	–	–	–	-	–
Publication history (Dates of submission, reviewing, publication)	–	–	–	–	–	+	–
Keywords	+	–	+	–	–	–	–
Icons of article tools (save, share, print, figures, request permissions, citation tools etc.)	–	–	+	–	–	–	–
Icons of Social Networks	–	–	+	–	–	–	–
Main Body (Introduction, Methods, Procedures Results, Discussion, Conclusion)	+	–	-	–	–	–	–
Visuals with captions	+	–	–	–	–	–	–
Hyperlinks to different parts of the article	-	–	–	–	+	–	–
References (or hyperlinks in the main body) to research data, figures, tables, previous research	+	+	+	–	–	–	–
Advertisement	–	–	–	–	–	–	+

Ending of the Table 3

1	2	3	4	5	6	7	8
Footnotes/Notes	–	+	+	–	–	–	–
Supplementary material	–	+	+	–	–	–	–
Acknowledgment	–	–	–	+	–	–	–
Author Contributions	–	–	–	+	–	–	–
Declaration of interest	–	–	–	+	–	–	–
List of References	–	–	+	–	–	–	–
Appendices	–	+	–	–	–	–	–

+ = function present; – = function absent

Source: compiled by Keith Topping, Liudmila A. Egorova, Natalia V. Nikashina, & Ekaterina V. Nagornova.

The means and medians for each variable are presented in Table 4 (below), which lists statistically significant different variables in the two corpuses. A further 27 variables were not significantly different between popular science and professional science. For each group, the total number of variables is also shown, which indicates how many differences in variables were not significant.

Table 4. **Significant Variables for Popular and Professional Science Corpuses**

Category of Variable	Popular science corpus (mean/median)	Science corpus (mean/median)
Summary Language Variables (4)		
Analytical thinking	94,39/95,23	97,07/97,34
Clout	59,88/59,90	54,90/54,86
Authentic	28,04/26,33	17,75/13,31
Language Metrics (3)		
Words/sentence	26,86/27,03	23,48/24,03
Dictionary words	73,74/74,37	58,55/57,45
Linguistic Dimensions (15)		
Total function words	42.59/42.18	33.29/33.19
Total pronouns	6.51/6.49	3.33/3.08
Personal pronouns	2.07/1.98	1.09/0.93
2nd person	0.20/0.17	0.03/0.01
3rd pers singular	0.53/0.44	0.17/0.00
3rd pers plural	0.81/0.72	0.24/0.18
Impersonal pronouns	4.44/4.40	2.23/2.16
Articles	8.19/7.85	6.76/6.47
Prepositions	15.25/15.07	13.62/13.45
Auxiliary verbs	5.67/5.32	4.15/4.11
Common Adverbs	3.27/3.31	1.58/1.45
Conjunctions	5.47/5.52	4.47/4.36
Negations	0.67/0.70	0.47/0.44
Other Grammar (6)		
Common verbs	9.93/9.78	6.37/5.93
Common adjectives	4.94/4.85	3.97/3.76
Comparisons	2.97/3.07	2.17/1.97
Interrogatives	1.26/1.22	0.52/0.47
Numbers	2.11/2.01	7.73/6.79

Continuation of the Table 4

Category of Variable	Popular science corpus (mean/median)	Science corpus (mean/median)
Quantifiers	2.28/2.37	1.59/1.47
Affective processes (6)		
Affective processes	2.99/2.84	2.12/1.90
Positive emotion	1.56/1.62	1.18/1.01
Negative emotion	1.38/1.05	0.92/0.86
Anger	0,21/0,17	0,09/0,06
Social processes (5)		
Social processes	5.47/5.53	3.06/2.68
Friends	0.23/0.20	0.12/0.03
Male references	0.40/0.31	0.13/0.03
Cognitive processes (7)		
Cognitive processes	11.02/10.69	8.16/8.61
Discrepancy	0.92/0.90	0.38/0.28
Tentative	2.47/2.26	1.45/1.32
Differentiation	2.74/2.81	1.77/1.64
Perceptual processes (4)		
Perceptual processes	2.10/2.01	1.31/1.06
Hear	0.65/0.58	0.09/0.03
Feel	0.32/0.23	0.25/0.14
Biological processes (5)		
Biological processes	4.10/3.96	2.54/2.40
Body	1.42/0.73	0.57/0.38
Drives (6)		
Drives	5.34/5.01	4.42/4.11
Affiliation	1.29/1.25	1.11/0.93
Achievement	1.43/1.41	1.01/0.85
Reward	0.75/0.72	0.56/0.48
Risk	0.58/0.51	0.40/0.37
Time orientations (3)		
Present focus	6.15/5.95	3.34/2.93
Future focus	0.90/0.75	0.53/0.46
Relativity (4)		
Relativity	13.08/12.67	11.26/10.60
Motion	1.73/1.63	1.25/1.03
Space	7.58/7.56	6.78/6.78
Time	3.87/3.67	3.20/2.62
Personal concerns (6)		
Work	3.86/3.75	2.63/2.39
Informal language (6)		
Informal language	0.16/0.14	0.75/0.65
Netspeak	0.05/0.00	0.62/0.56
Nonfluencies	0.06/0.06	0.14/0.10
Punctuation (12)		
Punctuation marks	15.74/15.53	23.58/22.83
Period	3.94/3.82	5.51/5.03
Colon	0.22/0.18	0.37/0.31
Semi-Colon	0.05/0.05	0.45/0.38

Ending of the Table 4

Category of Variable	Popular science corpus (mean/median)	Science corpus (mean/median)
Question mark	0.07/0.04	0.06/0.00
Dash	1.32/1.16	2.95/3.00
Quotation marks	1.59/1.44	0.34/0.16
Apostrophe	1.28/1.28	0.24/0.14
Parentheses	0.63/0.56	5.17/5.17
Other punctuation marks	0.52/0.23	2.57/2.07

Source: compiled by Keith Topping, Liudmila A. Egorova, Natalia V. Nikashina, & Ekaterina V. Nagornova.

All groups contained variables distinguishing Professional Science and Popular Science articles. Ranging the groups according to the percentage of these variables, we can single out 'Biological processes' and 'Personal concerns' as having less than 50% significant variables. Table 5 shows different groups indicating the percentage of significant variables for each.

Table 5. **Statistically Significantly Different Variables between Popular and Professional Science**

Name of group of variables	Total number of variables in the group	Number of variables showing statistically significant difference	Percentage of variables showing statistically significant difference, %
Relativity	4	4	100
Other Grammar	6	6	100
Linguistic Dimensions	15	13	87
Punctuation	12	10	83
Drives	6	5	83
Perceptual processes	4	3	75
Summary Language Variables	4	3	75
Language Metrics	3	2	67
Time orientations	3	2	67
Affective processes	6	4	67
Social processes	5	3	60
Cognitive processes	7	4	57
Informal language	6	3	50
Biological processes	5	2	40
Personal concerns	6	1	17
Total	92	65	68

Source: compiled by Keith Topping, Liudmila A. Egorova, Natalia V. Nikashina, & Ekaterina V. Nagornova.

On the whole, articles in the Popular Science Corpus are characterized by longer sentences, lower analytical thinking parameters, higher parameters of authenticity and clout, higher numbers of pronouns, articles, auxiliary verbs, conjunctions and negations. Time orientation words also distinguish two corpuses. Parameters connected with focus on the present/future are twice as high in popular

science articles. Their authors use more words out of the systemic dictionary (see the description in [41. P 5—7]), which belong to different categories and are validated to reflect various psychological/social/cognitive/biological processes and states, as well as words connected with the future. Professional science articles contain more informal words, Internet abbreviations and fillers though they are very rare. As for punctuation marks, the number of these is higher in professional science articles, and colons, semi-colons, parentheses, and dashes are more typical for professional science articles, while quotation marks and apostrophes are more typical for popular science texts.

Discussion

Summary

The obtained results show that professional science and popular science articles, being the core genres of professional science and popular science functional styles, differ in many and indeed most variables reflecting Language metrics, Linguistic dimensions, Perceptual processes, and Punctuation. The articles' structures reflect their multi-level communicative character — elements of all sectors (cognitive, linguistic, social and cultural, according to Khomutova & Petrov [38] are present and sustain the development of science communication involving audiences of different background. Interdiscursivity and multimodality enhance the properties of the text connected with putting the research into the general context of science communication and facilitating the perception of information. Distinctive features of professional science articles structure (such as Correspondence author; Affiliations of the authors; Information about the author(s) (with hyperlinks to their profiles); Access information; Copyright information; DOI (Digital Object Identifier); Highlights; Publication history (Dates of submission, reviewing, publication); Icons of article tools (save, share, print, figures, request permissions, citation tools etc.); Advertisement; Footnotes/Notes; Supplementary material; Acknowledgment; Author Contributions; Declaration of interest) confirm the validity of results.

Limitations and Strengths

This study chose to focus on science articles rather than other kinds of written science communication. It also chose a limited number of articles, and found difficulty in sourcing the requisite amount in one category. By no means all variables showed significant differences between the two corpuses, but this might have been expected. The LIWC methodology gives a means of analysing texts which might be useful for multiple purposes.

Relationship to Previous Literature

The articles are also built into the net of similar research creating a net of publications, which supports the trends in the development of digital science

communication (see [3; 9; 10]. Analysis of abridged versions of articles showed that such structural elements, as 'Headline (with the Hyperlink to the main body of the article)'; 'Summary'; 'Author'; 'Main Body'; 'Visuals with captions' distinguished professional and popular science articles, irrespective of the theme of research.

A popular science article is supposed to be persuasive and emotional, oriented on social processes, and focused on the present. Its addressee is a non-specialist, so the text should be easy to understand, interesting for the reader, emotional and coherent. The language used reflects this communicative intention. The LIWC analysis allowed defining the specific percentage of dictionary (non-terminological) words, time orientation and functional words, and different punctuation marks for popular science articles. Variables associated with social and psychological processes revealed the inner characteristics of popular science texts as more affective, evaluative (positive or negative), and tentative in comparison with science texts.

As for Linguistic dimensions, the corpuses differed in the proportion of pronouns, articles, prepositions, auxiliary verbs, conjunctions and negations they contained. Functional words played a significant role among interactive meta-discourse markers that organized, signposted and facilitated processing the text in general. Recent studies proved their significance for science articles (e.g., [56]). Psychological studies (e.g., [40]) proved their correlation with the interest in objects and things (indefinite articles), education, concern with precision (prepositions), and inhibition (negations). Popular science articles demonstrated a more developed system of connectors and substitutive words rather typical for journalistic and newspaper texts in contrast with professional science texts, with their uniformity, repetitions of nouns, rigid inner structure, and precise information presentation.

Popular science is addressed to non-specialists, so the texts are less terminologically charged, they contain more paraphrases, explicatory constructions, and comparisons. Quantifiers are preferred to numerals, and comparisons and evaluations to neutral factual presentation of data. That accounts for longer sentences, and a larger number of LIWC systemic dictionary words (words of high frequency describing processes) in popular science articles in contrast to professional science articles full of terms which are usually of lower frequency and do not constitute part of the LIWC dictionary.

Time orientation words also distinguished the two corpuses. Parameters connected with a focus on the future were almost twice as high in popular science articles. Popular science articles tended to render scientific information as the content of a story developing along a timeline, so they contained more references to the past, present, and future in general. However, the main objective of a popular science discourse was to attract the audience to urgent scientific news and events, which were usually associated with the changes in the future of mankind. Therefore, the number of verbs in future tenses and adverbial modifiers referring to the future was higher in popular science articles.

Professional science articles contain more punctuation marks, especially colons and semi-colons. However, popular science articles are characterized by a larger number of quotation marks and apostrophes. The usage of graphical means to structure the text reveals their significance in making scientific text coherent, with a more explicit hierarchy of principal and minor non-essential components. Popular science articles also use other means of cohesion, and the results are close to expressive writing [41].

Interpretation

If we compare the results of popular science and professional science articles using LIWC analysis based on different types of texts (novels, expressive writing, natural speech, newspaper articles, blogs, tweeter) [41], we see that science communication texts represent a peculiar type of writing with specific parameters. Popular science articles (Pop) are closer to the New York Times corpus than professional science articles (Pro) in variables reflecting the pragmatic aspect of communication, for example, total number of functional words (NY Times 42.39 — Pop 42.59 — Pro 33.29), common verbs (NY Times 10.23 — Pop 9.93 — Pro 6.37), affective processes (NY Times 3.82 — Pop2.99 — Pro 2.12), social processes (NY Times 7.62 — Pop 5.47 — Pro 3.07), and time orientation future focus (NY Times 0.80 — Pop 0.90 — Pro 0.53). These features reflect the specifics of journalistic texts addressed at ordinary non-specialist people. These facts are presented from the point of view of their importance for the everyday life of the readers, and benefits they may derive from them. A certain information redundancy is created to facilitate the perception of the journalistic text and maintain the interest of the reader. However, the data obtained show significance differences between them and popular science articles in terms of emotional tone ((NY Times 43.61 — Pop 32.47), and past focus (NY Times 4.09 — Pop 2.84). The LIWC analysis shows popular science texts as a separate entity, with specific language features reflecting the stage of developing a popular science functional style on the periphery of science and journalistic styles.

So, LIWC analysis can be applied not only to individual corpuses to measure textual specifics conditioned by individual or situational peculiarities, but also to distinguish corpuses of texts belonging to different functional styles. LIWC dimensions cover important aspects of textual features significant for the differentiation of these styles. The differences between science and popular science articles are relevant in terms of information structure, degree of formality and evaluation, means of cohesion, and time orientation. The LIWC analysis reveals the language features that reflect pragmatic components of the text meaning, such as the relations between members of communication, their characteristics, their attitude to the subject of communication, and their involvement in the process of communication.

The fact that most LIWC variables show statistically significant differences between science and popular science articles demonstrates their complex character,

and proves the formation of a popular science functional style as an independent entity with its own system of language means, selection and combination. The model of science/popular science cybertext helped to distinguish the components of electronic hypertexts fulfilling different functions and compare the compositions of science and popular science articles as core genres of contemporary science and popular science discourses.

Practical Implications & Future Research

Using LIWC can help attribute a text as a scientific or popular science text, which could be important in big data analysis. Another useful application for educational purposes would be to evaluate students' essays as belonging or not to a professional science style, and to show the students the features that are to be changed. At the present stage the following procedure might be suggested to assess the stylistic features of a science article: 1) Abridge the text, so that it includes Headline, Summary, Author, Main Body, Visuals with captions; 2) Using LIWC variables, check the number of words in a sentence, number of pronouns, articles, auxiliary verbs, conjunctions, negations and punctuation marks. Attribute the text to professional or popular science style on the basis of the results obtained.

This procedure could also be used to assess ease of comprehension, which is supposed to be greater in popular science articles. However, further research based on a bigger corpus of professional and popular science texts would be necessary to determine the parameters influencing the perception of different elements. Also, revealing genre specifics of articles in contrast with reviews/monographs/news items/headlines/abstracts, as well as comparing the results for different languages, might be beneficial both for educational purposes and discourse studies.

Conclusions

The development of mass communication led to the differentiation of a system of functional styles. Substyles are transformed into separate entities with their own peculiarities on all language levels, so the texts acquire specific features in terms of cognitive, linguistic, social and cultural aspects, or 'sectors'. This research examined language features which were statistically significant to distinguish professional and popular science articles as core genres of professional and popular science functional styles, thereby adding weight to this tendency.

The numerical method of analysis facilitated identification of the features statistically significant for distinguishing popular and professional science styles and defining their role in text formation. Word counting methods were reliable in identifying distinctive stylistic features, as they singled out different groups of words (functional words, time orientation words, informal words, words connected with psychological/biological/affective/cognitive processes), punctuation marks, number of words in a sentence (which reflected a sentence structure), the level of formality and means of cohesion typical for the chosen material. It has been shown

that language features that appeared due to the functions of textual elements, information structure, and audience characteristics formed a stable system of language means specific for popular science texts in the field of life sciences.

The findings may contribute to educational practices related to teaching English for Academic Purposes, as they showed the importance of identifying structural elements, usage of functional elements and punctuation marks to build up a coherent science text, which could properly fulfil its functions. Using computing tools may be beneficial for the preliminary assessment of academic texts in terms of their consistency with stylistic requirements.

References / Список литературы

1. Burns, T.W., O'Connor, D.J., & Stocklmayer, S.M. (2003). Science Communication: A Contemporary Definition. *Public Understanding of Science*, *12*(2), 183–202. https://doi.org/10.1177/09636625030122004 EDN: JTROXR
2. Cope, B., & Phillips, A. (eds.). (2014). *The Future of the Academic Journal*. Oxford: Chandos Publishing.
3. Malyuga, E.N. (Ed.). (2020). *Functional Approach to Professional Discourse Exploration in Linguistics*. Singapore: Springer Nature Singapore. https://doi.org/10.1007/978-981-32-9103-4
4. Boyd, D. (2010). Social Network Sites as Networked Publics: Affordances, Dynamics, and Implications. In: Z. Papacharissi (Ed.), *A Networked Self* (pp. 47–66). New York: Routledge.
5. Jewitt, C. (Ed.). (2011). *The Routledge Handbook of Multimodal Analysis*. New York: Routledge.
6. Spicer, S. (2014). Exploring Video Abstracts in Science Journals: An Overview and Case Study. *Journal of Librarianship and Scholarly Communication*, *2*(2), eP1110. https://doi.org/10.7710/2162-3309.1110
7. Hyland, K. (2015). *Academic Publishing: Issues and Challenges in the Construction of Knowledge*. Oxford: Oxford University Press.
8. Kress, G., & Van Leeuwen, T. (2020). *Reading Images: The Grammar of Visual Design*. London: Routledge.
9. Biber, D., & Conrad, S. (2019). *Register, Genre, and Style*. Cambridge: Cambridge University Press.
10. Haider, T., & Palmer, A. (2017). Modeling Communicative Purpose with Functional Style: Corpus and Features for German Genre and Register Analysis. In: *Proceedings of the Workshop on Stylistic Variation* (pp. 74–84). Copenhagen: Association for Computational Linguistics. https://doi.org/10.18653/v1/W17-4910
11. Kozhina, M.N. (2003). Functional Style. In: *Stylistic Encyclopedic Dictionary of the Russian Language* (pp. 116–118). Moscow: Flinta, Nauka. (In Russ.).
Кожина М.Н. Функциональный стиль // Стилистический энциклопедический словарь русского языка / под ред. М.Н. Кожиной. М. : Флинта, Наука, 2003. С. 116–118.
12. Sampson, J. (1997). "Genre", "Style" and "Register": Sources of Confusion? *Revue Belge de Philologie et d'Histoire*, *75*(3), 699–708.
13. Moessner, L. (2001). Genre, Text Type, Style, Register: A Terminological Maze? *European Journal of English Studies*, *5*(2), 131–138.
14. Galperin, I.R. (1958). *Essays on the Style of the English Language*. Moscow: Literature in Foreign Languages publ. (In Russ.).
Гальперин И.Р. Очерки по стилистике английского языка. М. : Изд-во литературы на иностранных языках, 1958.
15. Vinogradov, V.V. (1963). *Stylistics. Theory of Poetic Speech, Poetics*. Moscow: AN SSSR Publ. (In Russ.). *Виноградов В.В.* Стилистика. Теория поэтической речи. Поэтика. М. : Изд-во АН СССР, 1963.

16. Kozhina, M.N. (1972). On the Speech System of the Scientific Style in Comparison with Some Others. Perm: Permian University Publ. (In Russ.).
Кожина М.Н. О речевой системности научного стиля сравнительно с некоторыми другими. Пермь : Изд-во Пермского университета, 1972.
17. Klimova, B.F. (2013). Scientific Prose Style and Its Specifics. *Procedia – Social and Behavioral Sciences*, *83*, 52–55.
18. Khomytska, I., & Teslyuk, V. (2017). The Method of Statistical Analysis of the Scientific, Colloquial, Belles-Lettres and Newspaper Styles on the Phonological Level. In: N. Shakhovska (Ed.) *Advances Inintelligent Systems and Computing* (Vol. 512, pp. 149–163). Cham: Springer International Publishing. https://doi.org/10.1007/978-3-319-45991-2_10 EDN: XNREIT
19. Mandravickaitė, J. (2020). *Linguistic Characteristics of Functional Styles: A Quantitative Analysis* [PhD Thesis]. Vilnius: Vilnius University. URL: www.vu.lt/lt/naujienos/ivykiu-kalendorius (accessed: 15.02.2025).
20. Kozhina, M.N. (1966). *On the Specifics of Artistic and Scientific Speech in the Aspect of Functional Stylistics*. Perm: Permian University Publ. (In Russ.).
Кожина М.Н. О специфике художественной и научной речи в аспекте функциональной стилистики. Пермь: Изд-во Пермского университета, 1966.
21. Troyanskaya, E.S. (1984). Field Structure of the Scientific Style and Its Genre Varieties. In: *General and Specific Problems of Functional Styles* (pp. 16–27). Moscow: Nauka. (In Russ.).
Троянская Е.С. Полевая структура научного стиля и его жанровых разновидностей // Общие и частные проблемы функциональных стилей. М. : Наука, 1984. С. 16–27.
22. Budagov, R.A. (1976). *Man and His Language*. Moscow: Moscow State University publ. (In Russ.).
23. *Будагов Р.А.* Человек и его язык. М. : Изд-во Московского университета, 1976.
24. Maevsky, N.N. (1979). *Peculiarities of the Popular Science Style* [PhD Thesis]. Rostov-na-Don: Rostov State University publ. (In Russ.).
Маевский Н.Н. Особенности научно-популярного стиля: дис. ... канд. филол. наук. Ростов-на-Дону: Ростовский государственный университет, 1979.
25. Perrault, S. (2013). *Communicating Popular Science: From Deficit to Democracy*. New York: Palgrave Macmillan.
26. Maier, C.D., & Engberg, J. (2021). Harvard Business Review's Reframing of Digital Communication: From Professional Expertise to Practical Guidance. *Journal of Pragmatics*, *176*, 186–197. https://doi.org/10.1016/j.pragma.2021.02.005
27. Clyne, M. (2002). Contrastive Discourse Studies. *Cahiers de Praxématique*, 38, 59–84.
28. Egorova, L.A. (2018). Popular Science Discourse Development in the Cyberspace. *Advances in Language and Literary Studies*, *9*(5), 79–83.
29. Aarseth, E. (1997). *Cybertext: Perspectives on Ergodic Literature*. Baltimore and London: Johns Hopkins University Press.
30. Herring, S.C. (2007). A Faceted Classification Scheme for Computer-Mediated Discourse. *Language@Internet*, *4*(1). https://www.languageatinternet.org/articles/2007/761/
31. Vashunina, I.V., Ryabova, M.E., & Egorova, L.A. (2018). Polycode Hypertext in Polylingual Discourse of Intercultural Communications. *XLinguae*, *11*(2), 218–231. https://doi.org/10.18355/XL.2018.11.02.17 EDN: XXPGCD
32. Shiryaeva, T.A., Bagiyan, A.Y., & Arakelova, A.R. (2018). Linguistic Representation of Professional Values in Popular Science Discourse off IT: Axiological and Pragmatic Perspective. *International Journal of Applied Linguistics and English Literature*, *7*, 52–57. https://doi.org/10.7575/aiac.ijalel.v.7n.7p.52 EDN: XXPGCD
33. Petrushka, A. (2018). The Functional Model of Medium of Popular Science Content. *Informatio et Scientia. Information Science Research*, *1*(1), 1–16.
34. Khomutova, T.N., & Petrov, S.G. (2013). Popular Science Text: Integral Model. *Bulletin of the South Ural State University. Series: Linguistics*, *10*(2), 37–41. (In Russ.). EDN: RAITHD

Хомутова Т.Н., Петров С.Г. Научно-популярный текст: интегральная модель // Вестник Южно-Уральского государственного университета. Серия: Лингвистика. 2013. Т. 10. № 2. С. 37–41. EDN: RAITHD

35. Khomutova, T.N. (2015). Strategies of Research Discourse: An Integral Approach. *Bulletin of the South Ural State University. Series: Linguistics*, *12*(3), 15–22. (In Russ.). EDN: RAITHD *Хомутова Т.Н.* Стратегии научного дискурса: интегральный подход // Вестник Южно-Уральского государственного университета. Серия: Лингвистика. 2015. Т. 12. № 3. 15–22. EDN: RAITHD
36. Tausczik, Y.R., & Pennebaker, J.W. (2010). The Psychological Meaning of Words: LIWC and Computerized Text Analysis Methods. *Journal of Language and Social Psychology, 29*(1), 24–54.
37. Pennebaker, J.W., Booth, R.J., Boyd, R.L., & Francis, M.E. (2015a). *Linguistic Inquiry and Word Count: LIWC2015*. Austin, TX: Pennebaker Conglomerates.
38. Nurken, A. (2017). Distinctive Features of Popular Science Discourse. In: *International Scientific and Practical Conference World Science* (Vol. 2, No. 11(27), pp. 32–35). ROST. EDN: ZTPIIT
39. Soodmand Afshar, H., Doosti, M., & Movassagh, H. (2018). A Comparative Study of Generic Structure of Applied Linguistics and Chemistry Research Articles: The Case of Discussions. *Journal of Research in Applied Linguistics*, *9*(1), 28–56. https://doi.org/10.22055/rals.2018.13404
40. Macedo-Rouet, M., Rouet, J.-F., Epstein, I., & Fayard, P. (2003). Effects of Online Reading on Popular Science Comprehension. *Science Communication*, *25*(2), 99–128. EDN: JTTAVB
41. Swales, J.M. (1981). *Aspects of Article Introductions*. Birmingham: The University of Aston, Language Studies Unit.
42. Dudley-Evans, T. (1994). Genre Analysis: A Key to a Theory of ESP? *Ibérica*, *4*, 3–19.
43. Ruiying, Y., & Allison, D. (2003). Research Articles in Applied Linguistics: Moving from Results to Conclusions. *English for Specific Purposes*, *22*, 365–385. EDN: BJFVBH
44. Uzuner, S. (2008). Multilingual Scholars' Participation in Core/Global Academic Communities: A Literature Review. *English for Academic Purposes*, *7*, 250–263.
45. Stoller, F.L., & Robinson, M.S. (2013). Chemistry Journal Articles: An Interdisciplinary Approach to Move Analysis with Pedagogical Aims. *English for Specific Purposes*, *32*, 45–57.
46. Martín, P., Rey-Rocha, J., Burgess, S., & Moreno, A. I. (2014). Publishing Research in English-Language Journals: Attitudes, Strategies, and Difficulties of Multilingual Scholars of Medicine. *Journal of English for Academic Purposes*, *16*, 57–67.
47. Pennebaker, J.W., Boyd, R.L., Jordan, K., & Blackburn, K. (2015b). *The Development and Psychometric Properties of LIWC2015*. Austin, TX: University of Texas at Austin.

Information about the authors:

Keith Topping, Professor of Educational and Social Research, Director of the Centre for Peer Learning, Director of the Higher Education Effective Learning Project, Co-Director of the national Read On project and of the national Problem-Solving, University of Dundee (Nethergate, Dundee, Scotland, UK, DD1 4HN); *Research interests*: problematic behaviour in schools; *e-mail*: k.j.topping@dundee.ac.uk
ORCID: 0000-0002-0589-6796.

Liudmila A. Egorova, PhD in Philology, Associate Professor of Linguistics, is the Head of the Linguistics Department at the Institute of Foreign Languages, RUDN University (6 Miklukho-Maklaya st., Moscow, Russian Federation, 117198); *Research interests*: L2 acquisition, discourse studies and pragmatics in close connection with education, new educational technologies, methodology of linguistics; *e-mail*: egorova-la@rudn.ru
ORCID: 0000-0002-5159-1512, SPIN-code: 3088-4697, AuthorID: 245980.

Natalia V. Nikashina, PhD in Philology, Associate Professor of Linguistics at the Department of Foreign Languages in Theory and Practice at the Institute of Foreign Languages, RUDN University (6 Miklukho-Maklaya st., Moscow, Russian Federation, 117198); *Research*

interests: discourse and translation studies, methodology of teaching foreign languages, education and new educational technologies; *e-mail*: nikashina-nv@rudn.ru
ORCID: 0000-0002-7523-8318, SPIN- code: 7374-6068, AuthorID: 613790.
Ekaterina V. Nagornova, PhD in Philology, Associate Professor of Linguistics at the Department of Foreign Languages in Theory and Practice at the Institute of Foreign Languages, RUDN University (6 Miklukho-Maklaya st., Moscow, Russian Federation, 117198); *Research interests*: pragmatics, linguodidactics, cognitive linguistics, linguistylistics, grammar, educational technologies and development of new methods of teaching foreign languages; *e-mail*: nagornova-ev@rudn.ru
ORCID: 0000-0003-4371-6001, SPIN-code: 4371-2057, AuthorID: 391796.

Сведения об авторах:
Кит Топпинг, профессор Университета Данди (DD1 4HN, Великобритания, Шотландия, Данди, Незергейт), директор Центра по изучению взаимообучения (Centre for Peer Learning), руководитель проекта эффективного обучения в высшем образовании, со-руководитель национального проекта Read On и национального проекта Problem-Solving; *сфера научных интересов*: вопросы межкультурной коммуникации и адаптации, разработка культурно-релевантных рекомендаций в области образовательной политики и педагогической практики; *e-mail*: k.j.topping@dundee.ac.uk
ORCID: 0000-0002-0589-6796.
Егорова Людмила Алексеевна, кандидат филологических наук, доцент кафедры теории и практики иностранных языков, руководитель отделения «Лингвистика» института иностранных языков, Российский университет дружбы народов (117198, Российская Федерация, г. Москва, ул. Миклухо-Маклая, д. 6); *сфера научных интересов*: дискурсивные исследования, мультимодальный дискурс, психолингвистические аспекты усвоения второго языка, новые образовательные технологии*; e-mail*: egorova-la@rudn.ru
ORCID: 0000-0002-5159-1512, SPIN-код: 3088-4697, AuthorID: 245980.
Никашина Наталья Викторовна, кандидат филологических наук, доцент кафедры теории и практики иностранных языков института иностранных языков, Российский университет дружбы народов (117198, Российская Федерация, г. Москва, ул. Миклухо-Маклая, д. 6); *сфера научных интересов*: дискурсивные исследования, переводоведение, методика преподавания иностранных языков, проблемы образования и внедрения новых образовательных технологий; *e-mail*: nikashina-nv@rudn.ru
RCID: 0000-0002-7523-8318, SPIN-код: 7374-6068, AuthorID: 613790.
Нагорнова Екатерина Валерьевна, кандидат филологических наук, доцент кафедры теории и практики иностранных языков института иностранных языков, Российский университет дружбы народов (117198, Российская Федерация, г. Москва, ул. Миклухо-Маклая, д. 6); *сфера научных интересов*: прагматика, лингводидактика, когнитивная лингвистика, образовательные технологии и разработка инновационных методик преподавания иностранных языков; *e-mail*: nagornova-ev@rudn.ru
ORCID: 0000-0003-4371-6001, SPIN-код: 4371-2057, AuthorID: 391796.

RUDN Journal of Language Studies, Semiotics and Semantics
2313-2299 (Print), ISSN 2411-1236 (Online)
Вестник Российского университета дружбы народов.
Серия: ТЕОРИЯ ЯЗЫКА. СЕМИОТИКА. СЕМАНТИКА

2026 Vol. 17 No. 2 712–727
http://journals.rudn.ru/semiotics-semantics

ФУНКЦИОНАЛЬНАЯ ГРАММАТИКА И СЕМАНТИКА

FUNCTIONAL GRAMMAR AND SEMANTICS

DOI: 10.22363/2313-2299-2026-17-2-712-727
EDN: MCDWZD

Научная статья / Research Article

Динамика морфологических норм русского языка XXI века: элиминация системных лакун vs грамматические архаизмы

Э.Г. Куликова ✉, В.В. Барабаш

Российский университет дружбы народов, *Москва, Российская Федерация*
✉kulikova_ella21@mail.ru

Аннотация. Грамматическая система представляет собой три подсистемы — словообразование, морфологию и синтаксис, из которых наименее изученной в плане динамики норм представляется морфология. Статья дает описательно-аналитическую трактовку реально наблюдаемых языковых фактов в сфере морфологических форм русского языка, исследование их свойств, взаимосвязей и функционирования, то есть интерпретацию их с синхронной точки зрения. Доказывается, что многие из тех форм, которые представляются приметой последних десятилетий, являются элиминированными грамматическими лакунами. Но так как всякая синхрония условна, неизбежно, что и на рассматриваемом этапе развития русского языка некоторые из этих новых форм по сути оказываются грамматическими архаизмами. Особенностям корреляции этих феноменов — грамматических архаизмов и элиминированных лакун (представляющихся в синхронии новообразованиями), установлению причин реархаизации грамматических форм — посвящена исследовательская часть статьи. Цель настоящей статьи состоит в установлении глубинных закономерностей, определяющих возможности морфологических форм к системным изменениям, что детерминирует и динамику грамматической нормы. Для достижения указанной цели наиболее адекватным представляется сочетание системно-описательного (с приемами внутриязыкового сопоставления, социолингвистической и лингвопрагматической интерпретации грамматических форм) и функционального подхода в диахроническом ракурсе, использование герменевтического инструментария, предполагающего актуализацию истолкования и объяснения. Эмпирическим материалом для исследования послужили главным образом медийные тексты на русском языке XXI в. (тексты иных периодов использовались только для сопоставления). Понятие элиминированной лакуны существенно для словоизменительных категорий, которые обнаруживают тенденцию к расширению парадигмы (именное число), и для словообразовательных категорий, увеличивающих охват лексических единиц (род одушевленных существительных). Современные

© Куликова Э.Г., Барабаш В.В., 2026
This work is licensed under a Creative Commons Attribution-NonCommercial 4.0 International License
https://creativecommons.org/licenses/by-nc/4.0/legalcode

границы русского литературного языка предполагают в том числе и активную «реархаизацию» грамматических архаизмов в качестве интенциональных единиц нестандартного текста.
Ключевые слова: языковые изменения, нормативность, морфология, грамматические лакуны

Вклад авторов: вклад авторов равнозначен на всех этапах исследования и подготовки текста статьи.

Конфликт интересов: авторы заявили об отсутствии конфликтов интересов.

История статьи: дата поступления: 21.11.2025; дата приема в печать: 5.12.2025.

Для цитирования: *Куликова Э.Г., Барабаш В.В.* Динамика морфологических норм русского языка XXI века: элиминация системных лакун vs грамматические архаизмы // Вестник Российского университета дружбы народов. Серия: Теория языка. Семиотика. Семантика. 2026. Т. 17. № 2. С. 712–727. https://doi.org/10.22363/2313-2299-2026-17-2-712-727 EDN MCDWZD

Dynamics of Morphological Norms of the Russian Language in the XXI Century: Elimination of Systemic Lacunae vs Grammatical Archaisms

Ella G. Kulikova ✉, Victor V. Barabash

RUDN University, *Moscow, Russian Federation*
✉kulikova_ella21@mail.ru

Abstract. The authors present descriptive and analytical interpretation of actually observed morphological forms in the modern Russian language, the study of their features, relationships and functioning, that is, their interpretation from a synchronous point of view. The article proves that many of the forms that seem to be a sign of recent decades are eliminated grammatical lacunae. Since any synchrony is conditional, in modern Russian it is inevitable for some of these new forms to turn out grammatical archaisms. The research part is devoted to the peculiarities of the correlation of these phenomena — grammatical archaisms and eliminated lacunae (appearing in synchrony as neoplasms), and the detection of grammatical forms "rearchaization". The purpose of this article is to establish deep patterns determining the abilities of morphological forms for systemic changes, which determines grammatical norm dynamics. To achieve this purpose, the combination of system-descriptive (with techniques of intra-linguistic comparison, sociolinguistic and linguopragmatic interpretation of grammatical forms) and functional approach in a diachronic perspective, the use of hermeneutic tools involving the actualization of interpretation and explanation is supposed to be most appropriate. First of all, the empirical material was media texts in Russian of the XXI century. The concept of an eliminated lacuna is significant for inflectional categories that present a tendency to expand the paradigm (number of noun), and for word-formation categories increasing the coverage of lexical units (genus of animate nouns). Modern boundaries of the Russian literary language include active "rearchaization" of grammatical archaisms as intentional units of a non-standard text.
Keywords: linguistic changes, normativity, morphology, grammatical lacunae

Authors' contribution: The authors contributed equally to this article and preparation of the article text**.**

Conflict of interests: The authors stated that there were no conflicts of interest.

Article history: Received: 22.11.2025; accepted: 5.12.2025.

For citation: Kulikova, E.G., & Barabash, V.V. (2026). Dynamics of Morphological Norms of the Russian Language in the XXI century: Elimination of Systemic Lacunae vs Grammatical Archaisms. *RUDN Journal of Language Studies, Semiotics and Semantics, 17*(2), 712–727. (In Russ.). https://doi.org/10.22363/2313-2299-2026-17-2-712-727 EDN MCDWZD

Введение

В теоретической морфологии многие проблемы нередко представляются решенными, поскольку уже даны завершенные списки узуальных грамматических форм всех категорий, а в нарративной грамматике морфологические формы чаще всего расцениваются как имеющие зависимое и второстепенное значение для текстопорождения и потому также не подвергаются специальному исследованию. Как известно, явления, традиционно относимые к морфологии, нередко могут быть интерпретированы на синтаксическом уровне. Более того, в англоязычной традиции популярна идея устанавливать практически равенство между морфологией и синтаксисом, определять, таким образом, грамматику обобщенно и рассматривать явления морфологии как принадлежащие либо лексике, либо синтаксису. Цель настоящей статьи состоит в установлении глубинных закономерностей, определяющих возможности морфологических форм к системным изменениям, что детерминирует и динамику грамматической нормы. Для достижения указанной цели наиболее адекватным представляется сочетание системно-описательного (с приемами внутриязыкового сопоставления, социолингвистической и лингвопрагматической интерпретации грамматических форм) и функционального подходов в диахроническом ракурсе, использование герменевтического инструментария, предполагающего актуализацию истолкования и объяснения. Эмпирическим материалом для исследования послужили медийные тексты на русском языке XXI в.; тексты иных периодов использовались только для сопоставления.

Полагаем, что вопрос о соотношении морфологии и синтаксиса вновь приобретает остроту для русской грамматики в связи с тотальной аналитизацией. Исследование грамматических явлений в лингвистике требует постоянного обновления эмпирического материала, поскольку именно такой анализ позволит выявить сохранение или изменение тенденций в грамматической системе конкретного языка, в частности русского. Динамика социолингвистических факторов в социуме закономерно влияет не только на лексику, но, опосредованно, также и на грамматику.

Дискуссия

Толерантность языковой нормы и элиминирование лакун

Общим тезисом о современной языковой норме стала констатация ослабления (даже падения) порога нормативности. Основополагающая гуманитарная идея толерантности распространилась и на сферу языковой нормы: авторитарная нормативность уступила место нормативности эластичной,

рекомендательной, в высшей степени демократичной (что проявилось в необыкновенном расширении сферы действия лексикографической пометы «доп.» — допустимо: средний род слова «кофе» и под.). Классическая отечественная лингвистика, начиная от Л.В. Щербы и В.В. Виноградова, отстаивала идею непродуктивности чрезмерно строгой нормализации на том основании, что норма в себе таит свое отрицание, ибо она постоянно подтачивается стремлением к неординарному (а значит, не соответствующему строгой норме) способу выражения как лексической, так и грамматической семантики. Как писал В.Г. Костомаров [1. С. 27], неизбежно появление «правильных неправильностей», поскольку, «с одной стороны, люди стремятся к порядку и однозначности, а с другой — к свежему разнообразию и яркой выразительности». См. также: [2. С. 126—133]. В современной лингвистике о «правильных неправильностях» идет речь в рамках теории лакунарности [3—14]. В рамках этой теории достигнуто понимание того, что наличие грамматических лакун может компенсироваться в системе языка. Системные лакуны представляют собой часть языковой картины мира, отражающей определенные этапы в его познании; они (как и наличествующие единицы) фиксируют национальную уникальность каждой лингвокультуры. Заполнение вакантных ячеек системы принято называть элиминированием лакун; для наименования восполняющих единиц применяется не безупречное с точки зрения логики наименование «элиминированные лакуны» (противоречие, собственно, в том, что исходная семантика слова «лакуна» не предполагает наличествующей реальной единицы). Термин этот, однако, широко распространен (см. перечисленные работы по лакунологии), применяется как к лексическим единицам, так и к грамматическим (морфологическим) формам.

Элиминирование интраязыковых лакун

Представления о морфологии как области чистой формы давно ушли в прошлое, акценты грамматической мысли сместились, и общепризнанным стало положение о том, что учет содержательной стороны граммем ведет к более объемному и сущностному пониманию формального строя языка. «Семантика — как лексическая, так и грамматическая — не может быть оторвана от всех тонкостей концептуализации, а также не может быть оторвана от социокультурных и прагматических факторов» [15. С. 591]. Так, нетривиальные грамматические формы числа могут быть вызваны к жизни именно совокупностью таких факторов. Существительное «риск» оказалось в современном языке «доминантно плюральным» (этот термин традиционно применяют для «нестойких» pluralia tantum типа *глаза*, *ботинки* и под.), поскольку понятие о риске стало в высшей степени дифференцированным. Плюральное использование этого слова типично не только для экономического и политического дискурсов, где оно используется в значениях ‘вероятность убытка или недополучения доходов по сравнению с прогнозируемой ситуацией’ и ‘мера опасности’ (*правовые, валютные, финансовые, банковские риски* и под.;

ср.: *Наступило время — такое периодически случается в истории, — когда волей-неволей приходится выбирать сторону. И вместе с тем все сопутствующие выбору* ***риски***[1]. Показательно название лингвистической статьи «Речевые **риски** цифровой эпохи…» [16]. То есть слово, заимствованное еще в XVIII в., подверглось семантической и грамматической трансформации именно под влиянием экстралингвистических условий [17. С. 33–35], и элиминированная лакуна (плюратив «риски») стала не только средством выразительности-оживления, нарушения монотонности восприятия (как большинство элиминированных лакун), но приобрела частотность и нормативный статус.

Задолго до того, как лакунология оформилась как достаточно строгая теория, было очевидно, что потенции грамматической формализации неизмеримо богаче ее наличествующего содержания. А.А. Зализняк[2] писал о потенциальных плюративах к любому существительному неконкретной семантики; Г.О. Винокур[3] — о потенциальных коррелятах женского рода (феминитивах), которые не функционируют в узусе, однако при необходимости всегда могут быть образованы и адекватно поняты. Вполне ясно было и то, что такие потенциальные формы востребованы прежде всего художественным дискурсом с его стремлением преодолеть стандарт (ср. предложенную трактовку понятия «интенциональность грамматической формы» в работах А.В. Бондарко, которая разрабатывается затем во многих трудах по грамматической стилистике; ср. также коррелирующие понятия авторского употребления, преднамеренной аномалии и дискурсивного слова). Проблема, однако, в каждом конкретном случае состоит в том, чтобы объяснить, при каких условиях неузуальное образование, восполняющее пробел в системной парадигме, может приблизиться к норме.

Множественное гетерогенное от существительных неконкретной семантики, отражающее неоднородность объекта, становится обычным по мере того, как все более детально и подробно познается сама эта неоднородность, ср.: *Но в другом мире несколько иные законы. Там открываются* ***все памяти*** *— кем человек был до того, как попал на небо. Мы живем, не зная, что с нами было, а там — полная информация о каждом из нас. Причем и о будущем*[4].

Такие элиминированные лакуны могут приобретать только дискурсивную узуальность (что отражается даже в лексикографических пометах: *памяти*, мн. научное, *плесени*, мн. научное). Но нередко эти формы становятся достоянием и обиходного языка. На такую особенность современной речи обратила внимание И.Б. Левонтина: «Обществу потребления

[1] *Власов П.* Эффект Чебурашки // Культура. 2023. № 1. С. 2.

[2] *Зализняк А.А.* Грамматический словарь русского языка. Словоизменение. М. : Рус. словари, 2003.

[3] Цит. по: *Бондарко А.В.* К проблеме интенциональности в грамматике // Вопросы языкознания. 1994. № 2. С. 29–42.

[4] *Бурляев Николай*. Смерти не боюсь – я уже был на том свете // Экспресс-газета. Спецвыпуск. 2021. № 9.

предлагаются готовые номенклатуры всего — в том числе вкусов, запахов, удовольствий», и к традиционному плюративу «цвет*á*» прибавились «пять вкусов» и семь запахов». «Потребителю остается только ткнуть пальцем — выбрать. Все каталогизировано, посчитано, только что не пронумеровано» [18. С. 100].

Плюрализация онимов также может быть отнесена к элиминированию лакун, поскольку это регулярный процесс, в результате которого вовсе необязательно осуществляется деонимизация. Случаи типа *Молчалины блаженствуют на свете* теперь уже традиционно объясняются с помощью понятия эллиптического плюраля, обозначающего множество с маркированным подмножеством (о разработке этого понятия и его применении к плюральным онимам (см.: [19]). Помимо эллиптического плюраля объяснительной силой в данном случае обладает понятие приблизительного множественного. Ср. формулировку О. Есперсена [20. С. 220], который подчеркивал: «Когда в одной форме объединяются несколько предметов или лиц, не принадлежащих в точном смысле к одному виду, мы будем называть ее формой «приблизительного» множественного числа (plural of approximation)».

У антономазии такого типа находят черты метафоры (О.С. Ахманова), метонимии (И.Г. Гальперин, А.П. Квятковский) и синекдохи (Ж. Дюбуа). Наиболее типичное прагматическое созначение таких форм — пейоративность, которая характерна для всех типов онимов — антропонимов, топонимов и всех разновидностей искусственной ономастики; ср.: *А мне не нравятся фильмы, этот муляж весь модного кино, этот кинематический дизайн, — как я его называю. Потому что это не искусство кино, а скорее кинематический дизайн. Ну абсолютно* ***все эти «Девятые роты», и прочие «Жары»***[5]. *И в самом деле — это лет пятнадцать назад те, кто ехал проводить время за границу, считались богатеями, а остальные злобно шептали вслед — ничего, ничего, гуляйте там в своих* ***Шарм-эль-шейхах****, а мы и дома…*[6].

Антономасии на основе числа представляют собой своего рода универсалию, во всяком случае, они вполне типичны в диахронии для языков европейского стандарта [21], но в каждом конкретном языке они являются элиминированными лакунами, восполняющими числовую парадигму, и при этом передают оценочность, которая складывается в каждой лингвокультуре особым образом под влиянием экстралингвистических факторов.

Одну из ярких примет современных языковых изменений в сфере морфологии можно обозначить как свободное градуирование. Оно проявляется в расширении парадигматики компаративности. Если прежде неузуальные степени сравнения (причем не только от прилагательных всех разрядов, но и от имен существительных и местоимений) были чертой поэтических идиостилей (например — М. Цветаевой), то сегодня это обыденный медийный и рекламный прием, ср.: *Мне кажется, что действительность превосходит всякое*

[5] Интервью с Т. Москвиной // Особое мнение, 18 декабря 2020 г.
[6] *Москвина Т.В.* Общая тетрадь. М. : АСТ, Астрель, 2009.

воображение Алексея Балабанова. И он бы не мог придумать ничего такого, чего не было. Нет, он человек не больной, а он человек творческий, беспокойный. Другой бы отмахнулся, а вот он думает, переживает. ***Вот очень русский, очень интересный***[7]*. Это я намекала на фильм «Королевство кривых зеркал», где подхалимы-придворные величают короля Ягупопа «****наикоролейшим*** *из королей»*[8]*. Двухэтажное здание школы № 6 на улице Семинарской находится в историческом центре Рязани. «****Центральнее****» некуда: в 10 минутах ходьбы — Рязанский кремль, в восьми минутах — Рязанское епархиальное управление*[9].

Оживленные дискуссии вызывает сегодня элиминирование лакун — коррелятов к антропонимам мужского рода [4; 22–26], поскольку именно в последнее десятилетие активизировались образования типа «блогерка», «поэтка», «режиссерка», «авторка» и под. Они вызывают эмоциональные отрицательные оценки: Татьяне Толстой не нравится слово «авторка», как в свое время А. Ахматовой, М. Цветаевой, а затем и Б. Ахмадулиной претило слово «поэтесса», ныне вполне узуальное. Ср.: *А. Балуев: До этого в течение последних 7 лет из Ксении Анатольевны делали политическую даму. Началось все с Болотной, где она показала, что вполне себе* ***оппозиционерка и либералка*** [10]*. Федор Бондарчук, которому молодая жена-актриса, дебютировавшая в качестве* ***сценаристки*** *сериала «Псих», — Паулина Андреева несколько дней назад родила сына Ваню, стал гостем* ***журналистки*** *…* [11].

В последнем примере очевидно различие степени узуальности двух женских коррелятов, поскольку в приведенном контексте первый феминитив может быть заменен формой мужского рода — «сценарист», в то время как замена второго проблематична: вне сугубо официального стиля более приемлем именно феминитив («журналистка»). Ср. также: *Вот и я выползла на улицу, ковыляю по центру и думаю: что же это мне не то чтобы хорошо, а как-то… менее противно. И вдруг осенило — так сняли наконец это жуткое новогоднее освещение! Нет больше синих деревьев, красных волн, фиолетовых звезд, зеленых месяцев и прочих нечеловеческих красот, олицетворявших, наверное, мечту о прекрасном* ***кассирши*** *со станции Хацапетовка*[12].

В этом примере феминитив в обычной роли — пейоративной оценки, однако особенность последнего десятилетия состоит в том, что эта функция феминитивов перестает быть главной. Напротив, новые феминитивы, скорее всего, есть порождение четвертой волны феминизма, докатившейся до России, и они (по мысли тех, кто их внедряет) должны как раз символизировать равенство и достоинство.

---

[7] Интервью с Т. Москвиной // Особое мнение, 18 декабря 2020 г.
[8] *Москвина Т.В.* Общая тетрадь. М. : АСТ, Астрель, 2009.
[9] Собеседник. 2021. № 29. С. 6. Минюстом РФ включен в список иноагентов.
[10] Собеседник. 11 марта 2021 г. Минюстом РФ включен в список иноагентов.
[11] Экспресс-газета. 2021. № 11.
[12] *Москвина Т.В.* Общая тетрадь. М. : АСТ, Астрель, 2009.

Вопреки представлениям о том, что собирательные существительные представляют собой архаический концепт, который должен постепенно утратиться с упрочением грамматического противопоставления по числу, в рамках этого разряда наблюдается элиминирование лакун, то есть образование новых имен с морфемными показателями собирательности. Ср.: Заголовок «***Паровозня***». Лид: *партии выставляют на выборы в Госдуму кандидатов-«паровозов», которые позволяют им выиграть кампанию — а потом заменяют их на нужных людей*[13].

Если все рассмотренные языковые факты могут быть квалифицированы как элиминирование интраязыковых лакун, то масштабная аналитизация русского языка последних десятилетий имеет прямое отношение к межъязыковой лакунарности.

Элиминирование межъязыковых лакун

Как известно, далеко не все авторы признают оба типа лакун [26], однако бесспорно, что общий признак релятивности объединяет интраязыковые и межъязыковые лакуны. Поэтому можно считать, что сущностно они не различаются, а способ выявления по отношению к системе сопоставляемого языка или в рамках одного языка не может иметь решающего значения для признания/непризнания того или иного вида лакунарности. Полагаем, что нарастание черт аналитизма в грамматике русского языка может быть рассмотрено как усиление лакунарности (нарастание различий) по отношению к тем славянским языкам, которые не обнаруживают такой тотальной аналитизации (сербскому, чешскому и отчасти польскому), и как элиминирование лакун (приближение к грамматическому строю) аналитических языков типа английского.

С одной стороны, бесспорно, что «русский язык в ходе исторического развития сохранил и существенно расширил свои синтетические возможности и арсенал синтетических средств» [27. С. 239], а с другой стороны, процесс аналитизации, который начался еще в XVIII в. с появлением первых нарицательных несклоняемых заимствованных существительных, именно в XXI в. приобрел особый динамизм [28]. В то время как в других славянских языках заимствования из западноевропейских языков (английского, французского) с разной степенью последовательности включаются в соответствующую парадигму склонения, в русском языке нарастают черты аналитизма. В настоящее время несклоняемых нарицательных существительных в русском языке, по разным оценкам, от 3000 до 5000, растет число инициальных аббревиатур, не имеющих системы словоизменения, и, самое примечательное, беспрецедентно возросло число аналит-прилагательных. Эту категорию в конце XX в. ввел в научный обиход М.В. Панов [29], который убедительно

[13] Новая газета. 26 июля 2021 г. Минюстом РФ гл. редактор «Новой газеты» включен в список иноагентов.

доказал, что это особая часть речи (а не разряд с нулевым склонением внутри традиционного прилагательного, поскольку русское прилагательное как часть речи организуется именно как согласуемая единица — это его категориальный признак). М.В. Панов перечислил полтора десятка разрядов аналит-прилагательных, среди которых были и совсем малочисленные. За прошедшие два десятилетия положение дел в корне изменилось: практически любое имя, даже вновь появившееся в языке, может выполнять функцию аналит-прилагательного. Ср. массовые образования со словом «ковид» типа *ковид-госпиталь, ковид-диссиденты.*

Источником распространения таких конструкций иногда считают язык рекламы. И.Б. Левонтина вполне справедливо отмечает: «…авторы не рассчитывают, что покупатель в магазине будет читать этикетку внимательно. Его глаз, скользя по полкам с товарами, выхватывает отдельные слова. И тут лучше, чтобы ключевые слова были в начальной форме. <…> Русский язык с его богатой морфологией — чистое наказание для писателя этикеток» [18. С. 268–269]. Отсюда использование таких обозначений, как *йогурт малина* (вместо малиновый) и даже *пемолюкс лимон сода эффект.*

Хотя приведенное наблюдение сомнений не вызывает, надо иметь в виду, что у тотального распространения аналит-прилагательных и вообще аналитизации более глубокие корни. С точки зрения классиков компаративистики, утрата флексий — диахроническая универсалия, при том, что, конечно, процесс идет в разных языках с разной степенью интенсивности. Как известно, в XIX в. приводилась следующая аргументация: усиление мыслительного компонента в деятельности человека определяло ослабление формы его языкового выражения, ибо набравший силу дух не нуждается, как прежде, в «замысловатых» языковых формах»[14]. Аналитический английский неслучайно считается в высшей степени эффективным: несложная (но вполне достаточная для выражения всех семантических нюансов) грамматика, отсутствие «назойливого» дублирования грамматической семантики (неизбежное во флективном языке), максимальное соответствие принципу экономии языковых и мыслительных усилий. О движении грамматики русского языка «в сторону английского», как известно, упоминал еще В.И. Ленин (имея в виду несклоняемые аббревиатуры и заимствования типа *метро*). И в наше время мы наблюдаем это ускоренное движение. Помимо приведенных фактов, в ряду явлений аналитизации нередко рассматривают утрату флективности у сложных и составных числительных, которые в разговорной речи обнаруживают отчетливую тенденцию к противопоставлению только двух типов форм — одну для именительного и винительного падежей, другую — для всех остальных. Масштабность явления свидетельствует, что это не просто отдельные нарушения грамматических норм отдельными недоучившимися носителями языка, а процесс, имеющий глубокие основания.

[14] *Гегель Г.В.Ф.* Сочинения. Т. 8. М. : Государственное социально-экономическое издательство, 1935. С. 60.

Итак, грамматические изменения, которые могут быть соотнесены с грамматической лакунарностью, обнаруживаются во многих звеньях системы, и изменения эти чаще всего видятся как новые явления.

Результаты

Элиминация системных лакун vs грамматические архаизмы

Однако при диахроническом анализе оказывается, что многие из рассмотренных элиминированных лакун коррелируют с грамматическими архаизмами. Так, многие из феминитивов, которые на синхронном срезе представляются производными феминистских идей конца XX — начала XXI вв., были известны еще в XIX в. К примеру, известно, что феминитив «поэтка» вовсе не новейший неологизм; он был впервые употреблен еще в 1842 г. [22. С. 6–22; 23. С. 7–26].

В целях стилизации Ю.Н. Тынянов в своем повествовании о Пушкине применяет многие феминитивы, имевшие хождение в пушкинскую эпоху: *Коли б не нетерпение, так была б* ***музыкантка****, — сказала Марья Алексеевна. Модная французская* ***живописица*** *Виже–Леорань писала теперь каждый день портреты модных красавиц и написала в два присеста крохотный портрет Надежды Осиповны, очень милый, с локонами*[15].

Вопрос в том, почему эти образования в течение десятилетий (даже столетий) оставались на периферии, во многом лежит в плоскости экстралингвистики. Показательно, например, что чрезвычайно активизировались такие номинации в первые годы советской власти, о чем свидетельствует материал Толкового словаря русского языка под ред. Д.Н. Ушакова. Многочисленные номинации типа *трактористка*, *красноармейка*, *избачка*, *птицеводка* и под. знаменовали новое положение женщины в стране победившего социализма. Затем бум на такого рода номинации прошел. Очевидно, что и сегодняшняя волна феминитивов имеет своим истоком социальные причины — от степени распространенности соответствующей профессии среди женщин (*режиссерка*, *сценаристка*) до общих идей политкорректности и толерантности.

Что касается элиминированных числовых лакун, характерных для современных текстов, то они также нередко оказываются «квазинеологизмами», которые использовались еще в текстах XIX в. Ср. фрагмент из последней книги А. Ширвиндта: *Когда мне стукнуло 85, я подумал, что это вполне серьезная цифра, чтобы себя сформулировать. Не получается. Мешают трусость и инфантильность. Трусость закрывает глаза на количество* ***вин*** *(не выпитых вин, что тоже можно инкриминировать, а совершенных за этот немалый срок проступков), а инфантильность бросает в слезливое оправдание моральных и физических преступлений, совершенных в борьбе с действительностью*[16].

[15] *Тынянов Ю.Н.* Пушкин. Ленинград : Гослитиздат, 1938.
[16] *Ширвиндт А.А.* Опережая некролог. М. : КоЛибри; Азбука-Аттикус, 2022.

Множественное число от слова «вина» (в родительном падеже составляющее омоформу к слову «вино») представляется настолько необычным современному языковому сознанию, что автор делает этот шутливый комментарий в скобках и разъясняет смысл с помощью синонима *проступки*. Между тем это самая обычная форма в классической литературе XIX в., которая нередко встречается в текстах И.А. Крылова, Л.Н. Толстого, Н.С. Лескова, Н.А. Некрасова и памятна носителям русского языка по школьной программе (*Досуг мне разбирать вины твои, щенок* или *За каплю крови, общую с народом, мои вины, о родина, прости*). И если для плюратива «риски» в современных реалиях сложились условия, позволяющие этой форме войти в узус, то слово «вина» никаких семантических приращений, которые повлияли бы на числовое поведение, не приобрело. Поэтому элиминированная лакуна «вины» остается интенциональной формой, незаменимой для передачи прагматических оттенков смысла (иронии, пейоративности), для создания оригинального текста [31], но не пригодной для стандартного использования.

Иные из рассмотренных грамматических инноваций — более свободное градуирование (в том числе образование степеней сравнения от существительных), широкое распространение аналит-прилагательных — находят серьезное обоснование в диахронии. Если вспомнить, что современное прилагательное развилось из прежде недифференцированного имени (которое могло, в зависимости от синтаксической позиции, выступать в качестве как существительного, так и прилагательного), то сегодняшние аналит-прилагательные представляются воплощением такого недифференцированного имени (ср. *удачный бизнес* и *бизнес-план*). Новые образования по продуктивным моделям в сфере собирательных имен опираются на субъективно-качественное представление множества, которое свойственно классу собирательных имен и благодаря которому этот класс даже после оформления грамматической категории числа (а собирательность исторически предшествовала множественному числу как способ представления грамматической семантики плюральности) остался живым и продуктивным.

Полагаем, что элиминированные грамматические лакуны (как рассмотренные, так и многие другие, например, будущее время русских причастий, расширение зоны переходности глагола как отражение активного действия; об этих изменениях в русской грамматике (см.: [30; 31]) могут быть включены в новые грамматические описания (наряду с дискурсивным анализом грамматических форм), в том числе в новую академическую грамматику, контуры которой стремятся наметить многие грамматисты [32; 33]. Такие единицы способствуют точности и выразительности словоупотребления и в то же время препятствуют «нормативному хаосу» [34; 35], преодолевают нарастающее в социуме безразличие к точному выбору слова вообще и нюансам грамматической семантики в частности.

Заключение

Итак, понятие элиминированной лакуны существенно для словоизменительных категорий, которые обнаруживают тенденцию к расширению парадигмы (именное число), и для словообразовательных категорий, увеличивающих охват лексических единиц (род одушевленных существительных). 40 лет назад было высказано мнение о том, что развитие таких категорий, как словоизменительное число и словообразовательный именной род, идет по пути усиления грамматической регулярности, то есть наращивания парадигмы, или, говоря в терминах лакунологии, поступательного компенсирования (элиминирования) системных лакун. По прошествии времени стало очевидно, что это движение не столь однозначно и одномерно. Есть сферы, перманентно активные (антономазии на базе категории числа), есть области, где указанная тенденция действует избирательно — затрагивая только определенные лексико-семантические пласты. В последнем случае играет роль широкозначность семантики, возможность градуирования исходного понятия. Современные границы русского литературного языка предполагают в том числе и активную «реархаизацию» грамматических архаизмов в качестве интенциональных единиц нестандартного текста.

Список литературы

1. *Костомаров В.Г.* Стилистика, любовь моей жизни. СПб. : Златоуст; М. : Гос. ИРЯ им. А.С. Пушкина, 2019.
2. *Киров Е.Ф.* Рецензия на монографию В.Г. Костомарова «Стилистика, любовь моей жизни…». СПб. : Златоуст; М. : Гос. ИРЯ им. А.С. Пушкина, 2019 // Русистика. 2020. Т. 18. № 1. С. 126–133. https://doi.org/10.22363/2618-8163-2020-18-1-126-133 EDN: XIUUGV
3. *Акай О.М., Беляева И.В.* Дескриптивная лакунографическая грамматика в свете идей проективной лингвистики // Язык и речь в Интернете: личность, общество, коммуникация, культура: сб. статей V Международной научно-практической конференции. Москва, РУДН, 22–23 апреля 2021 г. : в 2 т. М. : РУДН, 2021. Т. 1. С. 64–67. EDN: HQUQFA
4. *Антонова Е.Б.* Проблемы внедрения феминитивов-неологизмов в языковую систему современного русского языка // Молодой ученый. 2020. № 26 (316). С. 227–233. EDN: MGQAXX
5. *Беляева И.В., Куликова Э.Г.* Заимствование морфем как элиминирование лакун: «инговые» формы в русском языке // Вестник Московского государственного лингвистического университета. Гуманитарные науки. 2019. № 6 (822). С. 108–121. EDN: UEYWQS
6. *Ибрагимова К.Г., Рыкова П.А.* Межъязыковая лакунарность в синхронном переводе в рамках языка для специальных целей (на материале английского и русского языков) // Филологические науки. Вопросы теории и практики. 2021. Т. 14. № 11. С. 3565–3570. https://doi.org/10.30853/phil210561 EDN: TLOYTC
7. *Куликова Э.Г., Акай О.М., Беляева И.В.* Элиминированные лакуны vs потенциальные грамматические формы: корреляция понятий // Вестник Российского нового университета. Серия «Человек в современном мире». 2020. №. 4. С. 98–104. https://doi.org/10.25586/RNU.V925X.20.04.P.098 EDN: ANYOCQ
8. *Мануковская Т.В. Смотрова И.В.*, *Сложеницына Л.В.* Понятие «Языковая лакуна» в современной лингвистике на основе славянских и романо-германских языков // Филологические науки. Вопросы теории и практики. 2018. № 2 (80). Ч. 1. С. 120–123. EDN: YODNLK

9. *Anokhina T*. The Linguistic Lacunicon: Cogntive Mapping in Schemes and Terms // The Journal of Education, Culture and Society. 2013. Iss. 1. P. 166–172. https://doi.org/10.15503/jecs20131.166.174 EDN: DURVMR
10. *Grodzki E.* Using Lacuna Theory to Detect Cultural Differences in American and German Automotive Advertising. Frankfurt am Main; Berlin : Lang, 2003.
11. *Jolowicz J.O*. Lacuna theory in intercultural communication: Focus on axiological lacunae // Вопросы психолингвистики. 2006. Iss. 3. С. 58–77. EDN: JYCTJV
12. *Kshenovskaya (Sagnayeva) U.L*. Cultural Gaps and Ways of Eliminating Them // Text Processing and Cognitive Technologies. 2011. Iss. 20. P. 107–110.
13. *Panasiuk I., Schröder H*. Lakunen-Theorie: Ethnopsycholinguistische Aspekte der Sprach- und Kulturforschung. Münster–BerlinHamburg–London–Wien : LIT Verlag, 2006.
14. *Sun Y.* (Un)translatability and Cross-Cultural Readability // Perspectives: Studies in Translatology. 2012. Vol. 5. Iss. 20. P. 231–247.
15. *Бородай С.Ю.* Язык и познание. Введение в пострелятивизм. М. : Садра: ЯСК, 2020.
16. *Иванчук И.А.* Речевые риски цифровой эпохи на пересечении функциональных стилей (федеральный и региональный аспекты нормализаторской деятельности) // Язык и речь в Интернете: личность, общество, коммуникация, культура: сборник статей V Международной научно-практической конференции. Москва, РУДН, 22–23 апреля 2021 г. : в 2 т. М. : РУДН, 2021. Т. 1. С. 364–365. EDN: AXQOLH
17. *Захарова Е.В*. Семантические трансформации отвлеченного существительного «риск» // Язык в различных сферах коммуникации: материалы II Международной научной конференции. Чита : Забайкальский государственный университет, 2016. С. 33–35. EDN: YJENWR
18. *Левонтина И.Б.* Русский со словарем. М. : АСТ: CORPUS, 2017.
19. *Акай О.М.* Феномен грамматической лакунарности: когнитивный и лингвопрагматический аспекты. Ростов-на-Дону: ЮФУ, 2020. EDN: TVCBNH
20. *Есперсен О*. Философия грамматики. М. : УРСС, 2002.
21. *Зимина Е.А., Мюллер, Ю.Э.* Оценочный потенциал антономазии в языке современной немецкоязычной прессы // Филологические науки в МГИМО. 2021. № 7. С. 27–36. https://doi.org/10.24833/2410-2423-2021-1-25-27-36 EDN: FXOSRC
22. *Беркутова В.В*. Феминитивы в русском языке: исторический аспект // Филологический аспект. 2018. № 11 (43). С. 6–22.
23. *Беркутова В.В.* Феминитивы в русском языке: лингвистический аспект // Филологический аспект. 2019. № 1 (45). С. 7–26. EDN: YWXSTB
24. *Гузаерова Р.Р., Косова В.А.* Специфика феминитивов в современной русском медиапространстве // Филология и культура. Philololgy and culture. 2017. № 4 (50). С. 11–15. EDN: YLSHMI
25. *Пугачева Е.В*. Феминитивы: неологическая картина (по данным словарей) // Русский язык: исторические судьбы и современность: VI Международный конгресс исследователей русского языка. Москва, 20–23 марта 2019 г. Труды и материалы. М. : Изд-во МГУ, 2019. С. 88–89.
26. *Секерина М.Е., Захарова Ю.Г.* Феминитивы в языке блогов // Гуманитарные научные исследования. 2017. № 4. Режим доступа: https://human.snauka.ru/2017/04/23665 (дата обращения: 20.10. 2025). EDN: YNDADV
27. *Шульга М.В.* О векторе развития русской морфологической системы // Вестник Волгоградского государственного университета. Серия 2. Языкознание. 2017. Т. 16. № 4. С. 232–242. https://doi.org/10.15688/jvolsu2.2017.4.23 EDN: YTSDDN
28. *Сенько Е.В*. Аналитическое усилие современного русского языка // Русский язык: исторические судьбы и современность: материалы V Международный конгресс исследователей русского языка. Москва, 18–21 марта 2014 г. М. : Изд-во Московского университета, 2014. С. 143–144. EDN: TTQBTZ
29. *Панов М.В.* Позиционная морфология русского языка. М. : Наука, 1999.

30. *Булатбаева К.Н., Саркисеева Г.К., Айтжанов К.Б.* Окказиональное в современном языке: новые аспекты анализа // Вестник Чувашского государственного педагогического университета им. И.Я. Яковлева. 2020. № 1(106). С. 10–16. https://doi.org/10.37972/chgpu.2020.51.80.002 EDN: WOFAOB
31. *Эпштейн М.Н.* От знания — к творчеству. Как гуманитарные науки могут изменять мир. М.; СПб.: Центр гуманитарных инициатив, 2016.
32. *Радбиль Т.Б.* Национальная специфика грамматики русского языка: когнитивный и лингвокультурологический аспекты // Русская грамматика: сборник тезисов Международного научного симпозиума. Москва, 13–15 апреля 2016 г. М. : Гос. ИРЯ им. А.С. Пушкина, 2016. С. 96–99. EDN: XGTGKF
33. *Рогова К.А*. На пути к новой академической грамматике // Мир русского слова. 2018. № 2. С. 5–12. EDN: XWLMHZ
34. *Хазагеров Г.Г*. Риторический манифест. Ростов-на-Дону: Изд-во Ростовской риторической школы, 2020.
35. *Брусенская Л.А., Куликова Э.Г.* Рецензия на «Риторический манифест» Г.Г. Хазагерова (Ростов-на-Дону: Изд-во Ростовской риторической школы, 2020. 45 с.) // Коммуникативные исследования. 2021. Т. 8. № 1. С. 185–193. https://doi.org/10.24147/2413-6182.2021.8(1).185-193 EDN: GXTAFS

References

1. Kostomarov, V.G. (2019). *Stylistics, the Love of My life.* Saint Petersburg: Zlatoust publ.; Moscow: Gos. A.S. Pushkin IRL publ. (In Russ.). https://doi.org/10.22363/2618-8163-2020-18-1-126-133 EDN: XIUUGV
2. Kirov, E.F. (2020). Review on the Monograph “Stylistics, the Love of My Life...” V.G. Kostomarov. *Russian Language Studies*, *18*(1), 126–133. (In Russ.). https://doi.org/10.22363/2618-8163-2020-18-1-126-133 EDN: XIUUGV
3. Akay, O.M., & Belyaeva, I.V. (2021). Descriptive Lacunographic Grammar in the Light of The Ideas of Projective Linguistics. In: *Language and Speech on the Internet: Personality, Ssociety, Communication, Culture*: Collection of Articles of the V International Scientific and Practical Conference. Moscow, RUDN University, April 22–23, 2021: in 2 vols. (Vol. 1. pp. 64–67). Moscow: RUDN publ. (In Russ.). EDN: HQUQFA
4. Antonova, E.B. (2020). Problems of Introducing Feminitives-Neologisms Into the Linguistic System of the Modern Russian Language. *Young Scientist,* 26 (316), 227–233. EDN: MGQAXX
5. Belyaeva, I.V., & Kulikova, E.G. (2019). Morpheme Borrowing as Elimination of Lacunae: “ing” Forms in Russian. *Vestnik of Moscow State Linguistic University. Humanities,* 6 (822), 108–121. (In Russ.). EDN: UEYWQS
6. Ibragimova, K.G., & Rykova, P.A. (2021). Interlanguage Lacunarity in Simultaneous Translation within the Framework of a Language for Special Purposes (Based on the Material of English and Russian Languages). *Philology. Theory & Practice*, *14*(11), 3565–3570. (In Russ.). https://doi.org/10.30853/phil210561 EDN: TLOYTC
7. Kulikova, E.G., Akay, O.M., & Belyaeva, I.V. (2020). Eliminated Lacunae Vs Potential Grammatical Forms: Correlation of Concepts. *Vestnik of the Russian New University. The series “Man in the Modern World”*, 4, 98–104. (In Russ.). https://doi.org/10.25586/RNU.V925X.20.04.P.098 EDN: ANYOCQ
8. Manukovskaya, T.V., Smotrova, I.V., & Kolesenitsyna, L.V. (2018). The Concept of “Linguistic Lacuna” in Modern Linguistics Based on Slavic and Romano-Germanic Languages. *Philology. Theory & Practice,* 2 (80), I, 120–123. (In Russ.). EDN: YODNLK
9. Anokhina, T. (2013). The Linguistic Lacunicon: Cogntive Mapping in Schemes and Terms. *The Journal of Education, Culture and Society*, 1, 166–172. https://doi.org/10.15503/jecs20131.166.174 EDN: DURVMR

10. Grodzki, E. (2003). *Using Lacuna Theory to Detect Cultural Differences in American and German Automotive Advertising*. Frankfurt am Main; Berlin: Lang.
11. Jolowicz, J.O. (2006). Lacuna Theory in Intercultural Communication: Focus on Axiological Lacunae. *Journal of Psycholinguistics*, 3, 58–77. EDN: JYCTJV
12. Kshenovskaya (Sagnayeva), U.L. (2011). Cultural Gaps and Ways of Eliminating Them. *Text Processing and Cognitive Technologies*, 20, 107–110.
13. Panasiuk, I., & Schröder, H. (2006). *Lakunen-Theorie: Ethnopsycholinguistische Aspekte der Sprach- und Kulturforschung*. Münster – BerlinHamburg – London – Wien: LIT Verlag.
14. Sun, *Y.* (2012). (Un)translatability and Cross-Cultural Readability. *Perspectives: Studies in Translatology*, *5*(20), 231–247.
15. Borodai, S.Y. (2020). *Language and Cognition. Introduction to Post-Relativism*. Moscow: OOO «Sadra»: YASK publ. (In Russ.).
16. Ivanchuk, I.A. (2021). Speech Risks of the Digital Age at the Intersection of Functional Styles (Federal and Regional Aspects of Normalizing Activities). In: *Language and Speech on the Internet: Personality, Society, Communication, Culture: Collection of Articles of the V International Scientific and Practical Conference*. (Moscow, RUDN University, April 22–23, 2021): in 2 vols. (Vol. 1. pp. 364–365). Moscow: RUDN University publ. (In Russ.). EDN: AXQOLH
17. Zakharova, E.V. (2016). Semantic Transformations of the Abstract Noun "Risk". In: *Language in Various Fields of Communication: Proceedings of the II International Scientific Conference* (pp. 33–35). Chita: Zabaykalsky State University publ. (In Russ.). EDN: YJENWR
18. Levontina, I.B. (2017). *Russian with a Dictionary*. Moscow: AST CORPUS publ. (In Russ.).
19. Akay, O.M. (2020). *The Phenomenon of Grammatical Lacunarity: Cognitive and Linguopragmatic Aspects*. Rostov-on-Don: YUFU publ. (In Russ.). EDN: TVCBNH
20. Espersen, O. (2002). *Philosophy of Grammar*. Moscow: URSS publ. (In Russ.).
21. Zimina, E.A., & Muller, Yu.E. (2021). The Evaluative Potential of Antonomasia in the Language of the Modern German-Language Press. *MGIMO Journal of Language Research and Teaching*, 7, 27–36. (In Russ.). https://doi.org/10.24833/2410-2423-2021-1-25-27-36 EDN: FXOSRC.
22. Berkutova, V.V. (2018). Feminitives in the Russian Language: Historical Aspect. *Filologicheskij Aspekt: Mezhdunarodnyj Nauchno-Prakticheskij Zhurnal,* 11 (43), 6–22. (In Russ.).
23. Berkutova, V.V. (2019). Feminitives in the Russian Language: Linguistic Aspect. *Filologicheskij Aspekt: Mezhdunarodnyj Nauchno-Prakticheskij Zhurnal*, 1 (45), 7–26. (In Russ.).EDN: YWXSTB
24. Guzaerova, R.R., & Kosova, V.A. (2017). The Specificity of Feminitives in the Modern Russian Media Space. *Филология и культура. Philololgy and Culture*, 4 (50), 11–15. (In Russ.). EDN: YLSHMI
25. Pugacheva, E.V. (2019). Feminitives: a Neological Picture (According to Dictionaries). In: *Russian Language: Historical Destinies and Modernity: The VI International Congress of Russian Language Researchers*. (Moscow, March 20–23, 2019). Proceedings and Materials (pp. 88–89). Moscow: Moscow State University publ. (In Russ.).
26. Sekerina, M.E., & Zakharova, Yu.G. (2017). Feminitives in the Language of Blogs. *Humanitarian Scientific Research*, 4. URL: https://human.snauka.ru/2017/04/23665 (accessed: 20.10.2025). (In Russ.). EDN: YNDADV
27. Shulga, M.V. (2017). About the Vector of Development of the Russian Morphological System. *Science Journal of Volgograd State University. Linguistics. Seriya 2. Yazykoznanie*, *16*(4), 232—242. (In Russ.). https://doi.org/10.15688/jvolsu2.2017.4.23 EDN: YTSDDN
28. Senko, E.V. (2014). The Analytical Effort of The Modern Russian Language. In: *Russian Language: Historical Destinies and Modernity: The V International Congress of Russian Language Researchers.* (Moscow, March 18–21, 2014). Proceedings and materials (pp. 143–144.). Moscow: Moscow State University publ. (In Russ.). EDN: TTQBTZ

29. Panov, M.V. (1999). *Positional Morphology of the Russian language*. Moscow: Nauka publ. (In Russ.).
30. Bulatbaeva, K.N., Sarkiseeva, G.K., & Aitzhanov, K.B. (2020). Occasionalisms in Modern Language: New Aspects of Analysis. *I. Yakovlev Chuvash State Pedagogical University Bulletin*, 1(106), 10–16. (In Russ.). https://doi.org/10.37972/chgpu.2020.51.80.002 EDN: WOFAOB
31. Epstein, M.N. (2016). *From Knowledge to Creativity. How the Humanities Can Change the World*. Moscow; Saint Petersburg: Center for Humanitarian Initiatives. (In Russ.).
32. Radbil', T.B. (2016). National Specifics of the Grammar of the Russian Language: Cognitive and Linguocultural Aspects. In: *Russian Grammar: A Collection of Abstracts of The International Scientific Symposium* (Moscow, April 13–15, 2016) (pp. 96–99). Moscow: A.S. Pushkin IRL publ. (In Russ.). EDN: XGTGKF
33. Rogova, K.A. (2018). Towards a New Academic Grammar. *World of the Russian Word*, 2, 5–12. (In Russ.). EDN: XWLMHZ
34. Khazagerov, G.G. (2020). *The Rhetorical Manifesto*. Rostov-on-Don: Rostov Rhetorician School publ. (In Russ.).
35. Brusenskaya, L.A., & Kulikova, E.G. (2021). Review on "Rhetorical Manifesto" by G.G. Khazagerov (Rostov-on-Don: Rostov rhetorician school publ. 2020. 45 p.). *Communication Studies*, *8*(1), 185–193. (In Russ.). https://doi.org/10.24147/2413-6182.2021.8(1).185-193 EDN: GXTAFS

Сведения об авторах:
Куликова Элла Германовна, доктор филологических наук, профессор, профессор кафедры массовых коммуникаций филологического факультета, Российский университет дружбы народов (117198, Российская Федерация, г. Москва, ул. Миклухо-Маклая, д. 6); *e-mail*: kulikova-eg@rudn.ru
ORCID: 0000-0002-5305-8789; SPIN-код: 6607-2130; AuthorID: 360662; Scopus AuthorID: 56511592000; ResearcherID: R-2961-2016.
Барабаш Виктор Владимирович, доктор филологических наук, профессор, декан филологического факультета, заведующий кафедрой массовых коммуникаций филологического факультета, Российский университет дружбы народов (117198, Российская Федерация, г. Москва, ул. Миклухо-Маклая, д. 6); *e-mail*: barabash-vv@rudn.ru
ORCID: 0000-0002-8909-2145; Scopus: 57191418034; SPIN-код: 5585-6215; AuthorID: 268606; Scopus AuthorID: 57202091203; ResearcherID: ABC-3166-2021.

Information about the authors:
Ella G. Kulikova, Dr. Sc. (Philology), Professor, Professor of Mass Communications Department, Faculty of Philology, RUDN University (6 Miklukho-Maklaya str., Moscow, Russian Federation, 117198); *e-mail*: kulikova-eg@rudn.ru
ORCID: 0000-0002-5305-8789; SPIN-code: 6607-2130; AuthorID: 360662; Scopus AuthorID: 56511592000; ResearcherID: R-2961-2016.
Victor V. Barabash, Dr. Sc. (Philology), Professor, Dean of Faculty of Philology, Head of Mass Communications Department, RUDN University (6 Miklukho-Maklaya str., Moscow, Russian Federation, 117198); *e-mail*: barabash-vv@rudn.ru
ORCID: 0000-0002-8909-2145; Scopus: 57191418034; SPIN-code: 5585-6215; AuthorID: 268606; Scopus AuthorID: 57202091203; ResearcherID: ABC-3166-2021.

RUDN Journal of Language Studies, Semiotics and Semantics
2313-2299 (Print), ISSN 2411-1236 (Online)
Вестник Российского университета дружбы народов.
Серия: ТЕОРИЯ ЯЗЫКА. СЕМИОТИКА. СЕМАНТИКА

2026 Vol. 17 No. 2 728–742
http://journals.rudn.ru/semiotics-semantics

DOI: 10.22363/2313-2299-2026-17-2-728-742
EDN: MEFHQQ

Научная статья / Research Article

Функциональный и коллокационный потенциал существительного *coup* в современном французском языке

А.М. Червоный[1], Ж.М. Арутюнова[2], Н.Ю. Нелюбова[2]

[1]Ростовский государственный экономический университет, *Таганрог, Российская Федерация*
[2]Российский университет дружбы народов, *Москва, Российская Федерация*
neliubova-nyu@rudn.ru

Аннотация. Актуальность рассмотрения функционального потенциала существительного *coup (m)* в современном французском языке определяется наличием ряда проблемных аспектов семантики и роли компонененты *coup (m)* в продуцировании коллокаций во французском языке и речи. Цель работы — выявление коллокаций французского языка, в которых *coup (m)* выступает основным структурным элементом, а также в рассмотрении их структурных особенностей и номенклатуры значений. Эмпирическим материалом для исследования послужили более 300 случаев как отдельного употребления существительного *coup (m)*, так и употребления этого слова во французских устойчивых словосочетаниях и выражениях, извлеченных из словарей и художественных произведений французских авторов, а также зафиксированных в процессе коммуникации с носителями французского языка. Были использованы метод сплошной выборки, методы семантического, компонентного, структурного и стилистического анализа. Как следует из проведенного исследования, продуцирование коллокаций существительным *coup (m)* объясняется экономным использованием речевых усилий при его произнесении, а также весьма широкой семантикой, что способствует активному процессу образования различных французских лексических единиц и сложных лексических конструкций с участием элемента *coup (m)*. В статье раскрывается важность изучения процесса образования и функционирования устойчивых словосочетаний, коллокаций как в научно-теоретическом аспекте, так и вследствие практической значимости владения ими в процессе формирования навыков иноязычной устной и письменной речи. В работе рассмотрена номенклатура значений данной лексической единицы, выявлены ее сочетательные возможности, проанализированы синтаксические структуры — коллокации, в которых она использована, проведена систематизация структурных и семантических особенностей коллокаций. *Coup (m)* во французском языке выступает определенной платформой — базовым элементом образования сложно оформленных лексических единиц.

Ключевые слова: семантика, сочетаемость, синтагма, коллокация, структура

Вклад авторов: равнозначен на все этапах исследования и подготовки статьи к публикации.

Конфликт интересов: авторы заявили об отсутствии конфликтов интересов.

© Червоный А.М., Арутюнова Ж.М., Нелюбова Н.Ю., 2026
This work is licensed under a Creative Commons Attribution-NonCommercial 4.0 International License
https://creativecommons.org/licenses/by-nc/4.0/legalcode

История статьи: дата поступления: 20.01.2026; дата приема в печать: 10.02.2026.

Для цитирования: *Червоный А.М., Арутюнова Ж.М., Нелюбова Н.Ю.* Функциональный и коллокационный потенциал существительного *coup* в современном французском языке // Вестник Российского университета дружбы народов. Серия: Теория языка. Семиотика. Семантика. 2026. Т. 17. № 2. С. 728–742. https://doi.org/10.22363/2313-2299-2026-17-2-728-742 EDN MEFHQQ

Functional and Collocational Potential of the Noun *Coup* in Modern French

Alexander M. Chervony[1], Zhanna M. Arutyunova[2], Natalia Yu. Neliubova[2] ✉

[1]Rostov State University of Economics, *Taganrog*, *Russian Federation*
[2]RUDN University, *Moscow*, *Russian Federation*
✉neliubova-nyu@rudn.ru

Abstract. This article examines the functional and collocational potential of the noun *coup* (masculine) in contemporary French. The relevance of this research stems from several unresolved issues concerning the role of *coup* in collocation formation, as well as its semantics and functioning within the French language and discourse. The primary objective of this study was to identify French collocations where *coup* serves as the core structural element, and to analyze their structural features and range of semantic meanings. The empirical data for this investigation comprised over 300 instances of both the standalone use of *coup* and its incorporation into established French word combinations and expressions. These examples were drawn from dictionaries and literary works by French authors, as well as from recordings of communication with native French speakers. To achieve the research objective, the study employed the method of continuous sampling, along with semantic, componential, structural, and stylistic analysis. The findings indicate that the noun *coup* readily participates in collocation production due to the economy of linguistic effort involved in its pronunciation and its exceptionally broad semantic scope. This, in turn, facilitates the active formation of diverse French lexical units and complex lexical constructions involving *coup*. The article highlights the importance of studying the formation and functioning of fixed expressions and collocations, both from a theoretical perspective and for their practical value in developing oral and written communication skills in a foreign language. The paper examines the range of semantic meanings of this lexical unit, identifies its combinatory possibilities, analyzes the syntactic structures — collocations — in which it is used, and systematizes the structural and semantic features of these collocations. In French, *coup* (m) serves as a foundational platform — a basic element for the formation of complex lexical units.

Keywords: noun *coup* (m), semantics, combinability, syntagma, collocation, structure

Authors' contribution: The authors contributed equally to this article and preparation of the article text.

Conflict of interests: The authors stated that there were no conflicts of interest.

Article history: Received: 20.01.2026; Accepted: 10.02.2026.

For citation: Chervony, A.M., Arutyunova, Zh.M., & Neliubova, N.Yu. (2026). Functional and Collocational Potential of the Noun *Coup* in Modern French. *RUDN Journal of Language Studies, Semiotics and Semantics, 17*(2), 728–742. (In Russ.). https://doi.org/10.22363/2313-2299-2026-17-2-728-742 EDN MEFHQQ

Введение

Как известно, в каждом языке существует группа слов, наиболее употребительных в речи. Во французском языке одной из таких лексических единиц является существительное *coup* (m) и его многочисленные производные словосочетания. Это слово — моносиллабическое, а, согласно закону Ципфа [1. P. 484–490], в любом естественном языке самые употребительные слова будут самыми короткими. Высокая частотность употребления этого слова объясняется, на наш взгляд, экономным использованием речевых усилий при его произнесении, весьма широкой семантикой, что способствует активному процессу образования различных французских лексических единиц и сложных лексических конструкций с участием элемента *coup* (m).

В настоящее время в современном французском языке очень широко используется форма этого слова *du coup (fam)*, которое заменяет в речи следующие лексические единицы: *ainsi* ‘таким образом, итак, следовательно’, donc ‘значит, итак’, *dès lors* ‘следовательно’, *tout à coup* ‘вдруг, неожиданно’, *en conclusion* ‘в заключение’, *c’est pourquoi* ‘вот почему, поэтому’, *par conséquent* ‘значит, следовательно’, *par suite* ‘следовательно, и поэтому’, *soudainement* ‘внезапно, неожиданно, вдруг’ и т.д.

На русский язык выражение *du coup* переводится союзами, союзными словами, наречиями: *таким образом*, *итак*, *следовательно*, *значит*, *вдруг*, *неожиданно*, *внезапно*, в *заключение*, *вот почему*, *поэтому*, *и поэтому* и т.д. *Du coup* эквивалентно даже целому выражению *si je comprends bien* ‘если я хорошо понимаю…’.

Полисемантизм *du coup*, с одной стороны, упрощает коммуникацию, но с другой — чрезмерно частотное его употребление во французской речи снижает выразительный потенциал языка. Высокая частотность употребления существительного *coup* (m) отмечается в разных регистрах языка и речи.

Этот субстантив в речи используется в первую очередь для выражения какого-либо движения, быстроты действия, динамизма происходящих событий. Так, существительное *coup* (m) весьма частотно во французском спортивном дискурсе при описании действий субъекта в рамках проведения спортивных игр и состязаний. Что совершенно естественно, ведь спорт — это прежде всего движение: *coup bien placé* [ajusté] ‘точный удар’; *coup de but* ‘удар от ворот’; *coup de coin* ‘угловой удар’; *coup franc direct* ‘штрафной удар’; *coup franc indirect* ‘свободный удар’; *coup coupé* ‘резаный удар’; *coup plat* ‘плоский удар’; *porter un coup* ‘нанести удар’; *coup bas* ‘удар ниже пояса’ (также перен. знач.); *coup de pied à la lune* ‘прыжок из передней стойки назад’.

Популярность этого слова подчеркивает тот факт, что оно входит в состав группы французских устойчивых выражений: *Il y a des coups de pied au derrière qui se perdent* (разг.) ‘кое-кто заслуживает хорошего пинка’. *Ne pas en foutre un coup* (прост.) *Ни черта не делать, палец о палец не ударить.* (ср. [2]). *....il aime boire un bon coup* (J.-P. Chabrol. *La dernière cartouche*) ‘…он любит

выпить как следует'. *Coup* (m) является также частью наречных выражений, например: *tout à coup* 'вдруг', *tout d'un coup* 'вдруг, внезапно'; *après coup* (loc adv) 'после (совершившегося факта), потом, слишком поздно; задним числом': *à coup sûr* (loc adv) 'наверняка, без риска'; *à coup perdu* (loc adv) 'наудачу' и т.д.

Теоретическая основа исследования

Следует отметить ведущую структурно-семантическую роль существительного *coup* (m) в выявленных устойчивых сочетаниях — французских фразеологизированных и нефразеологизированных коллокациях.

Термин *коллокация*[1], введенный в научный обиход Х. Пальмером, Д. Ферсом, выражает одно из основных понятий современной дидактики и комбинаторной лингвистики (см. [3–10]). Ввиду теоретической важности рассмотрения процесса образования и функционирования устойчивых словосочетаний, коллокаций, а также вследствие практической значимости владения устойчивыми словосочетаниями в ходе освоения иностранного языка *коллокации* стали предметом изучения многих зарубежных и отечественных лингвистов (см. [3–7; 9–13] и др.).

Проблема выбора слов, сочетаний слов и комбинирования в предложении, образования устойчивых словосочетаний, к которым относим и коллокации, лежит в плоскости синтагматики, синтагматических отношений[2] (см. [12; 14]). Исследование коллокаций затрагивает также проблемы семантики лексических единиц, их валентности[3], сочетаемости сем[4], лексико-грамматической и стилистической сочетаемости лексических единиц, их функционирования в речи, роли узуса и национальных особенностей образования коллокаций того или иного языка и т.д. [15–19].

Актуальность исследования. Как показал анализ языкового материала и обращение к научно-теоретическим источникам, рассмотрение особенностей коллокации по-прежнему остается актуальной темой лингвистических исследований [3–8; 11; 14; 20–22]. Для этого существуют определенные причины. Так, до настоящего времени отмечается значительная вариативность в определении понятия *коллокация*, предлагаются отличные друг от друга характеристики коллокаций, приводятся разные их классификации [8]. Возникшие разночтения объясняются, по нашему мнению, сложностью данного предмета исследования, его многоаспектностью и многогранностью, а также разными подходами лингвистов к изучению этого языкового явления [8. С. 20–25].

---

[1] *Ахманова О.С.* Коллокация // Словарь лингвистических терминов. М., 1966. С. 199.

[2] *Кубрякова Е.С.* Синтагматика // Лингвистический энциклопедический словарь. М. : Советская энциклопедия, 1990. С. 447–448.

[3] *Гак В.Г.* Валентность // Лингвистический энциклопедический словарь. М. : Советская энциклопедия, 1990. С. 79–80.

[4] Там же. С. 483.

Дополнительного изучения, по нашему мнению, требует механизм отбора ядерных элементов коллокации. Заслуживает также более глубокого внимания рассмотрение национальных особенностей образования устойчивых сочетаний того или иного языка, превалирующие в нем типы сочетаемости, следует подвергнуть дополнительному анализу структурные композиции коллокаций, изменение их значения в зависимости от контекста и коммуникативной ситуации, а также вследствие их перехода из одного стилистического уровня на другой. В своей работе мы будем опираться на мнение В.П. Захарова и М.В. Хохловой, согласно которому коллокации — это статистически устойчивые словосочетания, которые могут быть фразеологизированными и свободными [5. С. 138].

Цель работы состоит в выявлении коллокаций французского языка, в которых существительное *coup* (m) выступает основным структурным элементом, а также в рассмотрении их структурных особенностей и номенклатуры значений. *Задачи исследования*: 1. выявить сочетательные возможности субстантива *coup* (m); 2. проанализировать синтаксические структуры коллокаций, в которых использован субстантив *coup* (m); 3. рассмотреть семантику коллокаций с субстантивом *coup* (m); 4. выявить специфику семантики коллокации с *coup* (m) в зависимости от ее стилистической принадлежности; 5. провести структурную и семантическую систематизацию коллокаций рассматриваемого типа. *Объектом исследования* являются коллокации с субстантивом *coup* (m), их структурные, семантические и стилистические особенности.

Материал исследования. Эмпирическим материалом исследования послужили более 300 случаев как отдельного употребления существительного *coup* (m), так и употребление этого слова во французских устойчивых словосочетаниях и выражениях, извлеченных из словарей и художественных произведений французских авторов, а также зафиксированных в процессе коммуникации с носителями французского языка.

Методы исследования. В процессе работы над языковым материалом использовались семасиологический и ономасиологический подходы, что позволило рассмотреть объект исследования, двигаясь как от формы к содержанию, так и от содержания к форме выражения. Для реализации поставленной в исследовании цели были использованы метод сплошной выборки, методы семантического, компонентного, структурного и стилистического анализа.

Обсуждение и результаты

Во французском языке сочетание согласного/гласного звуков *cou* [ku] довольно часто используется в начале французских слов. Чтобы в этом убедиться, достаточно открыть любой словарь французского языка. Это буквосочетание, со ставшим непроизносимым в конце согласным *p*, претерпев грамматические и фонетические трансформации вследствие перехода базового существительного из латинского языка во французский (из классической латыни *colaphus* через форму народной латыни *colpus*), дало

французскому языку краткое однослоговое существительное мужского рода *coup* (m)[5].

Как известно, первичное значение французского существительного *coup* (m) ‘удар’. В своем первичном значении это существительное может также переводиться на русский язык как ‘толчок, пинок, тычок, тумак’. Согласно С.И. Ожегову, «*удар* — это короткое и сильное движение, непосредственно направленное на кого-что-нибудь, резкий толчок»[6].

Существительное *coup* (m) — моносиллабическое, не вызывающее затруднений в своем произношении и орфографии слово, оно омонимично французским существительным *cou* (m), *coût* (m), *coucou* (m) и глагольной форме *coud*(*s*) инфинитива *coudre*[7]. Особенностью этого слова является то, что оно является отглагольным существительным, выражающим в первичном значении резкое, быстрое действие, направленное на что-л. или на кого-л. *Coup (m)* относится в семантическом плане к периферии существительных, поскольку выражает характерные для глагола значения (в семантике этого слова содержится значение быстрого, резкого действия).

Сочетаемость слов [17; 18] проходит на синтагматическом уровне [12; 14. С. 35–45] и определяется их семантическим потенциалом, их так называемой семантической валентностью [8. С. 79-80]. Ввиду того, что *coup* (m) может выражать очень широкое, обобщенное действие, движение, в процессе коммуникации смысл передаваемого значения потребовал своего уточнения, дифференциации, что привело к использованию на синтагматическом уровне сочетания прилагательных, причастий, предлога de + существительное с указанной лексической единицей [16].

Как следует из анализа эмпирического материала, *coup* (m) является одним из продуктивных элементов коллокации французского языка. Нами были выявлены сочетаемостные возможности существительного *coup* (m) и рассмотрены синтаксические конструкции, в структуре которых эта лексическая единица использована. Рассмотрим случаи употребления словосочетания *coup* (m) с детерминативом — прилагательным, причастием, числительным.

Словосочетания без предлогов структурно могут быть представлены в следующем виде: Coup + Adj.; Adj. + Coup; Coup + P.P.

Эти словосочетания могут выражать как первичное значение *coup (m)*, так и вторичное, переносное [23]. Существуют словосочетания, выражающие значение одномоментно в первичном и вторичном значении.

а) Словосочетания с элементом *coup* (m), используемые в своем первичном значении: *coup sec* ‘короткий, отрывистый удар’; *grand coup* ‘решающий

[5] Coup (m) // Petit Robert: Dictionnaire alphabétique & anal ogique de la langue française par Paul Robert. A. Rey et J. Rey-Debove (Réd.). P. : Société du nouveau littré. Le Robert, 1981. P. 405.

[6] *Ожегов С.И., Шведова Н.Ю.* Толковый словарь русского языка. Российская АН Ин-т рус. яз., Российский фонд культуры. М. : Азъ Ltd., 1992. С. 855.

[7] Le Bescherelle 2 L’Art de l’Orthographe. Les 26 pièges de l’orthographe. Lexique de 2000 homonymes. Dictionnaire orthographique. P. : Hatier, 1980. P. 45.

удар’; *coups déloyaux* ‘удары ниже пояса, запрещенные удары’. Например: *Il a donné un coup sec à la tête de son adversaire.*

b) Словосочетания с элементом *coup* (m), используемые в переносном значении: *un gros coup* ‘*разг.* тяжкий удар’; *coup bas* ‘удар ниже пояса, низость, подлость’; *mauvais (sale) coup* ‘неприятность; предательский удар, подвох, подлость’. Например: *Son sourire malicieux signifiait qu’il avait préparé un mauvais coup.*

c) Словосочетания с элементом *coup* (m) используются как в прямом, так и переносном значении: *coup dur* ‘потрясающий удар; досадная неприятность, напасть, серьезная помеха’; *coup fourré* ‘одновременные уколы двух противников (в фехтовании); удар в спину, подвох’; *le coup dans le dos* ‘удар в спину’. Например: *Il avait dû prévoir de longue date le coup dur...* (P. Vialar. *Le Temps des imposteurs*).

В ряде словосочетаний данное существительное может полностью изменить свою семантику и выступать в отличном от изначального значения: *coup fumant* ‘*разг.* потрясное, классное дело’; *coup monté* ‘трюк, инсценировка’; *coup pourri* ‘*разг*. неудачное, пустое дело, дохлый номер’; *un coup réussi* ‘*ирон.* дело выгорело, удалось дело’. Например: *Il était réjoui, venant de réussir ce coup fumant* (Y. Gibeau. *Et la fête continue*)[8].

Словосочетания с предлогами структурно могут быть представлены следующим образом:

a) Coup + prép.+ Nom. : *coup dans l'aile* ‘*арго* (*coup dans l'aile [или dans le pif]*) смертельный удар’; *(avoir) coup dans le nez* ‘под хмельком, под градусом’; *coup dans le seau* ‘*разг.* неудавшаяся проделка, промашка’.

b) Coup + prép.+ Adj.: *coup à blanc* ‘холостой выстрел’.

c) Coup + prép.+ Adv.: *coup pour rien* ‘безуспешная попытка, неудавшееся предприятие, пустой номер’.

d) Adj.num. + Coup; Adj.num. + de + Coup; Adj.num. + Coup+les Adj. : *cent coups* ‘разгульный образ жизни, распутство’; *deux coups de cuillère à pot* ‘в два счета, мигом’; *deux coups les gros* ‘сразу, запросто’.

e) Adj. + Coup; Adj. + Coup + de + Nom : *cousu de coups* ‘избитый, покрытый синяками’; *joli coup de sabre* ‘*разг*. большой рот, рот до ушей’.

f) Adj. Dém + Coup+Adv. : *ce coup-ci* ‘в этот или на этот раз’.

Например: *Il avait bien profité de la vie autrefois, disat-on, puis ayant fait les quatre cents coups et ayant mangé tous ses héritages, il s’était marié et retiré* (P. Gamarra. *Rosalle Brousse*).

Как видим, существительное *coup* (m) обладает обширными сочетательными и коллокационными способностями. Оно может выражать разные субстанции, способно комбинироваться с разными частями речи и в предложении может выполнять разные синтаксические функции.

[8] *Гринева Е.Ф., Громова Т.Н.* Словарь разговорной лексики французского языка: (уа материале современ. худож. лит. и прессы). М. : Рус. яз., 1988. С. 175.

Как показывает анализ фактов языка, очень продуктивной является конструкция, имеющая структуру Coup + de + Nom. Значительная номенклатура этого типа словосочетания объясняется тем, что она позволяет выразить резкое действие / движение в соединении с орудием и т.д. По модели Coup + de + Nom французский язык, особенно его разговорный регистр, растиражировал многочисленную группировку указанной конструкции.

Это подтверждает мысль А. Фрея о том, что всякий языковой факт стремится к созданию других фактов, которые могут войти с ним в одну систему, и к объединению с ними. Потребность в ассимиляции — это лингвистическая форма инстинкта имитации, который является мощным фактором социальной жизни … В таком смысле можно говорить о «силе имитации, неотделимой от самой системы» (Systemzwang) [24. С. 34].

Коллокация Coup + de + Nom представляет широкую палитру значений, содержащих сему некоего действия или состояния, события/происшествия, возникших в результате какого-л. действия как субъекта разумного, одушевленного, так и субъекта природного, одушевленного-неодушевленного. Кроме того, выражаемые действия могут быть как однонаправленными: от субъекта к объекту, так и взаимонаправленными [25]. *Coup* (m) во французском языке выступает определенной платформой для образования сложно оформленных лексических единиц.

Лексемы с субстантивом *coup* (m) в своем первичном значении

Они представляют конкретные (ударные) действия субъекта.

1. Ударными элементами субъекта выступают:

a) Части тела (конечности человека): *coup de boule* ‘*прост.* удар головой в грудь или живот’; *un coup de main* ‘жест, движение руки’; *coup de pied* ‘пинок, удар ногой’; *coup de poing* ‘удар кулаком’; *coup de revers* ‘удар наотмашь’; *coup de botte* ‘пинок ногой’; *coup de tampon* ‘*разг.* тумак’; *coup de targette* ‘*арго* пинок’. Например: *D’un coup de main, il rebroussa la mèche qui lui barrait le front* (R. Martin du Gard. *Les Thibault*).

b) Инструменты: *coup de balai* ‘взмах метлы’; ‘*разг.* чистка, увольнение, сокращение’; *coup de bâton* ‘удар палкой’; *coup de bec* ‘клевок, удар клювом’; *coup de cachet* ‘*вор. жарг.* (смертельный) удар ножом в сердце, в живот т.д.’. Например: *Un coup de fouet claque, une poignée de sable et de terre retombe sur le capitaine* (J. Fréville. *Plein vent*)[9].

Коллокации, имеющие в своем составе существительное *coup* (m) с расширенной семантикой

Лексемы с *coup* (m) во вторичных значениях выражают различные действия одушевленного и неодушевленного субъектов, их физическое,

[9] Французско-русский фразеологический словарь (ФРФС) / сост. В.Г. Гак, И.А. Кумина, И.П. Лалаев [и др.]; под ред. Я.И. Рецкера. М., 1963.

коммуникативное и визуальное взаимодействия, события, явления, выражают моральную оценку, оценку успешности действий, физическое и психическое состояние субъекта, отношение субъекта к действительности, состояние субъекта, природные явления и т.д. Так, лексемы с ядерным константным элементом *coup (m)* могут выражать следующие значения:

1. Обобщенно представленные действия: *coups de bêche* ‘поиски, усилия’; *coup d’arrêt* ‘отпор’. Например: *Tous les coups de bêche des forces réactionnaires étaient vains.*

2. События. Происшествия. Коллокации могут содержать следующие коннотации:

a) нейтральная коннотация — *coup de politique* ‘политический акт, политическая акция’;

b) положительная коннотация — *coup de Bourse* ‘удачная биржевая сделка’; *coup de ventre* ‘*разг.* смелое предприятие’;

c) отрицательная коннотация — *coup d’arraché* ‘*арго* крупное ограбление, большая кража’; *coup de manche* ‘*вор. жарг.* домушничество’; *coup de raccroc* ‘*вор. жарг*. крупная добыча’; *coup de vague* ‘*арго.* карманная кража’.

Например: *On ne vole plus, on fait des coups de Bourse et on monte des sociétés par actions* (C. Farrère. *L’homme qui assassina*. 2015).

3. Поступок субъекта: *coup de caboche* ‘*прост.* неожиданная выходка, необдуманный поступок’; *coup de culot* ‘*разг.* рискованное дело, отчаянный шаг’; *coup de poker* ‘рискованная попытка, рискованный шаг; хитрая уловка’. Например: *Ce coup de caboche de Michel a étonné ses parents et tous ses amis.*

4. Физическое, психологическое состояние субъекта: состояние субъекта может быть результатом действий природного, техногенного и антропоморфного характера: *coup de bambou* ‘*разг.* 1) солнечный удар 2) внезапный приступ гнева

3) внезапный упадок сил’; *coup de colère* ‘приступ, порыв гнева’; *coup de désespoir* ‘приступ отчаяния’; *coup de folie* ‘приступ безумия’; *coup de fatigue* ‘приступ усталости’; *coup de sang* ‘1) прилив крови 2) приступ гнева, ярости’. Например: *Les oreilles du pharmacien lui tintaient à croire qu’il allait tomber d’un coup de sang* (G. Flaubert. *Madame Bovary*).

5. Употребление алкогольных напитков: *coup d’arrosoir* ‘*разг.* стаканчик вина’; *coup de chasselas* ‘*арго* выпивка; легкое опьянение’; *coup de l’étrier* ‘посошок’; *coup de médecin* ‘полстакана вина после первого блюда’; *coup du milieu* ‘*разг.* рюмка, выпитая за едой’. Например: *Après avoir bu le coup de l’étrier, tous les deux enfourchèrent leur bidet* (A. Theuriet. *La Chanoinesse*).

6. Коммуникация: *coup de casserole* ‘*прост.* донос’; *coup de chapeau* ‘приветствие, поклон; поспешный кивок’, ‘*разг.* знак восхищения’; *coup de cloche* ‘предостережение, предупреждение’; *coup de dent* ‘колкая шутка, колкость, злословие’; *coup de griffe* ‘колкая шутка, колкость’; *coup de plume* ‘укол, упрек в письме или печати’; *coup d’épingle* ‘*перен.* шпилька’; *coup de langue* ‘колкая шутка, злословие’; *coup de gueule* ‘крик, брань, галдеж, *прост.*

клевета’. Например: *Elle vengea d’un coup de langue Athanase et le cher chevalier de Valois* (H. de Balzac. *La vieille fille*).

7. Визуальный контакт: *coup d’oeil* ‘взгляд’; *coup de châsse* ‘*разг.* подмигивание, завлекательный взгляд’; *coup de périscope* ‘*прост.* (настороженный) взгляд’; *coup de sabord* ‘*арг.* (*coup de sabord [или de saveur]*) проницательный, испытующий взгляд’. Например: *Oh! le coup d’oeil terrible de la mère à l’enfant!* (A. Daudet. *Jack*).

8. Оценка действий субъекта: *coup d’éclat* ‘мастерской удар, удача, блестящий ход’; *coup de tête* ‘*уст.* мастерски, искусно, успешно проведенное дело’; ‘отчаянный, смелый поступок’; *coup de génie* ‘гениальный ход’; *coup de maître* ‘мастерской ход; успешно проведенное дело’. Например: *Leur liaison datait de ce coup de maître* (J. Cocteau. *Le grand écart*).

9. Профессиональное умение: *coup de crayon* ‘искусство рисования’; *coup de pinceau* ‘поспешный мазок кистью’; *coup de fion* ‘отделка, внешний лоск’; *coup de rabot* ‘окончательная отделка, последний штрих’. Например: *Ce vieux peintre a fait quelques coups de pinceau.*

10. Судьба. Удача. Неудача: *coup de bol* ‘*разг.* счастливый случай, везение, удача’; *coup de bonheur* ‘счастливый случай, удача’; *coup de cafard* ‘*разг.* промашка с отчаяния’; *coup de cul* ‘*разг.* удача, везение’; *coup de hasard* ‘чистая случайность, игра судьбы, удача’; *coup de pot* ‘*прост.* удача, везение’; *coup d’épée* ‘бесплодная попытка’; *coup de la fortune* ‘превратности судьбы’; *coup du sort* ‘случайность, каприз судьбы’; *coup de la grâce* ‘провидение, вмешательство’; *coup de malheur* ‘несчастный случай, несчастье, беда’; *coup de partie* ‘решающий шанс’. Например: *Il y avait eu là un de ces concours de circonstances extraordinaires qui font les grands coups du hasard* (E. Zola. *L’argent*).

11. Насилие. Столкновение субъектов: *coup de chien* ‘*разг.* драка, заваруха’; *coup de filet* ‘*разг.* облава’; *coup de force* ‘акт насилия; переворот; заговор’; *coup d'État* ‘государственный переворот, дело, совершаемое в интересах государства’; *coup d'autorité* ‘насильственные меры властей’; *coup de serviette* ‘*арго* облава’; *coup de raclette* ‘*арго.* облава’. Например: *Spinelle qui avait échappé au coup de filet du 13 avril devait venir au cours de la nuit avec des armes...* (S. de Beauvoir. *Tous les hommes sont mortels*).

12. Помощь. Содействие: *coup de main* (или *d’épaule, de coude, de pouce*) ‘помощь, поддержка подмога,’ *coup de piston* ‘*разг.* тайная помощь, протекция и т.д.’. Например: *Si tu as besoin de quoi que ce soit, tu fais signe, il ne faut pas te gêner. Je peux ... te donner un coup d’épaule* (G. Bayle. *Le Pompiste et le chauffeur*).

13. Действие неодушевленного субъекта: *coup d’aile* ‘взмах крыльев’; ‘*ав.* беспосадочный перелет’; ***d’un coup d’aile*** ‘одним махом’.

14. Звучание оружия и предметов быта: *coup de feu* (*de fusil, de pistolet*) ‘выстрел’; *coup de téléphone* (или *de fil*, *разг.*, *арго.* de tube, *арго* coup de treuil) ‘телефонный звонок’; *coup de partance* ‘прощальный салют корабля’.

Например: *Je suis bien sûr que les Allemands sont nos ennemis, car j'ai bien souvent fait le coup de pistolet contre eux* (P. Mérimée. *Chronique du règne de Charles IX*).

15. Предательство. Обман. Несоблюдение моральных норм и правил: *le coup de canapé* 'подстроенная мужчине встреча с молодой девицей, с целью его на ней женить'; *coup de canif* '*разг.* нарушение, несоблюдение, супружеская измена'; *coup de Jarnac* 'нечестный, вероломный поступок'; *coup du père François* '*разг.* предательский удар; нападение с целью ограбления'. Например: *Elle est propre, ta façon d'aimer. à coup de canif dans le contrat...* (A. Theuriet. *Charme dangereux*).

16. Явления природы: *coup de foudre* 'удар молнии, (неожиданность; любовь с первого взгляда)'; *coup de froid* 'резкое похолодание, переохлаждение'; *coup de mer* 'удар волны, сильный шторм'; *coup de soleil* 'солнечный луч, солнечный удар'; *coup de tonnerre* 'раскат_грома'; *coup de vent* 'порыв ветра, шквал'. Например: *Ce fut un coup de foudre pour Jean-Claude* (Erckmann-Chatrian. *L'invasion*).

Заключение

Процесс образования и функционирования коллокаций в языке, в том числе и во французском с субстантивом *coup* (m) или какой-либо другой лексической единицей — процесс перманентный. В своей работе мы рассмотрели функциональную динамику *coup* (m), его семантическую палитру, а также общие тенденции и механизм образования коллокаций французского языка с ядерным элементом *coup* (m), их структурное разнообразие, стилистическую принадлежность.

Как показало проведенное исследование, существительное *coup* (m) наделено значительным коллокационным и функциональным потенциалом. Элемент *coup* (m) обладает обширными сочетательными способностями. Оно может выражать разные субстанции, способно комбинироваться с разными частями речи и в предложении может выполнять разные синтаксические функции. *Coup* (m), употребляясь в сочетании с другими лексическими единицами, образует фразеологизированные и нефразеологизированные коллокации.

Субстантив *coup (m)* помимо своего первичного значения 'удар, толчок, пинок' и т.д. выражает также такие понятия, как 'действие, поступок, звонок, выстрел, дело, удача, неудача, глоток вина, рюмочка, попытка, раз' и т.д., вследствие чего он становится их языковым субститутом.

Семантика элемента *coup* (m) определяется в зависимости от синтагматического окружения. Определение смысла коллокации также зависит от контекста или коммуникативной ситуации, в которой использовались та или иная лексема с компонентом *coup* (m). Довольно высокая частотность употребления *coup* (m) отмечается в разных регистрах французского языка и речи. Семантический сдвиг, который ведет это существительное к удалению от своего

первоначального значения и даже к полному изменению семантики, зависит от значения лексической единицы, с которой *coup* (m) взаимодействует в структуре словосочетания, семантический сдвиг происходит также в результате транспозиции коллокации из одного стилистического уровня на другой стилистический уровень.

Таким образом, существительное *coup* (m) реализует в выявленных устойчивых сочетаниях — французских коллокациях свою ведущую структурно-семантическую роль. *Coup* (m) во французском языке выступает определенной платформой — базовым элементом образования сложно оформленных лексических единиц. Французские словосочетания с *coup* (m) объективируют свои лексико-грамматические и коммуникативные функции в устной и письменной речи.

Список литературы

1. *Zipf G.K.* Human Behavior and the Principle of Least Effort. Cambridge: Addi-son-Wesley, 1949.
2. *Архангельский В.Л.* Устойчивые фразы в современном русском языке. Ростов-на-Дону, 1964.
3. *Аликова С.В.* Об истории колокативных исследований // Международный журнал гуманитарных и естественных наук. Филологические науки. № 10–1 (61). С. 133–136.
4. *Крыштaлева В.Е.* К вопросу об определении понятия «коллокация» в отечественной и западной лингвистических традициях // Вестник Новосибирского государственного университета. Серия: История, филология. 2016. Т. 15. № 9: Филология. С. 34–41. EDN: XHOWHL
5. *Захаров В.П., Хохлова М.В.* Анализ эффективности статистических методов выявления коллокаций в текстах на русском языке // Компьютерная лингвистика и интеллектуальные технологии. 2010. № 9 (16). С. 137–143. EDN: WKKORV
6. *Коршунова А.В*. Окказиональные коллокации в типологическом и функционально-семантическом аспектах (на материале произведений русских, английских и французских поэтов эпохи романтизма): автореф. дис. … канд. филол. наук. Кемерово, 2018. EDN: OBCYCT
7. *Лощакова Н.П.* Понятие «Коллокация» в истории традиционной и корпусной лингвистики // Гуманитарные исследования. 2021. № 3 (79) С. 34–37.
8. *Палийчук Д.А.* Проблема определения понятия коллокация в современнй лингвистике // Евразийский гуманитарный журнал. 2022. № 1. С. 20–25. EDN: FNXNKD
9. *Sinclair J.* Corpus, Concordance, Collocation. Oxford : Oxford University Press, 1991.
10. *Tutin A., Grossmann F.* Collocations régulières et irrégulières: esquisse de typologie du phénomène collocatif // Revue Française de Linguistique Appliquée. 2002 [Электронный ресурс] Режим доступа: https://hal.science/hal-01336278/document (дата обращения 18.08.2025).
11. *Богоявленская Ю.В., Палийчук Д.А.* Подходы к исследованию коллокаций во французской корпусной лингвистике // Вестник Пермского национального исследовательского политехнического университета. Проблемы языкознания и педагогики. 2025. № 1. С. 19–28. https://doi.org/10.15593/2224-9389/2025.1.2 EDN: KUZTME
12. *Гак В.Г.* Языковые преобразования. М. : Studia philologica, 1998. EDN: XRIAFL
13. *Иорданская Л.Н., Мельчук И.А.* Смысл и сочетаемость в словаре / под ред. Е.Н. Саввиной. М. : Языки славянских культур, 2007. EDN: SULWVR
14. *Влавацкая М.В.* Синтагматика vs. комбинаторика: основы комбинаторной лингвистики // Научный диалог. 2017. № 1. С. 35–45. EDN: XRIAFL

15. *Архипова Н.Г.* Сочетаемость слова в лексикографическом описании: дисс. ... канд. филол. наук. М., 2000. EDN: NLRDFZ
16. *Евдокимов А.П.* Лексическая сочетаемость и проблемы переводной лексикографии (на материале немецких соответствий сочетаниям лексем типа «прилагательное + абстрактное существительное» русского языка): дисс. ... канд. филол. наук. Алма-Ата, 1981. EDN: NPCNSX
17. *Котелова Н.З.* Значение слова и его сочетаемость (к формализации в языкознании) / АН СССР. Ин-т рус. яз. Ин-т языкознания. Ленинград: Наука. Ленингр. отд-ние, 1975.
18. *Телия В.Н.* Семантический аспект сочетаемости слов и фразеологическая сочетаемость // Принципы и методы семантических исследований. М. : Наука, 1976. С. 244–267.
19. *Телия В.Н.* Семантика связанных значений слов и их сочетаемости // Аспекты семантических исследований. М. : Наука, 1980. С. 250–319.
20. *Влавацкая М.В.* Комбинаторная семасиология (семантика и сочетаемость слов) // Мир науки, культуры, образования. 2009. № 7. С. 29–34. EDN: LJWMXH
21. *Лапаева Т.А.* Устойчивые сочетания в системе языковых единиц // Вестник Новгородского государственного университета им. Ярослава Мудрого. 2007. № 44. С. 75–77.
22. *Некрасова О.А.* Лингвистические коллокации английского языка и факторы, влияющие на процесс их образования // Вестник Московского государственного областного университета. Серия: Лингвистика. 2017. № 5. С. 120–132. https://doi.org/10.18384/2310-712X-2017-5-120-132 EDN: XHAZFE
23. *Соссюр Ф. де.* Труды по языкознанию. М. : Прогресс, 1977.
24. *Фрей А.* Грамматика ошибок: пер. с фр./ вступ. ст. В.М. Алпатова. М. : КомКнига, 2006.
25. *Червоный А.М.* Функциональная динамика категории «языковой субъект» (на материале французского языка). Монография. Ростов-на-Дону: издательство Южного федерального университета, 2014. EDN: TOEVFB

References

1. Zipf, G.K. (1949). *Human Behavior and the Principle of Least Effort.* Cambridge: Addi-son-Wesley.
2. Arkhangelsky, V.L. (1964). *Set Phrases in Modern Russian*. Rostov-on-Don. (In Russ.).
3. Alikova, S.V. (2021). On the History of Colocative Studies. *International Journal of Humanities and Natural Sciences. Philological Sciences*, 10-1 (61), 133–136. (In Russ.).
4. Kryshtaleva, V.E. (2016). On the Definition of the Concept of “Collocation” in Domestic and Western Linguistic Traditions. *Questioning the Definition of the Notion “Collocation” in Russian and Occidental Linguistic Traditions. Vestnik NSU. Series: History and Philology, 15*(9): Philology, 34–41. (In Russ.). EDN: XHOWHL
5. Zakharov, V.P., & Khokhlova, M.V. (2010). Study of Effectiveness of Statistical Measures for Collocation Extraction on Russian Texts**.** *Computational Linguistics and Intellectual Technologies*, *9*(16), 137–143. (In Russ.). EDN: WKKORV
6. Korshunova, A.V. (2018). *Occasional Collocations in Typological and Functional-Semantic Aspects (Based on the Works of Russian, English and French Poets of the Romantic Era*) [PhD Thesis]. Kemerovo. (In Russ.). EDN: OBCYCT
7. Loshchakova, N.P. (2021). The Concept of Collocation in The History of Traditional and Corpus Linguistics. *Humanitarian Studia*, 3(79), 34–37. (In Russ.).
8. Paliychuk, D.A. (2022). The Problem of Defining “Collocation” in Modern Linguistics. *Eurasian Humanitarian Journal*, 1, 20–25. (In Russ.). EDN: FNXNKD
9. Sinclair, J. (1991). *Corpus, Concordance, Collocation*. Oxford: Oxford University Press.
10. Tutin, A., & Grossmann, F. (2002). Collocations régulières et irrégulières: esquisse de typologie du phénomène collocatif. *Revue Française de Linguistique Appliquée*. URL: https://hal.science/hal-01336278/document (accessed: 18.08.2025).

11. Bogoyavlenskaya, Yu.V., & Paliychuk, D.A. (2025). Approaches to the Study of Collocations in French Corpus Linguistics. *PNRPU Linguistics and Pedagogy Bulletin*, 1, 19–28. (In Russ.). https://doi.org/10.15593/2224-9389/2025.1.2 EDN: KUZTME
12. Gak, V.G. (1998). *Language Transformations*. Moscow: Studia philologica. (In Russ.).
13. Iordanskaya, L.N., & Melchuk, I.A. (2007). *Meaning and Combinability in the Dictionary*, E.N. Savvina (Ed.). Moscow: Languages of Slavic Cultures. (In Russ.). EDN: SULWVR
14. Vlavatskaya, M.V. (2017). Sintagmatics Vs. Combinatorics: Basic Principles of Combinatorial Linguistics. *Nauchnyi Dialog*, 1, 35–45. (In Russ.). EDN: XRIAFL
15. Arkhipova, N.G. (2000). *Word Collocation in Lexicographic Description* [PhD Thesis]. Moscow. (In Russ.). EDN: NLRDFZ
16. Evdokimov, A.P. (1981). *Lexical Compatibility and Problems of Translation Lexicography (Based on German Correspondences to Combinations of Lexemes of the Type "Adjective + Abstract Noun" of The Russian Language)* [PhD Thesis]. Alma-Ata. (In Russ.). EDN: NPCNSX
17. Kotelova, N.Z. (1975). *The Meaning of the Word and its Combinability (towards Formalization in Linguistics*). Leningrad: Nauka. Leningrad Branch. (In Russ.).
18. Teliya, V.N. (1976). Semantic Aspect of Word Compatibility and Phraseological Compatibility. In: *Principles and Methods of Semantic Research* (pp. 244–267). Moscow: Nauka. (In Russ.).
19. Teliya, V.N. (1980). Semantics of Related Meanings of Words and Their Compatibility. In: *Aspects of Semantic Research* (pp. 250–319). Moscow: Nauka. (In Russ.).
20. Vlavatskaya, M.V. (2009). Combinatorial Semasiology (Semantics and Collocability of Words). *Mir Nauki, Kultury, Obrazovaniya,* 7, 29–34. (In Russ.). EDN: LJWMXH
21. Lapaeva, T.A. (2007). Stable Combinations in the System of Linguistic Units. *Vestnik of the Novgorod State University*, 44, 75–77. (In Russ.).
22. Nekrasova, O.A. (2017). Factors Influencing the Process of English Linguistic Collocations. B*ulletin of Moscow Region State University. Series: Linguistics*, 5, 120–132. (In Russ.). https://doi.org/10.18384/2310-712X-2017-5-120-132 EDN: XHAZFE
23. Saussure, F. de. (1977). *Works on Linguistics*. Moscow: Progress. (In Russ.).
24. Frey, A. (2006). *Grammar of Errors*. Moscow: KomKniga. (In Russ.).
25. Chervony, A.M. (2014). *Functional Dynamics of the Category "Language Subject" (Based on the French Language).* Rostov-on-Don: Southern Federal University Publ. (In Russ.). EDN: TOEVFB

Сведения об авторах:

Червоный Александр Михайлович, доктор филологических наук, доцент, профессор кафедры немецкого и французского языков, Таганрогский институт имени А.П. Чехова (филиал) РГЭУ (РИНХ) (347900, Российская Федерация, г. Таганрог, ул. Петровская, д. 68); *сфера научных интересов*: теория языка, сравнительная типология, семантический синтаксис, словообразование, фразеология; *e-mail*: ckutrik@yandex.ru

ORCID: 0000-0002-4587-0195; SPIN-код: 1350-1560; AuthorID: 471853; ResearcherID: A-7307-2017.

Арутюнова Жанна Михайловна, кандидат филологических наук, доцент, доцент кафедры иностранных языков факультета гуманитарных и социальных наук, Российский университет дружбы народов (117198, Российская Федерация, г. Москва, ул. Миклухо-Маклая, д. 6); *сфера научных интересов*: лексикология и фразеология французского языка; e-mail: arutyunova-zhm@rudn.ru

SPIN-код: 4448-6745; AuthorID: 566258.

Нелюбова Наталия Юрьевна, доктор филологических наук, доцент, доцент кафедры иностранных языков филологического факультета, Российский университет дружбы народов (117198, Российская Федерация, г. Москва, ул. Миклухо-Маклая, д. 6); *сфера научных интересов*: теория языка, сопоставительная паремиология, лингвоаксиология; *e-mail*: neliubova-nyu@rudn.ru
ORCID: 0000-0002-6538-8267; SPIN-код: 9799-0555; AuthorID: 394192.

Information about the authors:

Alexander M. Chervony, Dr. Sc. (Philology), Associate Professor, Professor of the Department of German and French Languages, Taganrog Institute named after A.P. Chekhov (branch) of Rostov State University of Economics (68 Petrovskaya str., Taganrog, Russian Federation, 347900); *Research interests*: theory of language, comparative typology, semantic syntax, word formation, phraseology; *e-mail*: ckutrik@yandex.ru
ORCID: 0000-0002-4587-0195; SPIN-code: 1350-1560; AuthorID: 471853; ResearcherID: A-7307-2017.
Zhanna M. Arutyunova, Ph.D. in Philology, Associate Professor, Foreign Languages Department, Faculty of Humanities and Social Sciences, RUDN University (6 Miklukho-Maklaya Str., Moscow, Russian Federation, 117198); *Research interests*: French lexicology and phraseology; *e-mail*: arutyunova-zhm@rudn.ru
SPIN-code: 4448-6745; AuthorID: 566258.
Natalia Yu. Neliubova, Dr. Sc. (Philology), Associate Professor, Associate Professor, Department of Foreign Languages, Faculty of Philology, RUDN University (6 Miklukho-Maklaya Str., Moscow, Russian Federation, 117198); *Research interests*: theory of language, comparative paremiology, linguo-axiology; *e-mail*: neliubova-nyu@rudn.ru
ORCID: 0000-0002-6538-8267; SPIN-code: 9799-0555; AuthorID: 394192.

RUDN Journal of Language Studies, Semiotics and Semantics
2313-2299 (Print), ISSN 2411-1236 (Online)
Вестник Российского университета дружбы народов.
Серия: ТЕОРИЯ ЯЗЫКА. СЕМИОТИКА. СЕМАНТИКА

2026 Vol. 17 No. 2 743–756
http://journals.rudn.ru/semiotics-semantics

DOI: 10.22363/2313-2299-2026-17-2-743-756
EDN: MKYVLZ

Научная статья / Research Article

Концепт DIEU ‘БОГ’ как ценностная доминанта французской лингвокультуры

В.Н. Денисенко[1], Л.А. Сафаралиева[1]✉, Е.М. Недопекина[1], Чжан Кэ[2]

[1]Российский университет дружбы народов, *Москва, Российская Федерация*
[2]Восточный педагогический университет, *Чанчунь, Китайская Народная Республика*
✉kuznetsova-la@rudn.ru

Аннотация. Вера в Бога, безусловно, является одной из общечеловеческих ценностей, поразному осмысляющейся носителями национальных языков, что обуславливает различия в понятийных, образных и ценностных признаках концепта БОГ в тех или иных лингвокультурах. *Целью* настоящего исследования является реконструкция понятийных и ценностных признаков французского концепта DIEU ‘БОГ’ как ценностной доминанты культуры. *Материалом* исследования послужили толкования вышеупомянутой лексемы, взятые из нормативных толковых словарей французского языка, а также 85 устойчивых выражений, отобранных методом сплошной выборки из фразеологических и паремиологических словарей французского языка. Основными *методами* исследования стали: описательный, метод компонентного анализа, лингвокультурологический и лингвоаксиологический. Было обнаружено, что в понятийных признаках концепта отражаются три «ипостаси» Бога: Бог — сверхъестественное существо, Бог — человек и Бог — ценностная сущность окружающего мира. Ценностные установки, содержащиеся в паремиях с компонентом *dieu,* отражают, в большей степени, религиозное (католическое) восприятие Бога и отражают четкие установки поведения для носителей французского языка: *следует помогать нуждающимся, делиться материальными благами, тогда Бог вознаградит еще большим богатством; следует знать, что только Бог управляет всеми процессами в мире, человеку это неподвластно; следует верить в Бога, чтобы получить его помощь и поддержку в делах* и др.

Ключевые слова: семантика, аксиология, концепт, лингвокультурология, французский язык

Вклад авторов: В.Н. Денисенко — концепция исследования, методология исследования, написание части текста; Л.А. Сафаралиева — методология исследования, сбор и обработка материала, анализ полученных данных, написание части текста, оформление статьи; Е.М. Недопекина — сбор и обработка материала, анализ полученных данных, написание части текста; Чжан Кэ — сбор и обработка материала, анализ полученных данных, написание части текста. В.Н. Денисенко, Л.А. Сафаралиева, Е.М. Недепекина и Чжан Кэ – подготовка текста статьи.

Конфликт интересов: авторы заявили об отсутствии конфликтов интересов.

© Денисенко В.Н., Сафаралиева Л.А., Недопекина Е.М., Чжан Кэ, 2026
This work is licensed under a Creative Commons Attribution-NonCommercial 4.0 International License
https://creativecommons.org/licenses/by-nc/4.0/legalcode

Финансирование: Статья подготовлена в рамках реализации инициативной темы НИР № 050226-0-000 «Системный взгляд как основа филологической мысли», Российский университет дружбы народов имени Патриса Лумумбы (РУДН), Российская Федерация.

История статьи: дата поступления: 20.01.2026; дата приема в печать: 10.02.2026.

Для цитирования: *Денисенко В.Н., Сафаралиева Л.А., Недопекина Е.М., Чжан Кэ.* Концепт DIEU ‘БОГ’ как ценностная доминанта французской лингвокультуры // Вестник Российского университета дружбы народов. Серия: Теория языка. Семиотика. Семантика. 2026. Т. 17. № 2. С. 743–756. https://doi.org/10.22363/2313-2299-2026-17-2-743-756 EDN MKYVLZ

The Concept of DIEU ‘GOD’ as the Dominant Value of French Linguoculture

Vladimir N. Denisenko[1], Lyubov A. Safaralieva[1]✉, Ekaterina M. Nedopekina[1], Zhang Ke[2]

RUDN University, *Moscow, Russian Federation*
Northeast Normal University, *Changchun, The People's Republic of China*
✉kuznetsova-la@rudn.ru

Abstract. Faith in God is, of course, one of the universal human values, interpreted in different ways by native speakers of national languages, which causes differences in the conceptual, figurative and value features of the concept of GOD in various linguistic cultures. The purpose of this study is to reconstruct the conceptual and value features of the French concept of DIEU ‘GOD’ as the value dominant of culture. The study material was the interpretation of the above-mentioned lexeme taken from standard explanatory dictionaries of the French language, as well as 85 set expressions selected using the continuous sampling method from phraseological and paremiological dictionaries of the French language. The main research methods are descriptive, the method of component analysis, linguocultural and linguoaxological. It was found that the conceptual features of the concept reflect three “hypostases” of God: God is a supernatural being, God is a man, and God is a valuable object of the surrounding world. The values contained in the parodies with the dieu component reflect, to a greater extent, the religious (Catholic) perception of God and reflect clear behavioral guidelines for native French speakers: *one should help those in need, share material goods, then God will reward with even greater wealth; one should know that only God controls all processes in the world, man this is beyond control; one must believe in God in order to receive his help and support in business*, etc.

Keywords: semantics, axiology, concept, linguoculturology, French language, dominant cultural values

Authors’ contribution: Vladimir N. Denisenko — study conception, research methodology, writing part of the text; Lyubov A. Safaralieva — material collection and processing, data analysis, writing part of the text, article design; Ekaterina M. Nedopekina — material collection and processing, data analysis, writing part of the text; Zhang Ke — material collection and processing, data analysis, writing part of the text.Vladimir N. Denisenko, Lyubov A. Safaralieva, Ekaterina M. Nedopekina, Zhang Ke — preparation of the article text.

Conflict of interests: The authors stated that there were no conflicts of interest.

Funding: The article was prepared as part of the implementation of the initiative topic of Research No. 050226-0-000 “A systematic view as the basis of philological thought”, Patrice Lumumba Peoples’ Friendship University of Russia (RUDN University), Russian Federation.

Article history: Received: 20.01.2026; Accepted: 10.02.2026.

For citation: Denisenko, V.N., Safaralieva, L.A., Nedopekina, E.M., & Zhang, Ke (2026). The concept of DIEU 'GOD' as the dominant value of French linguoculture. *RUDN Journal of Language Studies, Semiotics and Semantics, 17*(2), 743–756. (In Russ.). https://doi.org/10.22363/2313-2299-2026-17-2-743-756 EDN MKYVLZ

Введение

Антропоцентрическая парадигма современной лингвистики обуславливает актуальность исследований, направленных на изучение языка во взаимосвязи с сознанием (когнитивная лингвистика) и культурой (лингвокультурология). По мнению Н.Н. Болдырева, «… контекст культуры является наиболее значимым, поскольку мысль о том, что познание мира человеком происходит в рамках определенной культуры не вызывает сомнений. Этим обусловлено особое внимание, уделяемое в современных исследованиях когнитивного направления, изучению взаимосвязи языка и сознания в широком культурном и социально-историческом контексте» [1. С. 690]. В связи с этим реконструирование фрагментов языковой картины мира того или иного народа — носителя языка, вербализующих особые ментальные образования — концепты, является одной из приоритетных задач современной лингвистической науки.

В условиях многополярного мира и укрепления межгосударственных отношений, роста межкультурной коммуникации актуальным становится исследование *аксиологического сознания* или «части языкового (и неязыкового) сознания, анализ которого особенно необходим в современных условиях все возрастающего общения носителей неидентичных языковых сознаний» [2. С. 102]. Безусловно, осмысление и описание ценностей культуры, как общечеловеческих[1], так и национальных, вербализуемых языковыми единицами, позволяет выявить сходства и различия в ценностных ориентирах поведения носителей различных языков. По мнению В.И. Карасика, выявление таких ориентиров или стереотипов поведения должно основываться на описании ценностных признаков лингвокультурных концептов, представляющих собой ц*енностные доминанты*[2] культуры или «доминанты поведения, свойственные представителям того или иного социума» [4. С. 62]. В.И. Карасик полагает, что «ценности осознаются, обусловлены эмоциями и *детерми-*

[1] Прим.: «Общечеловеческие ценности — это некоторые сущности, изучаемые в философии, политологии, психологии, культурологи и языкознании <…> это содержание, формируемое отдельным членом общества, при помощи которого он ориентируется относительно успешности/неуспешности своей целенаправленной деятельности, согласует правила достижения этих целей с ограничениями, существующими в конкретном обществе» [3. С. 62].

[2] Прим.: value dominants — «термин российского лингвиста В.И. Карасика (1996) — наиболее важные и существенные смыслы данной культуры, имеющие этническую специфичность» Цит. по: *Жукова И.Н.* Словарь терминов межкультурной коммуникации / под ред. М.Г. Лебедько, З.Г. Прошиной. М. : Флинта ; Наука, 2013.

нируют поступки <…> эмоциональная обусловленность ценностей базируется на переживаемой системе культурных приоритетов, а поступки вписываются в стереотипные действия, принятые в соответствующем обществе» [4. С. 60–61].

Следует отметить, что имеется значительное количество исследований, посвященных осмыслению аксиологического потенциала языковых единиц, выявлению и классификации ценностей национальной культуры, вербализуемых паремиологическими единицами языка: труды В.И. Карасика [5–7], О.В. Ломакиной [8; 9], Н.Ю. Нелюбовой [10; 11], В.Н. Денисенко [12; 13], Л.А. Сафаралиевой [14; 15] и др.

Настоящее исследование посвящено реконструкции понятийных и ценностных признаков концепта DIEU ‘БОГ’, репрезентирующего ценностную доминанту французской лингвокультуры и представляющего значимый фрагмент языковой картины мира носителей французского языка.

Ранее концепт БОГ уже становился предметом исследования лингвистов: Л.Н. Гишкаева, О.В. Ломакина и А.С. Макарова проанализировали веру в Бога как ценностную константу христианской и мусульманской лингвокультур [16], Е.В. Милетова проанализировала вербализацию концепта БОГ на материале английских религиозных текстов [17], В.М. Мокиенко осмыслил дихотимию «Бог-дьявол» в русинской паремиологии [18], М.Ч. Кремшокалова исследовала ценностный концепт БОГ в жанровой картине мира благожеланий и проклятий [19] и др. В настоящей статье впервые описываются предписания и установки поведения носителей французского языка, эксплицитно выражаемые паремиологическими единицами французского языка с компонентом-лексемой *dieu* ‘Бог’. По мнению многих исследований, именно пословичный фонд «…может помочь выявить картину в том числе современных национальных ценностных приоритетов… Аккумулируя представления об искомом характере отношений, пословицы остаются одним из важнейших критериев оценки» [20. С. 215].

Обсуждение и результаты

Выявление понятийных признаков концепта DIEU ‘БОГ’

Для выявления понятийных характеристик концепта DIEU ‘БОГ’ необходимо обратиться к толкованию лексемы *dieu*, отраженному в словарях французского языка.

В словаре *Littré* представлены следующие два значения лексемы *dieu*:

1. «Nom du principe, unique ou multiple, qui, dans toutes les religions, est placé au-dessus de la nature ‘название единственного или множественного принципа, который во всех религиях ставится выше природы’;

2. L’être infini créateur et conservateur du monde dans la religion chrétienne, et aussi dans le mahométisme, dans le judaïsme, et parmi ceux qu’on nomme deists

‘Бесконечное существо, являющееся создателем и хранителем мира в христианской религии, а также в исламе, иудаизме и среди тех, кого называют деистами’» [3].

Важно отметить, что во втором значении данное слово используется как имя собственное, пишется с заглавной буквы и употребляется без артикля.

В онлайн-версии словаря *LaRousse* обнаруживаем следующие определения лексемы *dieu*:

1. «Dans les religions polythéistes, être supérieur doué d’un pouvoir surnaturel sur les hommes; divinité ‘В политеистических религиях — высшее существо, наделенное сверхъестественной властью над людьми; божество’[4];

2. Représentation d’une divinité ‘Изображение или воплощение божества’;

3. Personne à laquelle on voue une sorte de vénération, un attachement passionné, ou que l’on considère comme supérieure ‘Человек, к которому испытывают благоговение, сильную привязанность или которого считают выше себя’;

4. Objet, chose, idée que l’on place au plus haut dans la hiérarchie des valeurs et à quoi on sacrifie tout ‘Объект, вещь или идея, занимающие вершину иерархии ценностей, которым все приносится в жертву’»[5].

Наконец, в третьем наиболее авторитетном толковом словаре французского языка *Le Robert* приводятся следующие толкования лексемы *dieu*:

1. «Être éternel, unique, créateur et juge ‘в монотеизме: Вечное, уникальное существо, творец и судья’

2. Être supérieur doué d’un pouvoir sur l’être humain et d’attributs particuliers ‘в политеизме: Высшее существо, наделенное особыми качествами и властью над людьми’»[6].

Таким образом, можно выделить следующие понятийные признаки концепта DIEU ‘БОГ’:

- ***высшее сверхъестественное существо;***
- ***создатель и хранитель мира;***
- ***существо, обладающее властью над людьми;***
- ***человек, которым восхищаются;***
- ***некая ценность, ради которой следует идти на жертвы.***

Интересно отметить, что в одном из толкований лексемы *dieu* в словаре *LaRousse* DIEU ‘БОГ’ отождествляется с человеком, имеющим способности,

[3] Dictionnaire Littré. Режим доступа: https://www.littre.org/definition/dieu (дата обращения: 20.01.2026).

[4] Прим.: В этом значении данная словарная единица может писаться как с заглавной, так и с прописной буквы; встречается в форме единственного и множественного числа, а также может использоваться в форме женского рода — богиня (*déesse*).

[5] LAROUSSE. Режим доступа: https://www.larousse.fr/dictionnaires/francais/dieu/25418 (дата обращения: 20.01.2026).

[6] Le Robert. Dico en ligne. Режим доступа: https://dictionnaire.lerobert.com/definition/dieu (дата обращения: 20.01.2026).

умения и таланты, которых нет у других людей. Причем, согласно этимологическому словарю *Orthodidacte*[7], во французском языке данное переносное значение стало актуализироваться только с XVII века. Первоначально в качестве имени исследуемого концепта носители французского языка использовали латинское слово *deus*, имевшее значение ‘бог, божество’, а переход к форме *dieu* произошел лишь после XII века. Лексема *deus* принадлежит к той же семье слов, что и *dies*, означающее ‘день’, которое встречается во французских названиях дней недели (*lun**di*** **‘**понедельник’, *mar**di*** ‘вторник’, *mercre**di*** ‘среда’ и т.д.), а также, предположительно, связана с греческим именем собственным *Zeus* ‘Зевс’.

Словари синонимов французского языка *Synonymes et antonymes français*[8] и *Le Robert*[9] предлагают следующую синонимы лексемы *dieu* (табл.).

Синонимы лексемы *dieu* / Synonyms of the *dieu* lexeme

Французская лексическая единица / French lexical unit	**Перевод на русский язык / Translation into Russian**
le Créateur	Создатель, Творец
le Grand Architecte	Великий Архитектор (Вселенной) — в массонстве
le Seigneur	Господин, Господь
l’Éternel	Вечный
le Tout-Puissant	Всемогущий
le Très-Haut	Всевышний
l’Être Suprême	Высшее Божество
le Sauveur	Спаситель
le **Père** éternel	Вечный **Отец**
esprit	Дух (Святой)
juge	Судья
maître	Владыка, Учитель
idole	идол
démiurge	демиург

Источник: составлено В.Н. Денисенко, Л.А. Сафаралиевой, Е.М. Недопекиной, Чжан Кэ.
Source: compiled by Vladimir N. Denisenko, Lyubov A. Safaralieva, Ekaterina M. Nedopekina, & Zhang Ke.

Анализ синонимов лексической единицы *dieu* позволяет выделить еще один понятийный признак реконструируемого концепта:

❖ ***Отец, учитель.***

Все вышеизложенное позволяет заключить, что в отличие от амбивалентного представления концепта БОГ в русской лингвокультуре [15],

[7] Orthodidacte Dictionnaire. Режим доступа: https://dictionnaire.orthodidacte.com/article/etymologie-dieu (дата обращения: 20.01.2026).
[8] Synonymes et antonymes français. Режим доступа: https://www.synonymes.com/synonyme.php?mot=dieu (дата обращения: 20.01.2026).
[9] Le Robert. Dico en ligne. Режим доступа: https://dictionnaire.lerobert.com/synonymes/dieu (дата обращения: 20.01.2026).

во французской лингвокультуре концепт DIEU ‘БОГ’ представлен тремя ипостасями (рис. 1).

В корпусе Национального центра текстовых и лексических ресурсов Франции[10] в качестве антонимов слова *dieu* приводятся лексемы *créature* ‘создание, существо’, *diable* ‘дьявол’, *homme* ‘человек’. Таким образом, DIEU ‘БОГ’ противопоставляется *всему созданному им*, в том числе человеку, а также *силам зла* как источник доброты и чистоты.

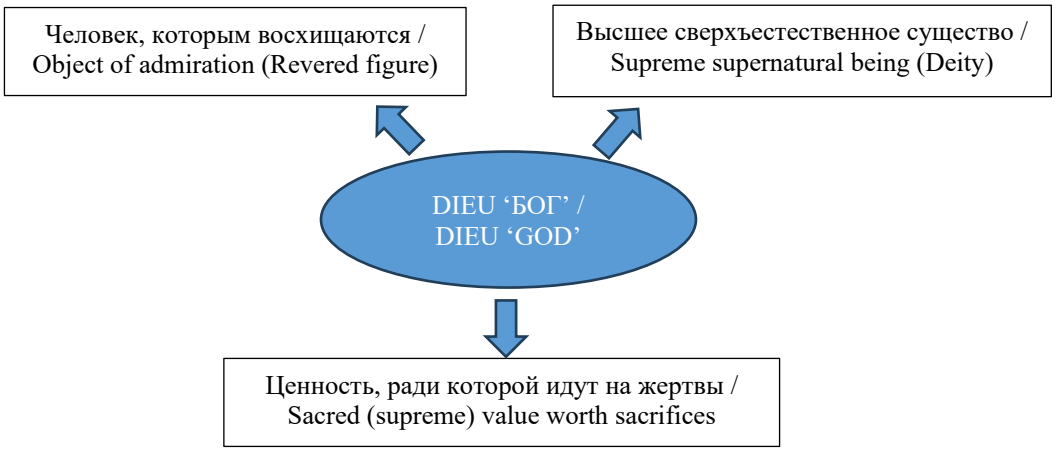


Рис. 1. Понятийное содержание концепта DIEU ‘БОГ’
Источник: составлено Л.А. Сафаралиевой.
Figure 1. The conceptual content of the DIEU concept ‘GOD’
Source: compiled by L.A. Safaralieva.

Антагонизм Бога и дьявола/чёрта выражается и на языковом уровне, о чем свидетельствует наличие во французском языке ряда устойчивых выражений: *devoir à Dieu et à diable* (досл.: ‘быть должным богу и чёрту — погрязнуть в долгах’); ne croire ni en (à) Dieu ni au diable (досл.: ‘не верить ни богу, ни чёрту — никому не верить’). Интересно, что в обоих выражениях слово *Dieu* пишется с заглавной буквы, как имя собственное, что подчеркивает, по всей вероятности, значимость Бога для носителей французского языка, а лексема *diable* ‘дьявол, чёрт’ записана со строчной буквы, что указывает на ничтожность дьявольских сил перед лицом Бога. Таким образом, можно выделить еще один понятийный признак концепта DIEU ‘БОГ’:

❖ ***Источник добра, противостоящий злым силам***.

Ценностные признаки концепта DIEU ‘БОГ’

Для реконструкции ценностных признаков концепта DIEU ‘БОГ’ были отобраны методом сплошной выборки фразеологические и паремиологические

[10] Centre National de Ressources Textuelles et Lexicales. Режим доступа: https://www.cnrtl.fr/antonymie/dieu (дата обращения: 20.01.2026).

единицы, содержащие лексему *dieu*. Всего было отобрано и проанализировано 85 устойчивых выражений.

Ценностная значимость концепта DIEU ‘БОГ’ переживается как всем говорящим на данном языке сообществом в целом, так и отдельными его индивидами. В зависимости от условий жизни, интересов, области приложений усилий, представлений о морали и норме каждого человека можно говорить о следующем ***спектре ценностей***, отражаемых французскими паремиологическими единицами с компонентом *dieu*:

– **ежедневные/бытовые** (забота о бытовом благополучии возлагается на Бога, но роль человека при этом тоже важна):

✓ *À qui se lève matin, Dieu prête la main досл.*: ‘Тому, кто встает утром, Господь подает руку — Под лежачий камень вода не течет’;

✓ *À toile ourdie Dieu envoie le fil* досл.: ‘Тому, кто ткет полотно, Бог посылает нить — Кто рано встает, тому Бог дает’;

✓ *Dieu donne le boeuf et non pas la corne* досл.: ‘Господь дает быка, а не рог — На Бога надейся, а сам не плошай’;

✓ *Aide-toi et Dieu t’aidera* досл.: ‘Помоги себе сам, и Бог тебе поможет — Сам трудись, да и Богу молись’;

✓ *Aux audacieux Dieu prête la main* досл.: ‘Бог протягивает руку смелым — На Бога надейся, а сам не плошай’;

✓ *Dieu est toujours pour les gros bataillons* досл.: ‘Бог всегда на стороне больших батальонов — Дорогу осилит идущий’;

– **моральные** (оценка поведения человека в обществе, соблюдение им нравственной чистоты и приверженности правде):

✓ *Dieu regarde les mains pures, non les mains pleines* ‘Бог смотрит на чистые руки, а не на руки, покрытые грязью’. Это высказывание передает мысль о том, что Бог благоволит людям с чистыми помыслами, совершающим добрые дела.

✓ *Dieu donne du bien aux hommes et non des hommes aux biens* ‘Бог дает людям добро, а не людей добру’. В данном случае под словом *добро* понимается ‘имущество, богатство’, таким образом делается акцент на порицании стяжательства и накопительства.

✓ *On peut voir le peu de cas que Dieu fait des richesses par les gens à qu’ il les donne* ‘Как мало люди ценят то, что им дано Богом’.

✓ *Qui donne aux pauvres, prête à Dieu* ‘Кто подает бедным, подает Богу’;

– **религиозные** (воздаяние славы Господу, забота о душе, отправляющейся в Царство Божье):

✓ *Pour la plus grande gloire de Dieu* ‘Во славу Божью’;

✓ *Paraître devant le bon Dieu* ‘Предстать перед Богом’;

✓ *Recommander son âme à Dieu* ‘Предать свою душу Богу’;

✓ *Se remettre entre les mains de Dieu* ‘Отдать себя в руки Божьи’;

– Бог и **пространство** его бытования:

✓ *C’est la maison du bon Dieu* ‘Это дом Божий’;

✓ *C’est la maison de Dieu où l’on ne boit ni ne mange* ‘Это дом Божий, в нем не пьют и не едят’.

Речь идет о месте нахождения Бога — храме, соборе, церкви, небе. Одним из синонимов лексемы *dieu* во французском языке является слово *Ciel* ‘Небо’: *Dieu, le Ciel nous préserve de quelque chose* ‘Бог / Небо нас бережет от чего-то’.

– Бог как **совесть** человека и/или его душа:

✓ *C’est le Coeur qui sent Dieu et non la raison* ‘Бога чувствует сердце, а не разум’;

✓ *En face de Dieu et des hommes* ‘Перед Богом и людьми’;

✓ *Entre Dieu et soi* ‘Между Богом и собой’;

– Бог как **властитель** или высшая сила:

✓ *L’homme propose et Dieu dispose* ‘Человек предполагает, а Бог располагает’;

✓ *La voix du peuple est la voix de Dieu* ‘Глас народа — глас Бога’;

✓ *Chacun pour soi, Dieu pour tous* ‘Каждый за себя, а Бог за всех’;

✓ *Un roi juste est l’ombre de Dieu sur la terre* ‘Честный царь — тень Бога на земле’;

✓ *Le Père des hommes et des dieus* ‘Отец людей и богов/святых’;

✓ *Il faut s’adresser à Dieu qu’à ses saintes* ‘Лучше обращаться к Богу, чем к святым’;

✓ *Il vaut mieux avoir affaire à Dieu qu’à ses saints* ‘Лучше иметь дело с Богом, чем с его святыми’.

В этих высказываниях отражается мысль о том, что Бог справедливо управляет миром и людьми, являясь наивысшей властью в мире, что именно он принимает самые важные решения.

– Бог — это **судьба**:

✓ *Dieu fait et défait le destin* ‘Бог создает и разрушает судьбу’;

✓ *L’homme propose et Dieu dispose* ‘Человек предполагает, а Бог располагает’;

✓ *Comme il plaît à Dieu* ‘как Богу будет угодно’;

✓ *Seul Dieu a donné la vie, seul il peut la retirer* ‘Бог жизнь дал, Бог ее и взял’;

✓ *La mort et la vie, Dieu en dispose* ‘Смертью и жизнью управляет Бог’;

– Бог как высшее **знание**:

✓ *Dieu le sait* ‘Только Богу известно’;

✓ *Dieu sait comme* ‘Бог знает как’;

✓ *Dieu sait que, si, combien* ‘Бог знает кто/сколько/ли’;

✓ *Dieu seul le sait* ‘Одному Богу ведомо’;

– Бог как **помощь** и поддержка:

✓ *L’homme s’agite mais Dieu le mène* ‘Человек суетится, но Бог его ведет’;

✓ *Que ceux qui vuelent un appui le cherche en Dieu* ‘Кто хочет опоры, ищет ее у Бога’;

✓ *Dieu vous assiste* ‘Бог в помощь’;

✓ *Avec l’aide de Dieu* ‘С Божьей помощью’;

✓ *Dieu veuille* ‘С Божьей помощью’.

В самые сложные минуты жизни человек обращается к Богу в молитве о помощи, поддержке и сотворении чуда.

– Бог как **любовь/милость/красота**:

✓ *Dieu châtie celui qu'il aime, comme un père châtie l'enfant qu'il cherit* 'Бог наказывает того, кого любит, подобно тому, как отец наказывает любимого ребенка';

✓ *Par la grâce de Dieu* 'По милости Божьей';

✓ *Dieu merci!* 'Слава Богу';

✓ *Beau comme un dieu* 'Красив как Бог'. Под любовью понимается также милость к людям, снисхождение к их слабостям;

– Бог как **наказание**:

✓ *La crainte du Dieu est le commencement de la sagesse* 'Боязнь Бога — начало мудрости';

✓ *On n'entendrait pas Dieu tonner sourd à n'entendre pas Dieu tonner* 'Мы бы не слышали грома Божьего, если бы были глухи к тому, что не слышим Его грома';

✓ *Acte de Dieu* 'Стихийное бедствие';

✓ *Du tonnerre (de Dieu)* 'Гром Божий';

✓ *Jugement de Dieu* 'Божий суд';

✓ *Le glaive de Dieu* 'Меч Бога';

– Бог как **прощение**:

✓ *Dieu ne veut pas la mort du pécheur* 'Бог не хочет смерти грешника';

✓ *Pardon de Dieu* 'Божье прощенье'.

Семантический и лингвокультурологический анализ приведенных выше паремиологических единиц позволяет выделить следующие ***ценностные установки*** поведения носителей французского языка:

❖ ***Бог помогает человеку в делах, но следует и самому прилагать усилия — нельзя сидеть сложа руки;***

❖ ***Следует помогать нуждающимся, делиться материальными благами, тогда Бог вознаградит еще большим богатством;***

❖ ***Человеку следует помнить, что он предстанет перед Богом, поэтому важно совершать добрые поступки;***

❖ ***Следует знать, что только Бог управляет всеми процессами в мире, человеку это неподвластно;***

❖ ***Бог любит каждого человека и прощает его грехи;***

❖ ***Следует верить в Бога, чтобы получить его помощь и поддержку в делах.***

Заключение

В результате проведенного анализа концепта DIEU 'БОГ' было обнаружено, что в восприятии образа Бога носителями французского языка в первую очередь, актуализируются следующие понятийные признаки: *высшее сверхъестественное существо*; *человек*, *которым восхищаются*; *некая ценность*,

ради которой следует идти на жертвы. При этом ценностные признаки концепта DIEU ‘БОГ’ верифицируют и усиливают следующие понятийные признаки: *существо*, *обладающее властью над людьми*; *источник добра*, *противостоящий злым силам.* Таким образом, в паремиологических единицах французского языка с компонентом *dieu* подчеркивается безграничная власть Бога, его доброе и чистое начало, то есть обнаруживается ярко выраженное влияние католической религии. Подчеркивается мысль о том, что в конце своего жизненного пути человек будет вознагражден Богом, но при условии, если будет вести достойную жизнь и трудиться. Представления о том, что человек или какой-то предмет окружающего мира могут приблизиться к Богу или стать им, в паремиологических единицах единичны.

Полученные в ходе проведенного исследования данные помогают реконструировать понятийные и ценностные признаки концепта DIEU ‘БОГ’ во французской лингвокультуре, что позволяет, с одной стороны, выявить национально окрашенные характеристики фрагмента французской языковой картины мира, а с другой стороны, использовать полученные данные в дальнейших исследованиях концепта DIEU ‘БОГ’ в сопоставлении с другими лингвокультурами.

Список литературы

1. *Болдырев Н.Н., Федяева Е.В.* Когнитивные методы исследования в лингвистике: концептуально-инференционный анализ // Вестник Российского университета дружбы народов. Серия: Теория языка. Семиотика. Семантика. 2023. Т. 14. № 3. С. 686–703. https://doi.org/10.22363/2313-2299-2023-14-3-686-703 EDN: LVSSMM
2. *Казакова Н.Ю.* Общечеловеческая ценность ДОМ в русском и немецком языковом сознании // Вестник Российского университета дружбы народов. Серия: Теория языка. Семиотика. Семантика. 2011. № 4. С. 102–110. EDN: PXZRWH
3. *Синячкин В.П.* Общечеловеческие ценности как инструмент анализа языкового сознания членов общества // Вестник Российского университета дружбы народов. Серия: Теория языка. Семиотика. Семантика. 2010. № 2. С. 62–66. EDN: MVIZXP
4. *Карасик В.И.* Аксиогенная ситуация как категория лингвокультурологии // Универсальное и национальное в языковой картине мира : сб. науч. ст. по итогам IV Междунар. науч. конф., Минск, 24–26 окт. 2019 г. : в 2 томах. Минск : МГЛУ, 2022. Т. 1. С. 59–67.
5. *Карасик В.И.* Щедрость как ценность русской лингвокультуры // Общая и русская лингвоаксикология. Ярославль : Канцлер, 2022. С. 14–39. EDN: CPIPWE
6. *Карасик В.И.* Лингвокультурные ценности в дискурсе // Иностранные языки в высшей школе. 2015. № 1 (32). С. 25–35. EDN: UASAOR
7. *Дигалев Н.В., Карасик В.И.* Лингвоаксиология спортивного состязания: ценности, нормы, традиции // Социальные и гуманитарные науки. Отечественная и зарубежная литература. Серия 6: Языкознание. 2025. № 1. С. 77–91. https://doi.org/10.31249/ling/2025.01.06 EDN: SANDJY
8. *Ломакина О.В.* Метаязык современной лингвоаксиологии: паремиологический уровень // Вестник Российского университета дружбы народов. Серия: Теория языка. Семиотика. Семантика. 2025. Т. 16. № 3. С. 622–637. https://doi.org/10.22363/2313-2299-2025-16-3-622-637 EDN: DQGURR
9. *Ломакина О.В.* Лингвоаксиологический анализ паремий как тип интегративного исследования (на примере ценностного концепта РОДИНА) // Когнитивные исследования языка. 2025. № 1–2 (62). С. 673–677. EDN: SOMMEK

10. *Нелюбова Н.Ю.* Представление ценностей в пословицах франкоязычных стран (на материале пословиц Мартиники) // Вестник университета (Российско-Таджикский (Славянский) университет). 2019. № 1 (65). С. 198–208. EDN: BWJTWG
11. *Денисенко В.Н., Ломакина О.В., Нелюбова Н.Ю.* Языковая маркированность аксиологических доминант франкоязычных паремий // Вопросы современной лингвистики. 2025. № 5. С. 19–30. https://doi.org/10.18384/2949-5075-2025-5-19-30 EDN: APIAGO
12. *Денисенко В.Н., Нелюбова Н.Ю.* Вербализация концепта *ЕДА* во французском паремиологическом фонде // Когнитивные исследования языка. 2024. № 1–2 (57). С. 31–34. EDN: APJPYR
13. *Сафаралиева Л.А., Денисенко В.Н.* Аксиологическое представление концепта МАТЬ в современной русской лингвокультуре // Вестник Российского университета дружбы народов. Серия: Теория языка. Семиотика. Семантика. 2025. Т. 16. № 1. С. 240–254. https://doi.org/10.22363/2313-2299-2025-16-1-240-254 EDN: LDPOGH
14. *Сафаралиева Л.А.* Аксиология старости в русской и испанской лингвокультурах: монография. СПб. : Наукоемкие технологии, 2025. EDN: QWKLNL
15. *Денисенко В.Н., Сафаралиева Л.А.* Реконструкция модели концепта «Бог» как ценностной доминанты русской культуры // Русистика. 2026. Т. 24. № 1. С. 138–153. https://doi.org/10.22363/2618-8163-2026-24-1-138-153 EDN: UWRENM
16. *Гишкаева Л.Н., Ломакина О.В., Макарова А.С.* Отражение веры в Бога как ценностной константы в пословичном фонде (на примере христианской и мусульманской лингвокультур) // Cuadernos de Rusística Española. 2021. Т. 17. С. 135–147. https://doi.org/10.30827/cre.v17.21112 EDN: OSOTVV
17. *Милетова Е.В.* Концепт «Бог» как ценностная доминанта религиозной коммуникации (на материале англоязычных проповедей) // Известия Российского государственного педагогического университета им. А.И. Герцена. 2024. № 213. С. 289–300. https://www.doi.org/10.33910/1992-6464-2024-213-289-300 EDN: LHQRMN
18. *Ломакина О.В., Мокиенко В.М.* Концептуальная дихотомия «Бог» — «дьявол» в русинской фразеологии и паремиологии (на славянском фоне) // Вестник Томского государственного университета. 2019. № 447. С. 55–62. EDN: PUZBYN
19. *Кремшокалова М.Ч.* Ценностный концепт «Бог» в жанровой картине мира благожеланий и проклятий // Филологические науки. Вопросы теории и практики. 2014. № 4–3 (34). С. 110–112. EDN: RYIAEL
20. *Семененко Н.Н.* Аксиология паремий в фокусе проблемы когнитивно-дискурсивного моделирования семантики русских пословиц // Вестник Российского университета дружбы народов. Серия: Теория языка. Семиотика. Семантика. 2020. Т. 11. № 2. С. 213–232. https://doi.org/10.22363/2313-2299-2020-11-2-213-232 EDN: WTIMBF

References

1. Boldyrev, N.N., & Fedyaeva, E.V. (2023). Cognitive Research Methods in Linguistics: Conceptual Inferential Analysis. *RUDN Journal of Language Studies, Semiotics and Semantics*, *14*(3), 686–703. (In Russ.). https://doi.org/10.22363/2313-2299-2023-14-3-686-703 EDN: LVSSMM
2. Kazakova, N.Y. (2011). Panhuman Value HOME/HOUSE in Russian and German Language Perception. *RUDN Journal of Language Studies, Semiotics and Semantics*, 4, 102–110. (In Russ.). EDN: PXZRWH
3. Sinyachkin, V.P. (2010). Panhuman Values as an Instrument to Analysis of the Language Perception by Members of the Society. *RUDN Journal of Language Studies, Semiotics and Semantics*, 2, 62–66. (In Russ.). EDN: MVIZXP
4. Karasik, V.I. (2022). Axiogenic Situation as a Category of Language and Culture Studies. In *Universal and national in the linguistic picture of the world: collection of scientific articles based on the results of the IV International Scientific Conference, Minsk, October 24–26, 2019:* in 2 vols. (Vol. 1. pp. 59–67). Minsk: MGLU Publ. (In Russ.).

5. Karasik, V.I. (2022). Generosity as a Value of Russian Linguocultures. In: *General and Russian Linguoaxiology* (pp. 14–39). Yaroslavl: Kantsler Publ. (In Russ.). EDN: CPIPWE
6. Karasik, V.I. (2015). Linguocultural Values in Discourse. *Foreign Languages in Tertiary Education*, 1 (32), 25–35. (In Russ.). EDN: UASAOR
7. Digalev, N.V., & Karasik, V.I. (2025). Linguistic Axiology of Sports Competition: Values, Norms, Traditions. *Social Sciences and Humanities. Domestic and Foreign Literature. Series 6. Linguistics*, 1, 77–91. (In Russ.). https://doi.org/10.31249/ling/2025.01.06 EDN: SANDJY
8. Lomakina, O.V. (2025). Metalanguage of Modern Linguaxiology: Paremiological Level. *RUDN Journal of Language Studies, Semiotics and Semantics, 16*(3), 622–637. (In Russ.). https://doi.org/10.22363/2313-2299-2025-16-3-622-637 EDN: DQGURR
9. Lomakina, O.V. (2025). Linguoaxiological Analysis of Paromias as a Type Integrative Research (On the Example of the Value Concept MOTHERLAND). *Cognitive Studies of Language*, 1–2 (62), 673–677. (In Russ.). EDN: SOMMEK
10. Nelyubova, N.Yu. (2019). Presentation of Values in the Proverbs of French-Speaking Countries (Based on Proverbs of Martinique). *University Bulletin (Russian-Tajik (Slavonic) University),* 1(65), 198–208. (In Russ.). EDN: BWJTWG
11. Denisenko, V.N., Lomakina, O.V., & Nelyubova, N.Yu. (2025). Linguistic Markedness of Axiological Dominants of French-Language Paroemias. *Key Issues of Contemporary Linguistics*, 5, 19–30. (In Russ.). https://doi.org/10.18384/2949-5075-2025-5-19-30 EDN: APIAGO
12. Denisenko, V.N., & Nelyubova, N.Yu. (2024). Verbalization of the *FOOD* Concept in the French Paremiological Fund. *Cognitive Studies of Language,* 1–2 (57), 31–34. (In Russ.). EDN: APJPYR
13. Safaralieva, L.A., & Denisenko, V.N. (2025). Axiological Representation of the Concept of MOTHER in Modern Russian Linguistic Culture. *RUDN Journal of Language Studies, Semiotics and Semantics*, *16*(1), 240–254. (In Russ.). https://doi.org/10.22363/2313-2299-2025-16-1-240-254 EDN: LDPOGH
14. Safaralieva, L.A. (2025). *The Axiology of Old Age in Russian and Spanish linguistic Cultures.* Saint Peterburg: Naukoyomkie tekhnologii publ. (In Russ.). EDN: QWKLNL
15. Denisenko, V.N., & Safaralieva, L.A. (2025). Reconstructed Conceptual Model 'God' as a Dominant Value of Russian Culture. *Russian Language Studies*, *24*(1), 138—153. (In Russ.). https://doi.org/10.22363/2618-8163-2026-24-1-138-153 EDN: UWRENM
16. Gishkaeva, L.N., Lomakina, O.V., & Makarova, A.S. (2021). Refection of Faith in God as a Value Constant in the State Fund (on the example of Christian and Muslim Linguistic Cultures). *Cuadernos de Rusística Española*, 17, 135–147. (In Russ.). https://doi.org/10.30827/cre.v17.21112 EDN: OSOTVV
17. Miletova, E.V. (2024). The Concept ‘GOD’ as a Value Dominant of Religious Communication: the Case of English Sermons*. Izvestia: Herzen University Journal of Humanities & Sciences*, 213, 289–300. (In Russ.). https://doi.org/10.33910/1992-6464-2024-213-289-300 EDN: LHQRMN
18. Lomakina, O.V. & Mokienko, V.M. (2019). The Conceptual Dichotomy “God” – “Devil” in Russian Phraseology and Paroemiology (Against the Slavic Background). *Tomsk State University Journal*, 447, 55–62. (In Russ.). https://doi.org/10.17223/15617793/447/7 EDN: PUZBYN
19. Kremshokalova, M.Ch. (2014). Value Concept “God” in Genre World View of Blessings and Damnations. *Philology. Theory & Practice*, 4–3 (34), 110–112. (In Russ.). EDN: RYIAEL
20. Semenenko, N.N. (2020). Axiology Proverbs in the Focus of the Problem of Cognitive-Discursive Modeling of the Semantics of Russian Proverbs. *RUDN Journal of Language Studies, Semiotics and Semantics*, *11*(2), 213–232. (In Russ.). https://doi.org/10.22363/2313-2299-2020-11-2-213-232 EDN: WTIMBF

Сведения об авторах:
Денисенко Владимир Никифорович, доктор филологических наук, профессор, заведующий кафедрой общего и русского языкознания филологического факультета, Российский университет дружбы народов (117198, Российская Федерация, г. Москва, ул. Миклухо-Маклая, д. 6); *сфера научных интересов*: разработка проблем семантики, фонетики, морфонологии и лексикологии современного русского языка, а также проблем общего и русского языкознания; *e-mail*: denisenko-vn@rudn.ru
ORCID: 0000-0001-6021-4068; SPIN-код: 4635-8342; Author ID: 310131; Scopus ID: 57193133860, AAG-9635-2020.
Сафаралиева Любовь Александровна, кандидат филологических наук, доцент, доцент кафедры общего и русского языкознания филологического факультета, Российский университет дружбы народов (117198, Российская Федерация, г. Москва, ул. Миклухо-Маклая, д. 6); *сфера научных интересов*: когнитивная лингвистика, ассоциативная лингвистика, лингвокультурология; *e-mail*: kuznetsova-la@rudn.ru
ORCID: 0000-0002-6960-9426; SPIN-код: 1089-5456, AuthorID: 1155571.
Недопекина Екатерина Михайловна, кандидат филологических наук, доцент кафедры общего и русского языкознания филологического факультета, Российский университет дружбы народов (117198, Российская Федерация, г. Москва, ул. Миклухо-Маклая, д. 6); *сфера научных интересов*: языковые контакты, лингвокультурология, переводоведение; *e-mail*: nedopekina-em@rudn.ru
ORCID: 0000-0002-5922-3669; SPIN-код: 2249-1912; AuthorID: 718637.
Чжан Кэ, кандидат филологических наук, старший преподаватель кафедры русского языка, Северо-Восточный педагогический университет (130024, Китайская Народная Республика, г. Чанчунь, ул. Жэньминь, д. 5268); *сфера научных интересов*: лексикология, семантика, лингвокультурология; *e-mail*: zhangk176@nenu.edu.cn
ORCID: 0000-0002-7765-2133; ResearcherID: D-7768-2019.

Information about the authors:
Vladimir N. Denisenko, Dr. Sc. (Philology), Professor, Head of the of the General and Russian Linguistics Department, RUDN University (6 Miklukho-Maklaya St., Moscow, Russian Federation, 117198); *Research interests*: semantics, phonetics, morphology and lexicology of the modern Russian language, problems of General and Russian linguistics; *e-mail*: denisenko-vn@rudn.ru
ORCID: 0000-0001-6021-4068; SPIN-code: 4635-8342; Author ID: 310131; Scopus ID: 57193133860, AAG-9635-2020.
Lyubov A. Safaralieva, PhD in Philology, Associate Professor, Associate Professor of the Department of General and Russian Linguistics, Faculty of Philology, RUDN University (6 Miklukho-Maklaya St., Moscow, Russian Federation, 117198); *Research interests*: cognitive linguistics, associative linguistics, linguoculturology; *e-mail*: kuznetsova-la@rudn.ru
ORCID: 0000-0002-6960-9426; SPIN-code: 1089-5456, AuthorID: 1155571.
Ekaterina M. Nedopekina, PhD in Philology, Associate Professor of the Department of General and Russian Linguistics, Faculty of Philology, RUDN University (6 Miklukho-Maklaya St., Moscow, Russian Federation, 117198); *Research interests*: language contacts, linguoculturology, translation studies; *e-mail*: nedopekina-em@rudn.ru
ORCID: 0000-0002-5922-3669; SPIN-code: 2249-1912; AuthorID: 718637.
Zhang Ke, PhD in Philology, Senior Teacher of the Department of Russian Language, Northeast Normal University (5268 Renmin Street, Changchun, The People's Republic of China, 130024); *Research interests*: lexicology, semantics, linguiculturology; *e-mail*: zhangk176@nenu.edu.cn
ORCID: 0000-0002-7765-2133; ResearcherID: D-7768-2019.

RUDN Journal of Language Studies, Semiotics and Semantics
2313-2299 (Print), ISSN 2411-1236 (Online)
Вестник Российского университета дружбы народов.
Серия: ТЕОРИЯ ЯЗЫКА. СЕМИОТИКА. СЕМАНТИКА

2026 Vol. 17 No. 2 757–777
http://journals.rudn.ru/semiotics-semantics

DOI: 10.22363/2313-2299-2026-17-2-757-777
EDN: MOUAEQ

Научная статья / Research Article

Модель этнокультурной семантической универсалии в эдусемиотике на примере казахского языка

С.В. Мыскин[1], И.В. Вашунина[2], М.П. Кабакова[3], К.М. Муса[3]

[1]Московский городской педагогический университет, *Москва, Российская Федерация*
[2]Всероссийская академия внешней торговли Минэкономразвития РФ, *Москва, Российская Федерация*
[3]Казахский национальный университет им. Аль-Фараби, *Алматы, Республика Казахстан*
✉myskinsv@mgpu.ru

Аннотация. Актуальность исследования эдусемиотики и психолингвистики учебного текста обусловлена мультикультурностью современного высшего образования. Вступая в профессию, студенты из разных этнокультур вовлекаются в единый профильный дискурс, включающий многообразие полимодальных учебных текстов. При этом зачастую не учитываются этнические особенности смыслового восприятия учебных текстов студентов, что может приводить к поверхностному усвоению учебного материала. Цель исследования — проверка гипотезы о существовании особых инвариантных и вариативных речевых структур, актуализация которых средствами казахского языка влияет на смысловое восприятие учебного текста. Теоретический анализ проблемы восприятия учебного текста позволил авторам обозначить методологическую основу исследования облигаторных связей языковых единиц учебного текста с их психологическими коррелятами. Описанная проблемная задача легла в основу разработки исследовательской стратегии семиотических, психологических и психолингвистических закономерностей восприятия учебного текста. В статье представлены некоторые результаты пилотного эксперимента изучения психологических и психолингвистических особенностей смыслового восприятия учебного текста на примере казахской лингвокультуры. В ходе эксперимента проверялась гипотеза о существовании особых инвариантных и вариативных речевых структур, актуализация которых средствами казахского языка влияет на смысловое восприятие учебного текста. Проверка гипотезы осуществлялась с использованием комплексного методического инструментария: ассоциативного и психосемантического эксперимента. В эксперименте участвовали студенты первого курса гуманитарных направлений подготовки Казахского национального университета им. аль-Фараби. Результаты эксперимента позволили выявить общекультурные (инвариантные) и этноспецифические (вариативные) семантические универсалии, обусловливающие смысловое восприятие учебного текста. Далее представлена психологическая интерпретация выявленных экспериментальных фактов с опорой на ключевые характеристики психологии казахского этноса и определены знаки, указывающие на наличие сформулированных универсалий.

© Мыскин С.В., Вашунина И.В., Кабакова М.П., Муса К.М., 2026
This work is licensed under a Creative Commons Attribution-NonCommercial 4.0 International License
https://creativecommons.org/licenses/by-nc/4.0/legalcode

Ключевые слова: общекультурная семантическая универсалия, этнокультурная семантическая универсалия, восприятие учебного текста, психология этноса, лингвокультура казахского этноса, психосемантический и психолингвистический эксперимент

Вклад авторов: равнозначен на все этапах исследования и подготовки статьи к публикации.

Конфликт интересов: авторы заявили об отсутствии конфликтов интересов.

История статьи: дата поступления: 20.01.2026; дата приема в печать: 10.02.2026.

Для цитирования: *Мыскин С.В.*, *Вашунина И.В.*, *Кабакова М.П.*, *Муса К.М.* Модель этнокультурной семантической универсалии в эдусемиотике на примере казахского языка // Вестник Российского университета дружбы народов. Серия: Теория языка. Семиотика. Семантика. 2026. Т. 17. № 2. С. 757–777. https://doi.org/10.22363/2313-2299-2026-17-2-757-777 EDN MOUAEQ

The Model of Ethnocultural Semantic Universality in Edusemiotics on the Example of the Kazakh Language

Sergey V. Myskin[1] ✉, Irina V. Vashunina[2], Mayra P. Kabakova[3], Kundyz M. Musa[3]

[1]Moscow City University, *Moscow, Russian Fedetation*
[2]VAVT Department of the Ministry of Economic Development of the Russian Federation, *Moscow, Russian Federation*
[3]Al-Farabi Kazakh National University, *Almaty*, *the Republic of Kazakhstan*
✉myskinsv@mgpu.ru

Abstract. The relevance of research in edusemiotics and psycholinguistics of educational texts is determined by the multicultural nature of contemporary higher education. Entering their profession, students from different ethnocultural backgrounds become involved in a unified professional discourse that encompasses a variety of multimodal educational texts. However, the ethnically specific features of students' semantic perception of educational texts are often not taken into account, which may result in superficial comprehension of the learning material. The aim of this study is to test the hypothesis that there exist specific invariant and variant speech structures whose actualization through the medium of the Kazakh language affects the semantic perception of the educational text. Theoretical analysis of the problem of educational text perception allowed the authors to identify the methodological basis for the study of the obligatory relationships between the linguistic units of an educational text and their psychological correlates. The described problematic task formed the basis for the development of a research strategy for semiotic, psychological and psycholinguistic patterns of perception of the educational text. The article presents some results of a pilot experiment to study the psychological and psycholinguistic features of semantic perception of educational texts within Kazakh linguoculture. During the experiment, the hypothesis of the existence of special invariant and variable speech structures was tested, the actualization of which by means of the Kazakh language affects the semantic perception of the educational text. The hypothesis was tested using appropriate methodological tools: associative and psychosemantic experiments. The experiment involved first-year humanities students at Al-Farabi Kazakh National University. The results of the experiment made it possible to identify general cultural (invariant) and ethnospecific (variable) semantic universals that determine the semantic perception of the educational text. This is followed by a psychological interpretation of the experimental findings based on key characteristics of the psychology of the Kazakh ethnic group and signs have been identified indicating the presence of proposed universals.

Keywords: general cultural semantic universal, ethnocultural semantic universal, perception of educational text, ethnic psychology, Kazakh linguoculture, psychosemantic

Authors' contribution: The authors contributed equally to this article and preparation of the article text.

Conflict of interests: The authors stated that there were no conflicts of interest.

Article history: Received: 20.01.2026; Accepted: 10.02.2026.

For citation: Myskin, S.V., Vashunina, I.V., Kabakova, M.P., & Musa, K.M. (2026). The Model of Ethnocultural Semantic Universality in Edusemiotics on the Example of the Kazakh Language. *RUDN Journal of Language Studies, Semiotics and Semantics, 17*(2), 757–777. (In Russ.). https://doi.org/10.22363/2313-2299-2026-17-2-757-777 EDN MOUAEQ

Введение

Эдусемиотика — это новая научная дисциплина, подраздел семиотики, занимающаяся изучением знаковой репрезентации образовательной среды (преподавания, учебных материалов и интерпретации знаний обучающимися). Сам термин «edusemiotics» предложен Инной Семецки в 2017 г. как результат сложения английских слов «education» (образование) и «semiotics» (семиотика) [1], то есть он может переводиться как «образовательная семиотика». Исследования представителей этой научной дисциплины базируются на положениях, что основной задачей образования является передача знания, любое знание передается в знаковой форме (в виде текста, рисунка, графика, мультимедиа), а сам процесс обучения представляет собой непрерывный семиозис — интерпретацию новых знаков на базе старых с образованием новых интерпретант [2].

Семиотический подход к процессу обучения открывает новые перспективы исследований, нацеленных на оптимизацию представления знаний, за счет чего происходит, помимо прочего, улучшение коммуникации между учителем и учащимися, и все вместе приводит к более глубокому пониманию и успешному усвоению учебного материала обучающимися. В этой области существует ряд работ, однако на сегодняшний день не поднимался вопрос об этноспецифике восприятия учебного текста. В настоящей статье авторы обсуждают пилотную модель этнокультурной семантической универсалии.

Построение любой модели требует выделения как составных элементов объекта, так и их отношений, связей. Применительно к проблеме смыслового восприятия данное условие предстает в виде решения задачи выявления совокупности знаков учебного текста и их психологических коррелятов. Обозначенная проблемная задача и легла в основу разработки теоретической модели этнокультурной семантической универсалии. В связи с этим целесообразно раскрыть семиотический, психолингвистический и психологический теоретические подходы к разработке модели.

Сама модель является способом представления знания, а смысл процессов репрезентации рассматривается семиотикой [3. P. 345]. Также на сегодняшний день семиотика стала «обычной наукой» для анализа текстов [4. P. 149], что позволяет использовать семиотический подход к учебным

текстам. Одним из направлений эдусемиотики является анализ имеющихся учебных материалов, а также разработка новых учебников, тексты которых наиболее эффективно означивают знание, предлагаемое к усвоению. Под эффективностью здесь понимается прагматический потенциал учебных материалов, их «способность» создать максимально удобные условия для усвоения знания обучающимися. В ходе обучения применяется специфический подход к языку — прагматический, язык используется с единой общей целью передачи знаний — это универсалия любого образовательного процесса. Семиотика же — это подход к языку, который лучше всего соответствует прагматизму [5. P. 209].

Язык позволяет человеку устанавливать понятийные связи и отношения между объектами окружающей действительности, так происходит их осмысление. Интерпретация знаков учебного текста ведет к конструированию его смысла, результат же осмысления проявляется как в понимании текста, так и непонимании. И.А. Зимняя отмечает по этому поводу: «Непонимание — это не отсутствие осмысления, а лишь его отрицательный результат, свидетельствующий, что процесс не достиг исхода, адекватного ситуации общения» [6. С. 5–6]. В этом случае перед адресантом возникает задача создания такого учебного текста, который в ходе восприятия воссоздавался бы в сознании реципиента без смысловых потерь или искажений, и тем самым обеспечивал его понимание. Методологический аппарат лингвистики предлагает для этого учитывать морфологические и синтаксические особенности текста, а также его семантическую структуру, т.е. план выражения и план содержания языкового знака. С позиции общей семантики данные планы могут быть рассмотрены исходя из двух типов — референциального и грамматического. Референциальная семантика объемлет все, что связано с передачей свойств референтов, сигнификатов и денотатов [7. С. 43]. Грамматическая семантика включает порядок следования морфем и слов, номинативные предложения и перечислительные конструкции, т.е. описательный порядок слов [8. С. 145–146]. Также к этому типу относят семантическое единство частей текста, специфику графического оформления, изобразительные средства и др.

В психолингвистическом аспекте ключевую функцию в смысловом восприятии, по нашему мнению, выполняют прагматические характеристики, обусловливающие процесс интерпретации реципиентом речевого сообщения в контексте совместной деятельности и общения.

Таким образом, для нашего анализа можно выделить следующие значимые семантические параметры: *формальные* — предикативно раскрывающие свойства предмета (денотата) текста; *содержательные* — представленные структурным единством и смысловой соотнесенностью описания предмета (денотата) учебного текста; *прагматические* — указывающие на значимость включения предмета в совместную деятельность. Данные языковые параметры — маркеры учебного текста (именно так мы их и обозначим, подчеркивая их знаковую природу) совокупно характеризуют общекультурные семантические универсалии единые для всех профессиональных языков.

Далее рассмотрим параметры, значимые для изучения этнокультурной семантики учебного текста. Понимание текста воспринимающим заключается в толковании им «основной идеи получаемого сообщения в соответствии с замыслом коммуникатора» [9. С. 35]. Основная идея или смысл бытового/разговорного или художественного сообщения определяется паралингвистическими средствами, а также контекстом коммуникативной ситуации. Профессиональное же сообщение предполагает иной смысловой сценарий, поскольку реализуется в научно-учебном стиле. Научно-учебный стиль как «твердая» языковая система характеризуется высокой степенью устойчивости значений языковых единиц, отсутствием контекстной зависимости, закрытостью словаря, строгими алгоритмами генерирования текстов, низкой степенью субъективности и вероятностности при воспроизводстве и интерпретации [10]. Следует также учесть, что профессиональный язык представляет собой специальный тезаурус, усвоение которого возможно анализировать лишь на лексико-семантическом уровне, по поверхностному и глубинному планам слов. На поверхностном уровне адекватность усвоения термина оценивается по соответствию морфологических признаков различной модальности, а глубинный уровень — по пересечению или совпадению семантических полей слов. В психологии семантическое поле трактуется как комплекс ассоциаций, возникающих вокруг слова, которое является ядром этого поля [11]. Проблема семантической (ассоциативной) структуры слова или текста широко рассмотрена и в лингвистике (см. [12—14] и др.). Так, исследователи полагают, что поскольку семантическое поле представляет собой набор ключевых слов, то его можно отнести к разряду текстов. В свою очередь, Ю.А. Сорокин расценивает ассоциативную структуру как свернутое описание ситуации, которая разворачивается в тексте, созданном на ее основе [14]. Применительно к учебному тексту это означает, что «лишь пересекающиеся или совпадающие семантические поля терминов обладают текстообразующим потенциалом, способным передать смысл профессионального текста» [15. С. 43] и обеспечить конструирование целостного фрагмента профессиональной картины мира в сознании субъекта. Следовательно, понимание (усвоение) учебного текста возможно анализировать по воспроизведению комбинаций ключевых слов, обусловленных пересечением или совпадением их семантических полей. Для нашего исследования это выступает важнейшим параметром этнокультурной семантической универсалии.

Таким образом целью данной статьи выступает экспериментальный поиск и нахождение таких семиотических, психологических и психолингвистических закономерностей восприятия учебного текста, которые бы обеспечивали константность его содержания вне зависимости от меняющихся коммуникативных условий.

Описание эксперимента

Эксперимент направлен на выявление этнокультурных семантических универсалий смыслового восприятия учебного текста студентами вузов. Основной замысел эксперимента раскрывается в следующих положениях:

— эффективный учебный текст — это текст, который достигает своей учебной цели, передает необходимую информацию в доступной и понятной форме, а также вызывает у читателя нужные реакции или действия;

— общекультурные семантические универсалии зафиксированы в языковом сознании представителей профессиональной группы в виде параметров учебного текста;

— параметры учебного текста, функционирующие в сознании представителей профессиональной группы, обусловлены особенностями общения и деятельности, сформированными в ходе накопления ими профессиональных знаний и опыта;

— этнокультурное содержание сознания его носителей является аффективно окрашенным и выступает фактором вариативного конструирования параметров учебного текста в процессе его восприятия представителями данной этнической группы;

— этнокультурный компонент сознания проявляется при восприятии учебного текста любой модальной структуры, что указывает на его инвариантный и универсальный характер;

— этнокультурные особенности проявляются в знаковой форме, эти знаки могут быть выявлены и интерпретированы.

Замысел задает следующую структуру экспериментального исследования. *Цель:* выявление этнокультурных особенностей смыслового восприятия учебного текста. *Объект:* смысловое восприятие учебного текста этнической группой студентов вуза. *Предмет:* текстовые параметры, обусловливающие оптимальность смыслового восприятия учебного текста этнической группой студентов вуза. *Гипотеза (основная)*: оптимальность смыслового восприятия учебного текста обусловлена этнокультурными семантическими универсалиями. *Гипотеза (альтернативная)*: оптимальность смыслового восприятия учебного текста обусловлена лишь общекультурными семантическими универсалиями. *Методы эксперимента*: направленный ассоциативный эксперимент, психосемантический эксперимент, методы математической обработки данных, сравнительно-сопоставительный анализ [16], этнопсихологический анализ.

Эксперимент реализовывался в два этапа. На экспертном этапе определялась общекультурная семантическая универсалия, а на оценочном этапе — этнокультурная. При выборе стимульного языкового материала авторы руководствовались следующими теоретическими допущениями относительно учебного текста. Содержание учебного текста может быть охарактеризовано структурой денотатов — терминов с фиксированным понятийным значением.

При этом коннотации у специальных терминов отсутствуют. При исследовании восприятия учебного текста параметром лингвистической относительности можно пренебречь, поскольку языковая научная (профессиональная) картина мира считается универсальной. Однако смысловое восприятие происходит в результате усвоения и интерпретации не отдельных слов, а текстовых фрагментов и целого учебного текста. Свойства структуры и системы разных языков влияют на восприятие текста. Также были выделены объективные критерии выбора стимульного материала. Учебный текст должен отвечать следующим критериям: быть полимодальным, т.е. содержать вербальный и невербальный компоненты; быть непрофильным по отношению к направлению образования экспериментальной группы студентов, но при этом соответствовать смежному направлению; предметным в индивидуальном плане, т.е. иметь возможность без труда для испытуемых определить денотат. На основании обозначенных критериев для проведения эксперимента был выбран учебный текст по физиологии человека[1].

В соответствии с замыслом эксперимента данный стимульный материал был представлен в трех видах с разным соотношением вербальной и невербальной составляющих: текст с частичной креолизацией (включающий и текст и изображение), не креолизованный текст (включающий только языковую часть) и полностью креолизованный текст (включающий лишь изображение с подписью к нему). Предполагалось предъявление каждого вида стимульного материала разным подгруппам испытуемых, включенных в общую этнокультурную экспериментальную выборку. Таким образом, в каждой этнокультурной группе были выделены по три подгруппы: «Изображение + Текст», «Изображение» и «Текст». Стимульный материал для разных этнокультурных групп был представлен на их национальном языке.

Для экспериментального доказательства существования общекультурных и этнокультурных семантических универсалий необходимо выявить семиотическую характеристику (выделить объективный признак) исследуемого явления. Как отмечают П. Пиционе и М.Ф. Фреда, «психологические процессы основаны на непрерывном процессе семиотического опосредования» [17. P. 193].

Применительно к нашему эксперименту данные формулировки имеют следующее прикладное значение. Семантическая универсалия объективно представлена *константными сочетаниями* в массиве соответствующих данных. Константа — это переменная, сохраняющая свое относительное постоянство при изменении других переменных. В ассоциативном эксперименте это повторяющиеся слова-реакции на стимул у разных респондентов, в психосемантическом эксперименте — повторяющиеся сочетания шкал (парамет-

[1] Текст взят из: *Батуев А.С., Никитина И.П., Журавлев В.Л., Соколова Н.Н.* Малый практикум по физиологии человека и животных : уч. пособие / под ред. А.С. Батуева. СПб. : Изд-во С.-Петербургского ун-та, 2001.

ров оценки), включенных в один из факторов, при восприятии разного стимульного материала. Выявление отмеченных объективных признаков позволяет считать их знаками, интерпретантой которых и будет семантическая универсалия.

Исследование общекультурной семантической универсалии

В ходе экспертного этапа решалась задача определения общекультурных семантических универсалий посредством выделения значимых параметров учебного текста в направленном ассоциативном эксперименте. В группу экспертов вошли преподаватели и студенты вуза старших курсов по педагогическим, психологическим и психолого-педагогическим направлениям. Общее количество испытуемых составило 100 человек. Экспертам предлагалось дать множественную вербальную реакцию на стимул «Характеристики эффективного учебного текста». Собранный в ходе ассоциативного эксперимента языковой материал был обработан посредством частотного анализа. Затем были отобраны те характеристики, частота которых составляла 3 и выше. Опираясь на основные текстовые категории по И.Р. Гальперину, все параметры учебного текста были сгруппированы на формальные, содержательные и прагматические [18. С. 13] (табл. 1).

Таблица 1. **Общекультурные семантические универсалии смыслового восприятия учебного текста**
Table 1. **General Cultural Semantic Universals of Semantic Perception of an Educational Text**

Форма / Form	Содержание / Content	Прагматика / Pragmatics
богат лексическим материалом / rich in lexical material	актуальность информации / relevance of information	адекватность объема текста / adequacy of text volume
воспроизводимость / reproducibility	анализ разных точек зрения / relevance of information	интерактивность (наличие заданий, вопросов по теме (т.е. обратная связь) / interactivity (the presence of tasks, questions on the topic (i.e. feedback)
гармоничный / гармоничный	достоверность информации / reliability of information	носит развивающий характер / is of a developmental nature
грамотно изложенный / well-written	завершенность / completeness	обладающий воспитательными аспектами / possessing educational aspects
доступность изложения / accessibility of presentation	качественный текст / high-quality text	оказывает эмоциональное воздействие / has an emotional impact
использование наглядной опоры / use of visual support	конкретность / concreteness	полифункциональность / polyfunctionality

Окончнаие табл. 1 7 Ending of the Table 1

Форма / Form	**Содержание / Content**	**Прагматика / Pragmatics**
лаконичность / laconicism	креативность / creativity	получение новых знаний с опорой на фоновые знания / acquiring new knowledge based on background knowledge
наличие выделений в тексте (курсив, подчеркивание, жирным, цветом) / the presence of highlights in the text (italics, underlining, bold, color)	логика изложения / logic of presentation	смысловая ценность / semantic value
наличие заранее подготовленного плана / having a pre-prepared plan	методологичность изложения / methodological presentation	способствует обогащению речи обучающихся / contributes to the enrichment of students' speech
структурированность / structuredness	наличие интриги / presence of intrigue	учитывающий особенности аудитории / taking into account the characteristics of the audience
чередуется с интерактивными формами взаимодействия / alternates with interactive forms of interaction	наличие ключевых слов / presence of keywords	эффективность применения информации в жизни / the effectiveness of applying information in life
экспрессивность /	наличие примеров из практики / availability of practical examples	
эстетичность / aesthetics	новизна текста / novelty of the text	
	объективность изложения / novelty of the text	
	содержит проблемную информацию / contains problematic information	

Источник: исследование С.В. Мыскина, И.В. Вашуниной, М.П. Кабаковой и К.М. Муса. / *Source*: compiled by Sergey V. Myskin, Irina V. Vashunina, Mayra P. Kabakova, & Kundyz M. Musa.

Сгруппированные таким образом параметры учебного текста были использованы в качестве шкал в трех анкетах семантического дифференциала, разработанных для следующего экспериментального этапа.

Исследование этнокультурной семантической универсалии казахского этноса

В качестве экспериментальных групп оценочного этапа были выбраны студенты первых курсов гуманитарных вузов по психологическим и психолого-педагогическим направлениям подготовки. Поскольку эксперимент направлен на выявление этнокультурных особенностей смыслового

восприятия учебного текста, в выборку были включены две группы студентов из Московского городского педагогического университета (МГПУ) и Казахского национального университета им. аль-Фараби (КазНУ). Всего в эксперименте участвовали 150 студентов (75 русских и 75 казахов). Это в дальнейшем обеспечило проведение сравнительно-сопоставительного анализа полученных данных и выделение этнокультурных универсалий. Отметим, что в данной статье описаны результаты исследования, полученных *только* на группе казахских студентов.

На оценочном этапе различным студенческим группам предлагалось оценить учебный текст при помощи анкет семантического дифференциала. Данные протоколов оценки переводились в цифровую форму и обрабатывались при помощи факторного анализа. Выбор процедуры факторного анализа обусловлен одним из положений его концепции — измеряемые эмпирические показатели считаются следствием других, глубинных, скрытых от непосредственного измерения характеристик (латентных переменных) [19]. Для удобства презентации полученные данные представлены в обобщенной табличной форме. Первым шагом выступала оценка степени применимости факторного анализа к данным выборкам при помощи критерия адекватности выборки Кайзера-Мейера-Олкина (КМО). Несмотря на минимально допустимые размеры выборок, все значения превысили допустимый уровень > 0,5, что указывает на приемлемость данных по выборкам для обработки факторным анализом.

Следующая процедура обработки заключалась в составлении факторных матриц и поиск в их структуре константных сочетаний шкал, включенных в один из факторов. Для примера приведем сравнение трех факторных матриц, полученных посредством вращения «Варимакс», по трем подгруппам группы КазНУ по признаку «Форма» (рис. 1).

Изображение + Текст

	Компонент			
	1	2	3	4
Структурированность	,846			
Наличие подготовленного плана	,800			
Использование наглядной опоры	,726			
Текстовые выделения	,535	,513	,501	
Эстетичность		,758		
Интерактивные формы		,753		
Воспроизводимость		,648		
Экспрессивность	,590	,640		
Доступность изложения		,575	,530	
Гармоничный			,924	
Богат лексикой			,561	,537
Грамотно изложен				,807
Лаконичность				,615

Изображение

	Компонент			
	1	2	3	4
Грамотно изложен	,831			
Экспрессивность	,741			
Доступность изложения	,730			
Гармоничный	,689			
Воспроизводимость	,639			-,516
Интерактивные формы	,599		,536	
Наличие подготовленного плана		,770		
Богат лексикой		,711		
Использование наглядной опоры		,704		
Текстовые выделения		,598		,550
Лаконичность			,876	
Структурированность			,619	
Эстетичность				,774

Текст

	Компонент		
	1	2	3
Текстовые выделения	,887		
Наличие подготовленного плана	,855		
Воспроизводимость	,729		
Интерактивные формы	,632		
Использование наглядной опоры			
Гармоничный		,907	
Доступность изложения		,738	
Богат лексикой	,562	,650	
Эстетичность		,502	
Лаконичность			,826
Экспрессивность			,780
Грамотно изложен		,553	,614
Структурированность			,586

Рис. 1. Сравнительный анализ факторных матриц по трем подгруппам группы КазНУ по признаку «Форма»

Источник: исследование С.В. Мыскина, И.В. Вашуниной, М.П. Кабаковой и К.М. Муса.

Figure 1. Comparative analysis of factor matrices for three subgroups of the KazNU group on the basis of “Shape”

Source: compiled by Sergey V. Myskin, Irina V. Vashunina, Mayra P. Kabakova, & Kundyz M. Musa.

Из представленных данных можно выделить следующие константные сочетания шкал, объединенных в одном из факторов и повторяющихся во всех матрицах (табл. 2).

Сочетание шкал *текстовые выделения* и *наличие подготовленного плана* включены в первый фактор и актуальны при восприятии языкового материала без и вместе с визуальным изображением, а при восприятии только визуального изображения переходят во второй фактор. Пара шкал *воспроизводимость* и *интерактивные формы* включены в первый фактор при восприятии либо только языкового материала, либо только визуального изображения. В стимульном материале, сочетающем вербальную и невербальную части, эта пара включена во второй фактор.

Таблица 2. **Константные сочетания шкал в факторных структурах по трем подгруппам группы КазНУ по признаку «Форма»**
Table 2. **Constant Combinations of Scales in Factor Structures for Three Subgroups of the KazNU Group based on the "Shape" Attribute**

№	**Константные сочетания шкал / Constant combinations of scales**	**Изображение + Текст / Image + Text**	**Изображение / Image**	**Текст / Text**
1	Текстовые выделения / Text highlights Наличие подготовленного плана / Having a prepared plan	Фактор 1 / Factor 1	Фактор 2 / Factor 2	Фактор 1 / Factor 1
2	Воспроизводимость / Reproducibility Интерактивные формы / Interactive forms	Фактор 2 / Factor 2	Фактор 1 / Factor 1	Фактор 1 / Factor 1

Источник: исследование С.В. Мыскина, И.В. Вашуниной, М.П. Кабаковой и К.М. Муса.
Source: compiled by Sergey V. Myskin, Irina V. Vashunina, Mayra P. Kabakova, & Kundyz M. Musa.

В соответствии с обозначенным выше нами критерием константности сочетания «*текстовые выделения* и *наличие подготовленного плана», «воспроизводимость* и *интерактивные формы»* можно отнести к вербальным параметрам этнокультурной семантической универсалии казахской этнокультуры по признаку «план выражения». Анализ факторных матриц по другим признакам также позволил выделить константные сочетания шкал (рис. 2 и 3).

Как видно из рис. 2, план содержания содержит следующие константные сочетания шкал: *«логика изложения»*, *«креативность»* и *«методологичность»*; *«достоверность»* и *«конкретность»*; *«актуальность информации»* и *«новизна текста»*.

Анализ рис. 3 позволяет выделить следующие константные сочетания шкал по признаку прагматика: *«новые знания и новые фоновые знания»* и *«учет особенностей аудитории»*; *«адекватный объем текста»* и *«применение информации в жизни»*.

Изображение + Текст

	Компонент			
	1	2	3	4
Проблемная информация	,810			
Актуальность информации	,775			
Новизна текста	,764			
Примеры из практики	,704			
Разные точки зрения	,577			
Логика изложения		,785		
Креативность		,754		
Методологичность		,714		
Интрига		,636		
Ключевые слова			,799	
Завершенность			,794	
Объективность изложения			,716	
Конкретность				,762
Качественный текст				,748
Достоверность информации				,660

Изображение

	Компонент				
	1	2	3	4	5
Методологичность	,764				
Креативность	,705				
Качественный текст	,691				
Логика изложения	,689				
Интрига	,656				
Завершенность	,478				
Ключевые слова		,724			
Проблемная информ		,703			
Объективность изложения		,654			
Примеры из практики		,647			
Достоверность информации			,836		
Конкретность			,801		
Актуальность информации				,893	
Новизна текста				-,642	
Разные точки зрения					,922

Текст

	Компонент			
	1	2	3	4
Достоверность информации	,872			
Качественный текст	,849			
Завершенность	,803			
Конкретность	,768			
Ключевые слова	,640			
Логика изложения		,847		
Креативность		,802		
Методологичность		,791		
Объективность изложения		,692	,576	
Проблемная информация			,847	
Актуальность информации	,527		,728	
Разные точки зрения			,661	,532
Новизна текста			,503	
Примеры из практики				,774
Интрига		,508		,635

Рис. 2. Сравнительный анализ факторных матриц по трем подгруппам группы КазНУ по признаку «Содержание»

Источник: исследование С.В. Мыскина, И.В. Вашуниной, М.П. Кабаковой и К.М. Муса.

Figure 2. Comparative analysis of factor matrices for three subgroups of the KazNU group on the basis of "Content"

Source: compiled by Sergey V. Myskin, Irina V. Vashunina, Mayra P. Kabakova, & Kundyz M. Musa.

Изображение + Текст

	Компонент			
	1	2	3	4
Новые знания и новые фоновые знания	,905			
Полифункциональность	,780			
Обогащение речи	,758	,508		
Учет особенностей аудитории	,658	,575		
Смысловая ценность		,821		
Интерактивность		,720		
Адекватный объем текста			,809	
Развивающий характер			,809	
Применение информации в жизни			,643	
Воспитательные аспекты				,933
Эмоциональное воздействие	,624			,637

Изображение

	Компонент			
	1	2	3	4
Новые знания и новые фоновые знания	,866			
Смысловая ценность	,705			
Развивающий характер	,702			
Учет особенностей аудитории	,635			
Полифункциональность	,523			
Эмоциональное воздействие		,891		
Обогащение речи		,698		
Применение информации в жизни			,778	
Воспитательные аспекты			,744	
Адекватный объем текста			,673	
Интерактивность				,923

Текст

	Компонент		
	1	2	3
Смысловая ценность	,867		
Обогащение речи	,765		
Учет особенностей аудитории	,710		
Интерактивность	,641		
Развивающий характер	,545		
Новые знания и новые фоновые знания	,475		
Полифункциональность		,847	
Эмоциональное воздействие		,783	
Воспитательные аспекты	,572	,626	
Применение информации в жизни			,881
Адекватный объем текста	,587		,594

Рис. 3. Сравнительный анализ факторных матриц по трем подгруппам группы КазНУ по признаку «Прагматика»

Источник: исследование С.В. Мыскина, И.В. Вашуниной, М.П. Кабаковой и К.М. Муса. /

Figure 3. Comparative analysis of factor matrices for three subgroups of the KazNU group on the basis of "Pragmatics"

Source: compiled by Sergey V. Myskin, Irina V. Vashunina, Mayra P. Kabakova, & Kundyz M. Musa.

Поскольку функцией факторного анализа выступает снижение размерности массива данных для удобства их интерпретации, то выделенные константные сочетания требуют обобщения. «Текстовые выделения» предполагают соблюдение правил типографского оформления (шрифт, насыщенность, заголовки и пр.), а «наличие подготовленного плана» — структурированности и логичности изложения. «Воспроизводимость» проявляется в понятности изложения, тема-рематической прогрессии «без скачков», а «интерактивные формы» — наличие заданий в зависимости от вида учебного текста. Как можно отметить, все дефиниции шкал характеризуют некую строгую заданность восприятия формы учебного текста. Скажем так, предварительную инструкцию его познания. Исходя из этого, выделенные нами константные

сочетания шкал по признаку «план выражения» можно обозначить как «Инструктивность».

Константные сочетания шкал по признаку «план содержания» включают следующие дефиниции. «Логика изложения» проявляется в последовательности движения от центрального объекта к периферийным, обозначении смысловых связей частей темы в общей структуре текста. «Креативность» понимается как отступление от стандарта визуального представления содержания текста (терминов, определений, «лида» и пр.), а «методологичность» предполагает соблюдение общепринятых научных требований к приемам представления материала о соответствующей предметной области. «Актуальность информации» определяется соответствием текущему моменту, ситуации, а «новизна текста» раскрывается при соотнесении известной и новой информации. Обобщение данных дефиниций может быть осуществлено по критерию системности, поскольку каждая отражает некую связь одного аспекта познания с другим. Таким образом, константные сочетания шкал по признаку «план содержания» можно обозначить как «Иерархия текста».

Следующий блок константных сочетаний шкал включает: «учет особенностей аудитории» как разработку текста на основе предварительного изучения характеристик адресата; «новые знания и новые фоновые знания» предполагает введение в учебном тексте новой информации с учетом знаний обучающихся; «адекватный объем текста» соответствует основному тексту темы и «применение информации в жизни» как возможность использования информации вне учебного процесса. Обобщение сделанных дефиниций возможно по критерию развития потенциала реципиента, его способностей, формирования нужных компетенций. Следовательно, константные сочетания шкал по признаку «план прагматики» можно обозначить как «Компетентность».

Итак, модель казахской этнокультурной семантической универсалии выражается следующей схемой (рис. 4).

Рис. 4. Модель казахской этнокультурной семантической универсалии
Источник: исследование С.В. Мыскина, И.В. Вашуниной, М.П. Кабаковой и К.М. Муса.
Figure 4. The model of the Kazakh ethnocultural semantic universality
Source: compiled by Sergey V. Myskin, Irina V. Vashunina, Mayra P. Kabakova, & Kundyz M. Musa.

Учитывая симультанный характер процесса восприятия речи, отношения между элементами модели носят двустороннюю направленность. Сукцессивность (последовательность) же будет проявляться при воспроизводстве учебного текста.

Интерпретация результатов исследования

Данную схему возможно объяснить с позиции когнитивно-стилевого подхода. Значение инструктивности релевантно *полезависимому* когнитивному стилю. Ведь предварительная инструкция по взаимодействию с учебным текстом обеспечивает преодоление контекста, влияние которого затрудняет свободную и быструю переработку информации. Иерархия текста раскрывается через такие стили как: *конкретность* — зависимость от статуса и авторитета, нетерпимость к неопределенности, превалирование стереотипных решений; *рефлексивность* — более замедленный темп принятия решения, связанный с многократной перепроверкой и уточнением гипотез на основе тщательного анализа всех возможных вариантов; *когнитивная сложность* — склонность к созданию многомерных моделей реальности, выделяя в них множество взаимосвязанных сторон. Компетентность объемлет: *толерантность к нереалистическому опыту* — открытость новому опыту и способность идти на компромиссы, оценка опыта по его фактическим характеристикам; *фокусирующий контроль* — внимание к фрагментам, представленные явными характеристиками ситуации; *узкий диапазон эквивалентности* — ориентация на различия объектов, обращая внимание главным образом на их детали и отличительные признаки.

В соответствии с общей идеей исследования необходимо выявить социально-психологические корреляты этнокультурной семантической универсалии. В связи с этим рассмотрим психологические особенности национального менталитета казахского народа и попытаемся найти их отражение в специфике процесса восприятия учебного текста. Основным методическим приемом установления культурно-психологического соответствия будет выявление в казахском языковом сознании устойчивых выражений, изречений, словесных оборотов и пр., отражающих общепринятые паттерны поведения, специфику межличностных и ценностных отношений.

Психология казахов зарождалась в быту и нравах кочевых народов Центральной Азии. Примечательны в этом плане труды ученых В.В. Бартольда[2], Б.Я. Владимирцова[3], Г.А. Федорова-Давыдова[4] и мн. др. выдающихся этнографов.

---

[2] *Бартольд В.В.* История культурной жизни Туркестана. Ленинград : Изд-во РАН СССР, 1927.

[3] *Владимирцов Б.Я.* Работы по истории и этнографии монгольских народов. М. : Восточная литература, 2002.

[4] *Фёдоров-Давыдов Г.А.* Искусство кочевников и Золотой Орды: Очерки культуры и искусства народов Евразийских степей и золотоордынских городов. М. : Искусство, 1976.

Общепринято считать, что национальные психологические черты любого народа определяются бытовыми, геополитическими, климатическими, этнографическими и историческими условиями жизни его представителей. Менталитет Казахстана также сложился под влиянием географических и социально-экономических условий: «кочевой быт, безграничные степные просторы, континентальный климат, преобладание скотоводческих хозяйств» [20. С. 372]. Для подобного образа жизни характерна особая коммуникативная культура с малой интенсивностью контактов. Ведь скотоводство вынуждало казахов целыми месяцами не видеться с людьми из другого рода, племени, а зачастую и со своими соплеменниками. Годовые кочевки порой насчитывали сотни и тысячи километров передвижения. Данные обстоятельства легли в основу возникновения особых строгих принципов речевой коммуникации. Константность, устойчивость коммуникативных ритуалов, основанных на хорошем знании друг друга и его роли в общине в соответствии с возрастом и занимаемым положением. Правила общения прививались с раннего детства. Именно это служило залогом установлению быстрых бесконфликтных связей и позитивному отношению, а отступление от правил морально порицалось. Ведь жизнь в территориально-разрозненных, природно-суровых и социально-тяжелых условиях в психологическом плане могла поддерживаться лишь высокой сплоченностью, что исторически формировало у казахского народа толерантное сознание и гуманизм [20. С. 373–374].

По нашему мнению, эти принципы проявляются и при восприятии учебного текста, а именно в опоре на четкие правила, план, инструкцию при его познании. Подтверждение этому мы встречаем в языковом сознании казахов в виде пословиц, поговорок, народных мудростей, сказаний. Например: *Жолаушыны сусындатпай тұрып бұйымтайын сұрама* ‘Не напоив путника, не спрашивай о деле’, *Қонақ келсе ет қызарар, ет қызармаса, бет қызарар* ‘Когда гость приходит, мясо варят, нет мяса — лицо хозяина горит со стыда’, *Сыйларға асың болмаса, сипарға тілің болсын* ‘Если нечем гостя угостить, доброе слово для него найди’, *Қонақ келсе есікке, жүгіріп шық, кешікпе* ‘Если гость придет, беги к двери, не опоздай’. Древние традиции строго регламентировали начало коммуникации с любым путником: его полагалось встретить, не спрашивая ни о чем, пригласить в юрту, накормить и лишь после этого можно поинтересоваться кто он и откуда. Вековой закон *Қонақ атаңнан да үлкен* ‘Гость старше отца’ обязывал хозяина первым здороваться, прерывать свою работу, предлагать свои услуги и помощь [21. С. 46]. Можно предположить, что данные традиции способствовали формированию общего впечатления о коммуниканте, составлению некого плана дальнейшей беседы, задавали общие коммуникативные рамки. Социально-психологическое значение такого вступления в общение заключается в блокировании неадекватных, косвенных ассоциаций, импульсивных вербальных реакций, которые бы вы-

ступали помехами для диалектического обсуждения предмета будущей беседы, совместной деятельности. Как видно из приведенных примеров, наличие предварительной инструкции или четкого начального плана познания нового выступает общим этнокультурным принципом для казахской коммуникативной культуры и является неотъемлемым для понимания.

Категория «иерархия текста» в структуре этнокультурной семантической универсалии казахов также имеет психологическую природу. Иерархия текста традиционно понимается как особая организация текста на основе «коммуникативных отношений подчинения и сочинения»[5]. Последовательное разворачивание содержания от главной мысли к второстепенным позволяет удерживать логику изложения, обоснования. В психологии казахов данный принцип представлен особым отношением к старшим, авторитетам, влиятельным людям. Ведь мысль, выраженная в слове, отражает и образ жизни. Примечательна в этом отношении культурная традиция казахов чтить интересы «трех народов» — собственного рода (*өз жұрты*), рода матери (*нағашы жұрты*) и рода жены (*қайын жұрты*). При этом важным считалось знание имен своих предков в семи поколениях, их историю жизни. Знание своей родословной является устойчивой этнической традицией казахов. Из поколения в поколение бережно передавались реликвии общины, что формировало почитание и уважение к старшим, предкам и роду в целом. В воспитании детей почитание и уважение к старшим культивировалось с ранних лет [22. С. 66–67]. Мысль о том, что дети должны стать достойными наследниками родителей, быть надежной опорой в старости и продолжать традиции их рода, и по сей день является ключевой ценностью воспитания подрастающего поколения казахского народа. В языковом сознании казахов можно встретить соответствующие пословицы: *Жеті атасын білген жетімдік көрмейді* ‘Знающий своих предков до седьмого колена сиротой не будет’, *Әкең өлсе де, әкенді көрген өлмесін* ‘Пока живы знавшие отца — жива и память о нем’, *Ата — асқар тау, ана — бауырыдағы бұлақ, бала — жағасындағы құрақ*» ‘Отец — неприступная гора, мать — родник у подножия горы, дитя — тростник, растущий на берегу реки’. Подобное бережное отношение к традициям, обычаям, предкам и в современных условиях жизни позволяет сохранять казахам целостность картины мира, прогнозировать события, исход случившегося. Приведенные обоснования возможно использовать как принцип разворачивания содержания учебного текста: от главной идеи, авторитетных формулировок, весомой аргументации к предикативным суждениям.

Третья составляющая этнокультурной семантической универсалии казахов — это «компетентность». В нашем понимании компетентность трактуется как некое неотъемлемое требование к качествам, знаниям и умениям человека для эффективного выполнения профессиональной деятельности. Применительно к учебному тексту это означает обязательное указание на

[5] Толковый переводоведческий словарь. М. : Флинта: Наука, 2003.

практическое применение знаний. Стремление казахов к прагматизации любых усвоенных знаний обусловлено их историей, национальным образом жизни. Достаточно суровая степная жизнь требовала от людей физической силы и крепости, выносливости, гибкости ума, социальной зрелости и психологической стабильности. В таких условиях мужчина мог рассчитывать лишь на себя и должен был уметь все, что позволяло организовывать жизнь его семьи. Он был и воином, и пастухом, и охотником, и главой семьи, и представителем рода[6]. В казахском языковом сознании это отражено в поговорке *Жігітке жетпіс өнер де аз* ‘Джигиту и семидесяти ремесел мало’.

Необходимо отметить, что выявление общекультурных и этнокультурных семантических универсалий достигалось в процессе вычленения типичных признаков комбинаторики параметров при восприятии учебных текстов различной модальности. Эти признаки могут рассматриваться как знаки с семантикой соответствующих универсалий, результатом проведенного исследования в семиотическом аспекте является создание интерпретант новых знаков. Знаки могут рассматриваться как материально воплощенные различия, которые изменяют физическое поведение [23. P. 81]. Константность присутствия этих признаков при оценке любого текста указывает на их универсальность.

Заключение

Результаты проведенного междисциплинарного (семиотического, психосемантического, психолингвистического, когнитологического и этнопсихологического) исследования восприятия учебного текста подтвердили основную гипотезу. Прежде всего, экспериментально были выявлены общекультурные (инвариантные) и этнокультурные (вариативные) семантические компоненты, обусловливающие смысловое восприятие учебного текста. Общекультурная универсалия представлена совокупностью параметров учебного текста, функционирующая в языковом сознании представителей всей профессиональной группы. Этноспецифический семантический компонент выражен комбинаторным соотношением отдельных параметров учебного текста, специфичным для каждой этнолингвокультуры. Причем комбинаторика параметров имеет этноспецифику на уровне плана формы (выражения) и содержания языкового знака, а также в аспекте прагматики.

Обработка результатов методами математической статистики вскрыла этноспецифические различия в факторных структурах. Для казахской лингвокультуры этноспецифическая универсалия представлена сочетанием «инструктивности + иерархии текста + компетентности». Именно эти параметры совокупно участвуют в формировании смысла воспринимаемого

[6] *Жаскайрат М., Кабакова М., Лысенкова Ю., Танирбергенов С., Абуова Г., Исмаилов Е., Мустафин Е., Макан К.* Руководство для мужчин. Как быть осознанным отцом и мужем. Нур-Султан, 2020. С. 14–17.

текста. Также данные семантические структуры позволяют выделить некоторые особенности этнокультурного когнитивного стиля казахов.

В целом, результаты проведенного междисциплинарного исследования показали, что общекультурные и этноспецифические универсалии действуют по принципу генераторов и селекторов семантических связей. Можно предположить, что общекультурные универсалии относятся к генераторам и стимулируют установление всевозможных семантических связей по внешним и смысловым признакам. Этноспецифические универсалии действуют как селекторы и выделяют те семантические сочетания, которые характерны для конкретной лингвокультуры и национальной психологии.

Промежуточные итоги данного исследования, по мнению авторов, подтверждают существование облигаторных связей языковых единиц учебного текста с их социально-психологическими коррелятами, которые влияют на понимание этого текста, а также знаковый характер этих связей, создающий предпосылки для их исследования в рамках эдусемиотики.

Список литературы

1. *Semetsky I.* Introduction: A primer on Edusemiotics // Edusemiotics – A Handbook. Singapore : Springer, 2017. P. 1–15. https://doi.org/10.1007/978-981-10-1495-6
2. *Semetsky I.* (Ed.). Edusemiotics: A Handbook. Singapore : Springer, 2017. https://doi.org/10.1007/978-981-10-1495-6
3. *Nadin M.* Meaning and semiosis: To know is by necessity purposeful // Open semiotics. Vol. 1: Epistemological and Conceptual Foundations, A. Biglari (Ed.). L'Harmattan, 2023. P. 345–367
4. *Volli U.* Towards an open semiotic toolbox // Open semiotics. Vol. 1: Epistemological and Conceptual Foundations, A. Biglari (Ed.). L'Harmattan, 2023. P. 149–160.
5. *Auxier R.E.* The pragmatic theory of signs as a strategy for developing unlimited semiosis// Open semiotics. Vol. 1: Epistemological and Conceptual Foundations, A. Biglari (Ed.). L'Harmattan, 2023. P. 209–224.
6. *Зимняя И.А.* Смысловое восприятие речевого сообщения // Смысловое восприятие речевого сообщения (в условиях массовой коммуникации) / отв. ред. Т.М. Дридзе, А.А. Леонтьев. М. : Наука, 1976. С. 5–33.
7. *Чебанов С.В.*, *Мартыненко Г.Я.* Семиотика описательных текстов. Типологический аспект. СПб. : Изд-во Санкт-Петербургского университета, 1999.
8. *Гвоздев А.Н.* Современный русский язык. Синтаксис. М., 1958.
9. *Дридзе Т.М.* Интерпретационные характеристики и классификация текстов (с учетом специфики интерпретационных сдвигов) // Смысловое восприятие речевого сообщения (в условиях массовой коммуникации) / отв. ред. Т.М. Дридзе, А.А. Леонтьев. М. : Наука, 1976. С. 34–45.
10. *Валгина Н.С.* Теория текста. М. : Логос, 2003.
11. *Лурия А.Р.* Язык и сознание. М. : Изд-во Московского университета, 1970.
12. *Залевская А.А.* Слово в лексиконе человека. Психолингвистическое исследование. Воронеж : Воронеж. гос. ун-т, 1990.
13. *Рогожникова Т.М.* Взаимосвязь ассоциативного мира слова и текста // Проблемы семантики: психолингвистические исследования : сб. научных трудов. Тверь, 1991. С. 47.
14. *Сорокин Ю.А.* Психолингвистические аспекты изучения текста : дис. ... д-ра филол. н. М. : Наука, 1985.

15. *Мыскин С.В.* Типология профессиональных коммуникативных ситуаций // Вестник Российского университета дружбы народов. Серия: Вопросы образования: языки и специальность. 2015. № 2. С. 40–56. EDN: TVVNMD
16. *Нистратов А.А., Тарасов Е.Ф.* Психосемантический эксперимент как инструмент анализа смысла и значения слова // Вопросы психолингвистики. 2017. № 2 (32). С. 124–133. EDN: YZGBIZ
17. *Picione R. de L., Freda M.F.* The contribution of semiotics to psychology: A focus on dynamic sense-making processes // Open semiotics. Vol. 4: Life and its Extensions, A. Biglari (Ed.). L'Harmattan, 2023. P. 193–212.
18. *Гальперин И.Р.* Текст как объект лингвистического исследования. М. : КомКнига, 2007.
19. *Загоруйко Н. Г.* Прикладные методы анализа данных и знаний. Новосибирск : ИМ СО РАН, 1999. EDN: PGXPTM
20. *Егинесова А.К.*, *Ерубаева А.Р.* Психологические особенности национального менталитета казахского народа // Международный журнал экспериментального образования. 2015. № 4 (2). С. 371–376.
21. *Радлов В.В.* Избранные сочинения. Алматы, 1952.
22. *Кабакова М.П.* Психологические особенности казахской семьи: монография. Алматы: Қазақ университеті, 2018.
23. *Cariani P.* Systems semiotics: A transdisciplinary framework for informational relations in natural and artificial systems // Open semiotics. Vol. 1: Epistemological and Conceptual Foundations, A. Biglari (Ed.). L'Harmattan, 2023. P. 81–104.

References

1. Semetsky, I. (2017). Introduction: A primer on Edusemiotics. In: *Edusemiotics – A Handbook*, I. Semetsky (Ed.) (pp. 1–15). Singapore: Springer. https://doi.org/10.1007/978-981-10-1495-6
2. Semetsky, I. (Ed.) (2017). *Edusemiotics: A Handbook*. Singapore: Springer. https://doi.org/10.1007/978-981-10-1495-6
3. Nadin, M. (2023). Meaning and Semiosis: To Know is by Necessity Purposeful. In: *Open Semiotics*, Vol. 1: Epistemological and Conceptual Foundations, A. Biglari (Ed.) (pp. 345–367). L'Harmattan.
4. Volli, U. (2023). Towards an Open Semiotic Toolbox. In: *Open Semiotics*, Vol. 1: Epistemological and Conceptual Foundations, A. Biglari (Ed.) (pp. 149–160). L'Harmattan.
5. Auxier, R.E. (2023). The Pragmatic Theory of Signs as a Strategy for Developing Unlimited Semiosis. In: *Open Semiotics*, Vol. 1: Epistemological and Conceptual Foundations, A. Biglari (Ed.) (pp. 209–224). L'Harmattan.
6. Zimnaya, I.A. (1976). Semantic Perception of a Speech Message. In: *Semantic Perception of a Speech Message (in the Context of Mass Communication)*, T.M. Dridze, & A.A. Leontiev (eds.) (pp. 5–33). Moscow: Nauka. (In Russ.).
7. Chebanov, S.V., & Martynenko, G.Ya. (1999). *Semiotics of Descriptive Texts. Typological Aspect*. Saint Petersburg:. Saint Petersburg University publ. (In Russ.).
8. Gvozdev, A.N. (1958). *Modern Russian Language*. *Syntax*. Moscow. (In Russ.).
9. Dridze, T.M. (1976). Interpretative Characteristics and Classification of Texts (Taking into account the Specifics of Interpretative Shifts). In: *Semantic Perception of a Speech Message (in the Context of Mass Communication)*, T.M. Dridze, & A.A. Leontiev (eds.) (pp. 34–45). Moscow: Nauka. (In Russ.).
10. Valgina, N.S. (2003). *Text Theory*. Moscow: Logos. (In Russ.).
11. Luria, A.R. (1970). *Language and Consciousness*. Moscow: Moscow University Publ. (In Russ.).
12. Zalevskaya, A.A. (1990). *The Word in the Human Lexicon. Psycholinguistic Research*. Voronezh: Voronezh State University publ. (In Russ.).

13. Rogozhnikova, T.M. (1991). The Interrelation of the Associative World of Words and Text. In: *Problems of Semantics: Psycholinguistic Research*: Collection of articles (p. 47). (In Russ.).
14. Sorokin, Yu.A. (1985). *Psycholinguistic Aspects of Text Study* [PhD Thesis]. Moscow: Nauka. (In Russ.).
15. Myskin S.V. (2015). Typology of Professional Communicative Situations. *Polylinguality and Transcultural Practices*, 2, 40–56. (In Russ.). EDN: TVVNMD
16. Nistratov, A.A., & Tarasov, E.F. (2017). Psychosemantic Experiment as a Tool for Analyzing the Meaning and Meaning of a Word. *Journal of Psycholinguistics*, 2(32), 124–133. (In Russ.). EDN: YZGBIZ
17. Picione, R. de L., & Freda, M.F. (2023). The Contribution of Semiotics to Psychology: A Focus on Dynamic Sense-Making Processes. In: *Open Semiotics*, Vol. 4: Life and its Extensions, A. Biglari (Ed.) (pp. 193–212). L'Harmattan.
18. Galperin, I.R. (2007). *Text as an Object of Linguistic Research*. Moscow: KomKniga. (In Russ.).
19. Zagoruiko, N.G. (1999). *Applied Methods of Data and Knowledge Analysis*. Novosibirsk: IM SB RAS. (In Russ.). EDN: PGXPTM
20. Eginesova, A.K., & Yerubayeva, A.R. (2015). Psychological Features of the National Mentality of the Kazakh People. *International Journal of Experimental Education*, 4 (2). 371–376. (In Russ.).
21. Radlov, V.V. (1952). *Selected Works*. Almaty. (In Russ.).
22. Kabakova, M.P. (2018). Psychological Features of the Kazakh Family: a monograph. Almaty: Kazakh University publ. (In Kazakh).
23. Cariani, P. (2023). Systems Semiotics: A Transdisciplinary Framework for Informational Relations in Natural and Artificial Systems. In: *Open Semiotics*, Vol. 1: Epistemological and Conceptual Foundations, A. Biglari (Ed.) (pp. 81–104). L'Harmattan.

Сведения об авторах:

Мыскин Сергей Владимирович, доктор филологических наук, кандидат психологических наук, доцент, профессор дирекции образовательных программ, Московский городской педагогический университет (129226, Российская Федерация, г. Москва, 2-й Сельскохозяйственный проезд, д. 4, корп. 1); *сфера научных интересов*: психолингвистика, психосемантика, когнитивные исследования языка и речи, языковая безопасность, деструктивная речевая коммуникация, языковая личность; *e-mail*: myskinsv@mgpu.ru
ORCID: 0000-0002-2665-6089; SPIN-код 2807-7271; Author ID: 391199.

Вашунина Ирина Владимировна, доктор филологических наук, доцент, заведующий кафедрой романо-германских языков, Всероссийская академия внешней торговли Министерства экономического развития Российской Федерации» (119285, Российская Федерация, г. Москва, Воробьевское шоссе, д. 6А); *сфера научных интересов*: психолингвистика, речевое воздействие, теория текста, креолизованный текст, смысловое восприятие речевого сообщения, экспериментальный исследования речи; *e-mail*: vashunina@yandex.ru
ORCID: 0000-0002-3482-8215; SPIN-код 7055-4402; Author ID 286574.

Кабакова Майра Победовна, кандидат психологических наук, профессор, профессор кафедры общей и прикладной психологии факультета философии и политологии, Казахский национальный университет им. аль-Фараби (050040, Республика Казахстан, г. Алматы, просп. Аль-Фараби, д. 71); *сфера научных интересов*: этнопсихология семьи и брака; социальная и педагогическая психология; психология общения и управления; *e-mail*: pobedovna_70@mail.ru
ORCID: 0000-0002-5374-3305.

Муса Кундыз Мусаевна, докторант PhD, старший преподаватель кафедры общей и прикладной психологии факультета философии и политологии, Казахский национальный университет им. аль-Фараби (050040, Республика Казахстан, г. Алматы, просп. Аль-Фараби, д. 71); *сфера научных интересов*: этнопсихология семьи и брака; социальная и педагогическая психология; психология общения и управления; *e-mail*: kundyz2@mail.ru
ORCID: 0009-0004-9937-8384.

Information about the authors:

Sergey V. Myskin, Dr.Sc. (Philology), PhD in Psychology, Associate Professor, Professor of the Directorate of Educational Programs, Moscow City University (4 buil. 1 2nd Agricultural Passage, Moscow, Russian Federation, 129226); *Research interests*: psycholinguistics, psychosemantics, cognitive studies of language and speech, language safety, destructive speech communication, linguistic personality; *e-mail*: myskinsv@mgpu.ru
ORCID: 0000-0002-2665-6089; SPIN code 2807-7271; Author ID: 391199.
Irina V. Vashunina, Dr.Sc. (Philology), Associate Professor, Head of the Department of Romano-Germanic Languages at the VAVT Department, Ministry of Economic Development of the Russian Federation (6A Vorobyevskoe Shosse, Moscow, Russian Federation, 119285); *Research interests*: psycholinguistics, speech impact, text theory, creolized text, semantic perception of speech messages, experimental speech research; *e-mail*: vashunina@yandex.ru
ORCID: 0000-0002-3482-8215; SPIN code 7055-4402; Author ID 286574.
Maira P. Kabakova, PhD in Psychology, Professor, Professor of the Department of General and Applied Psychology, Faculty of Philosophy and Political Science, Al-Farabi Kazakh National University (71 Al-Farabi Ave., Almaty, the Republic of Kazakhstan, 050040); *Research interests*: ethnopsychology of family and marriage; social and pedagogical psychology; psychology of communication and management; *e-mail*: pobedovna_70@mail.ru
ORCID: 0000-0002-5374-3305.
Kundyz M. Musa, PhD student, Senior Lecturer at the Department of General and Applied Psychology, Faculty of Philosophy and Political Science, Al-Farabi Kazakh National University (71 Al-Farabi Ave., Almaty, the Republic of Kazakhstan, 050040); *Research interests*: ethnopsychology of family and marriage; social and pedagogical psychology; psychology of communication and management; *e-mail*: kundyz2@mail.ru
ORCID: 0009-0004-9937-8384.

RUDN Journal of Language Studies, Semiotics and Semantics
2313-2299 (Print), ISSN 2411-1236 (Online)
Вестник Российского университета дружбы народов.
Серия: ТЕОРИЯ ЯЗЫКА. СЕМИОТИКА. СЕМАНТИКА

2026 Vol. 17 No. 2 778–795
http://journals.rudn.ru/semiotics-semantics

DOI: 10.22363/2313-2299-2026-17-2-778-795
EDN: NCXYUW

Научная статья / Research Article

Эволюция политической медиариторики во Вьетнаме 1945–2025 гг.: лингвистический анализ

В.А. Сербин[1,2]

[1]Российский университет дружбы народов, *Москва, Российская Федерация*
[2]Московский государственный лингвистический университет, *Москва*, *Российская Федерация*
✉vietvladimir@gmail.com

Аннотация. Актуальность исследования эволюции политической медиариторики Вьетнама с 1945 по 2025 г. в контексте последовательной смены медиарежимов обусловлена необходимостью лингвистического описания механизмов адаптации официального дискурса к условиям цифровой трансформации коммуникации в политических системах с устойчивым режимом. Цель исследования — выявить закономерности диахронической трансформации политического дискурса при сохранении его идеологического ядра. Материалом послужили три авторских корпуса официальных текстов и публичных речей совокупным объемом 2 393 000 знаков, охватывающих ключевые периоды истории Вьетнама. Анализ проводился на морфосинтаксическом, лексико-семантическом и дискурсивно-прагматическом уровнях с применением корпусных инструментов (AntConc, UDPipe) и качественного кодирования (NVivo) в рамках интегративного лингвоцентричного подхода, объединяющего политическую медиариторику, теорию медиатизации и концепцию платформенной логики. Выделены три этапа эволюции: революционно-мобилизационный монолог (1945–1975), институционально-медиатизированный диалог (1986–2010) и платформенно-алгоритмический дискурс (2010–2025). Установлено, что формальные параметры высказывания (синтаксическая структура, лексический состав, прагматические стратегии) трансформируются под влиянием сменяющихся медиасред, тогда как идеологическое ядро (роль Коммунистической партии, риторика национального единства, социалистические ценности) остается структурно неизменным. Предложена модель адаптивной идеологической устойчивости, в которой медиасреда функционирует как технологическая оболочка, обеспечивающая воспроизводство смыслового ядра в новых коммуникационных форматах. Научная новизна состоит в системном диахроническом описании вьетнамской политической медиариторики как целостного историко-лингвистического континуума и в развитии концепции «алгоритмического идиолекта» — институционального речевого регистра, оптимизированного под алгоритмическую логику цифровых платформ.

Ключевые слова: политический дискурс, диахрония дискурса, алгоритмический идиолект, медиатизация, вьетнамский язык

Конфликт интересов: автор заявил об отсутствии конфликтов интересов.

© Сербин В.А., 2026
This work is licensed under a Creative Commons Attribution-NonCommercial 4.0 International License
https://creativecommons.org/licenses/by-nc/4.0/legalcode

История статьи: дата поступления: 20.01.2026; дата приема в печать: 10.02.2026.

Для цитирования: *Сербин В.А.* Эволюция политической медиариторики во Вьетнаме 1945—2025 гг.: лингвистический анализ // Вестник Российского университета дружбы народов. Серия: Теория языка. Семиотика. Семантика. 2026. Т. 17. № 2. С. 778–795. https://doi.org/10.22363/2313-2299-2026-17-2-778-795 EDN NCXYUW

Evolution of Political Media Rhetoric in Vietnam 1945–2025: A Linguistic Analysis

Vladimir A. Serbin[1,2] ✉

[1]RUDN University, *Moscow, Russian Federation*
[2]Moscow State Linguistic University, *Moscow*, *Russian Federation*
✉vietvladimir@gmail.com

Abstract. The relevance of investigating the evolution of Vietnam's political media rhetoric from 1945 to 2025 against the backdrop of successive media regime changes is grounded in the need for a linguistic account of the mechanisms whereby official discourse adapts to the digital transformation of communication in political systems characterized by durable regimes. This study aims to identify the patterns of diachronic transformation of political discourse while preserving its ideological core. Three original corpora of official texts and public speeches totalling 2.393.000 characters were analyzed at the morphosyntactic, lexico-semantic, and discursive-pragmatic levels using corpus tools (AntConc, UDPipe) and qualitative coding (NVivo) within an integrative linguocentric framework combining political media rhetoric, mediatization theory, and platform logic. Three stages of evolution are identified: the revolutionary-mobilizational monologue (1945–1975), the institutionally mediated dialogue (1986–2010), and the platform-algorithmic discourse (2010–2025). It is demonstrated that formal parameters of utterance (syntactic structure, lexical composition, and pragmatic strategies) transform under the influence of successive media environments, while the ideological core (the leading role of the Communist Party, national unity rhetoric, and socialist values) remains structurally stable. A model of adaptive ideological stability is proposed, in which the media environment operates as a technological shell ensuring reproduction of the semantic core across new communication formats. The scientific novelty lies in the systematic diachronic description of Vietnamese political media rhetoric as a unified historico-linguistic continuum and in operationalizing the concept of the "algorithmic idiolect" — a standardized institutional speech register optimized for the algorithmic logic of digital platforms.
Keywords: political discourse; political media rhetoric; discourse diachrony; algorithmic idiolect; mediatization; Vietnamese language

Conflict of interests: The author stated that there were no conflicts of interest.

Article history: Received: 20.01.2026; Accepted: 10.02.2026.

For citation: Serbin, V.A. (2026). Evolution of Political Media Rhetoric in Vietnam 1945–2025: A Linguistic Analysis. *RUDN Journal of Language Studies, Semiotics and Semantics, 17*(2), 778–795. (In Russ.). https://doi.org/10.22363/2313-2299-2026-17-2-778-795 EDN NCXYUW

Введение

В условиях цифровизации публичной коммуникации и распространения платформенно-алгоритмических форм политических сообщений особую

значимость приобретает лингвистическое описание того, каким образом официальная политическая риторика адаптируется к новым медиасредам. Смена доминирующих медиаканалов от печатной прессы к телевидению, а затем к цифровым форматам коммуникации ставит вопрос о дискурсивно-языковой адаптации власти, поскольку каждый медиарежим задает свою логику для организации политического высказывания.

В медиалингвистике и политической лингвистике данные процессы осмысляются через анализ политической медиариторики, включая диахронические подходы, направленные на выявление закономерностей языковых изменений в зависимости от медиаканала [1–7]. Эти работы реализованы преимущественно на материале западных медиапространств (Франция, Великобритания) и опираются на модель конкурентной публичной сферы.

На этом фоне особый интерес представляет опыт Вьетнама, демонстрирующего иную форму взаимодействия политики и медиа. За период 1945–2025 гг. страна прошла путь от антиколониальной борьбы и консолидации общества к масштабным социально-экономическим трансформациям и интеграции в глобальный контекст. При этом базовая политическая модель сохраняет устойчивость, что позволяет рассматривать вьетнамский политический дискурс как систему, в которой меняются медиариторические формы, но не идеологическое ядро.

Вьетнамские исследования политического дискурса охватывают широкий спектр тем — от анализа революционной пропаганды и агитационных плакатов до парламентского дискурса и цифровых кампаний [8–21]. Однако, как отмечает Нгуен Тхи Минь Нгует, «политическая лингвистика во Вьетнаме пока не оформилась как самостоятельное направление» [22. С. 82], что указывает на фрагментарность данного направления.

Цель статьи — выявить и описать закономерности эволюции политической медиариторики во Вьетнаме в 1945–2025 гг. с учетом смены медиаканалов, на основе корпусного анализа официальных текстов и публичных речей. Для достижения этой цели в работе решаются следующие задачи:

1) разработать и обосновать периодизацию вьетнамской политической медиариторики в соответствии со сменой доминирующих медиарежимов;

2) проанализировать динамику языковых и риторических изменений на морфосинтаксическом, лексико-семантическом и дискурсивно-прагматическом уровнях;

3) выявить механизмы адаптации медиариторических форм к задачам воспроизводства идеологического ядра.

Рабочая гипотеза состоит в том, что в условиях стабильной политической модели политическая медиариторика Вьетнама изменяется главным образом на уровне медийных форм и языковых средств, тогда как ее идеологическое ядро остается устойчивым и последовательно воспроизводится в разных типах медиа. При этом периодизация эволюции определяется сменой медиарежимов и ключевыми социально-политическими событиями.

Научная новизна исследования заключается в том, что политическая медиариторика Вьетнама рассматривается как целостный историко-лингвистический континуум, а также в развитии концепции «алгоритмического идиолекта» как особой формы институционального речевого регистра, адаптированного к условиям платформенно-алгоритмической коммуникации [1].

Материал и методы исследования

Методологическую основу составляет интегративный лингвоцентричный подход, опирающийся на четырехфазную модель медиатизации политики [23; 24]. От инструментального использования медиа как канала (фазы 1–2) к интернализации медиалогики политическими акторами (фазы 3–4). В контексте настоящей работы эти фазы соотносятся с тремя периодами: инструментальное использование печати и радио (Период 1), адаптация к телевизионным форматам (Период 2), интернализация платформенной логики (Период 3).

При анализе периода 2010–2025 гг. дополнительно привлекается теория платформенной логики [25–28]. Видимость сообщения здесь определяется не редакторским решением, а алгоритмическими сигналами вовлечения (время просмотра, реакции, репосты), что порождает принципиально новые прагматические ограничения для дискурса. На уровне лингвистических наблюдений это проявляется в синтаксической модульности, семантической компрессии и стратегической эмоционализации — признаках, описанных ниже как «алгоритмический идиолект».

Интегративной дисциплинарной рамкой выступает политическая медиариторика, трактующая политический медиадискурс как совокупность институционально обусловленных речевых практик, реализуемых в форме поликодовых медиатекстов [3; 5]. Интерпретация результатов опирается на социосемантический подход к анализу репрезентации социальных акторов [29].

Лингвистический анализ проводится на трех уровнях: морфосинтаксическом (длина предложения, синтаксическая сложность, пассивные конструкции), лексико-семантическом (идеологемы, оценочная лексика, метафоры) и дискурсивно-прагматическом (лингвопрагматические стратегии; жанровориторические формы).

При обработке материала сочетались количественные и качественные методы. Частотный анализ и работа с конкордансами выполнялись в AntConc, морфосинтаксическая разметка и расчет показателей синтаксической сложности выполнялись с использованием UDPipe, качественно-количественное кодирование проводится в NVivo.

Под «алгоритмическим идиолектом» понимается стандартизированный институциональный речевой регистр, оптимизированный под критерии платформ (видимость, вовлеченность, цитируемость). Для его идентификации применяются четыре критерия:

1) модульность — синтаксически завершенные сегменты 12–15 слов;

2) цитатная пригодность — ритмическая завершенность фрагмента (параллелизм, анафора, бинарная структура);

3) семантическая компрессия — высокая плотность идеологических маркеров при минимальном синтаксисе;

4) трансмедийная устойчивость — воспроизведение фрагмента в нескольких каналах (пост — заголовок ——хештег/видео).

Каждый из указанных критериев проверялся на корпусном материале трех периодов.

Эмпирическую базу составляют три авторских корпуса официальных текстов совокупным объемом 2 393 000 печатных знаков. Отбор текстов проводился по следующим принципам: а) жанровая репрезентативность — для каждого периода включены тексты основных жанров (выступления, программные документы, медиапубликации); б) источниковая верифицируемость — все тексты извлечены из официальных изданий и архивов; в) хронологический баланс — тексты распределены по десятилетиям внутри каждого периода для исключения концентрации вокруг отдельных событий.

Для первого периода (1945–1975), 980 000 зн., включающего выступления, обращения, статьи и письма из полного собрания сочинений Хо Ши Мина[1], а также публикации официальной партийной прессы (*Nhân Dân* ‘Народ’, *Cứu Quốc* ‘Спасение Родины’ 1945–1963, *Sự Thật* ‘Правда’ 1945–1952, *Quân đội Nhân dân* ‘Вьетнамская народная армия’ 1945–1970). Ограничение по хронологии отдельных изданий обусловлено отсутствием открытого доступа к части архивов и учитывается как фактор репрезентативности при интерпретации результатов.

Для второго периода (1986–2010), 610 000 зн., партийно-государственные документы и публичные выступления. В корпус вошли материалы докладов VI—X съездов КПВ, статьи газеты *Nhân Dân*, а также материалы телеэфира и стенограммы пресс-конференций (контент VTV и др.).

Для третьего периода (2010–2025), 803 000 зн., наряду с традиционными медиа (пресса, телевидение) включает официальные цифровые сообщения, адаптированные к платформенной среде (правительственные порталы, официальные аккаунты и каналы лидеров и организаций в Facebook, YouTube, TikTok и др.).

Во всех трех корпусах анализировалась текстовая и графически оформленная вербальная составляющая медиатекста. Аудиовизуальные элементы (интонация, монтаж, визуальный ряд) не включены в предмет исследования и рассматриваются как перспективное направление дальнейших изысканий.

Эволюция политической медиариторики во Вьетнаме

На основе медиально-исторической логики развития публичной сферы политического дискурса Вьетнама выделяются три основных этапа. Каждый

[1] Hồ Chí Minh Toàn Tập: тт. 4–15. URL: https://hochiminh.vn/tac-pham-cua-ho-chi-minh/ho-chi-minh-toan-tap?categoryId=102001160 (accessed: 19.01.2026).

этап соотнесен с доминирующим типом медиарежима и характеризуется специфическими языковыми и риторическими чертами. Эти этапы условно обозначены как революционно-мобилизационный монолог (1945–1975), институционально-медиатизированный диалог (1986–2010) и платформенно-алгоритмический дискурс (2010–2025). Период 1976–1985 гг. рассматривается как переходный и отдельно не выделяется; основные качественные сдвиги фиксируются на рубежах 1986 и 2010 гг.

Революционно-мобилизационный монолог (1945–1975)

Политическая медиариторика этого периода сформировалась как жестко монологичная система, в которой средства массовой информации усиливали единый голос власти. В терминологии Штрембека это соответствует первой фазе медиатизации: медиа выступают как технический канал, полностью подчиненный политическому актору [23]. Ключевой фигурой дискурса выступал Президент Хо Ши Мин, чей личный стиль, отмеченный простотой и доверительностью, стал символическим центром, задающим канон официального высказывания, который затем воспроизводился всеми институциональными медиа.

Основными каналами коммуникации были печатная партийная пресса и радио. Важным дополнением служили массовые агитплакаты, чьи лаконичные императивные лозунги (*Tiến lên! Toàn thắng ắt về ta!* ‘Вперед! Победа будет за нами!’) в сочетании с яркой визуальной символикой формировали в городском и сельском пространстве непрерывную среду мобилизации, выступая, по определению исследователей, «живым свидетельством особого исторического периода»[2]. Адресат конструировался как коллективный субъект («соотечественники», «армия», «трудящиеся»): массовые обращения типа *Thưa toàn thể đồng bào* (‘Уважаемые соотечественники’) создавали эффект прямого разговора лидера с каждым гражданином.

На морфосинтаксическом уровне наблюдалось расслоение между устной и письменной формами при сохранении общего повелительного тона. Устные выступления Хо Ши Мина отличались краткостью и ритмичностью (средняя длина предложения — 10–20 слов). Хо Ши Мин сознательно избегал канцеляризмов: вместо *tài chính* ‘финансы’ употреблял *tiền* ‘деньги’, вместо *ảnh hưởng tiêu cực* ‘негативное воздействие’ — *xấu* ‘плохо’. Напротив, тексты официальной прессы демонстрировали противоположную тенденцию — синтаксическую развернутость и книжность (35–40 слов). Общей чертой обоих регистров было доминирование побудительных конструкций с модальными частицами *hãy* ‘пусть, давайте’ и *phải* ‘должно, надо’, что делало высказывание безусловным требованием к действию: *Chúng ta phải ra sức thực hiện…* ‘Мы должны приложить все силы, чтобы выполнить…’.

[2] Режим доступа: https://baotanglichsu.vn/en/Articles/3181/72477/the-value-of-the-1945-1975-propaganda-painting-collection-of-the-vietnam-national-museum-of-history.html (дата обращения: 19.01.2026).

Риторический эффект обеспечивали анафора, задававшая императивный ритм: *Phải… phải… phải…* ‘следует…следует…следует…’, и антитеза, противопоставлявшая тяжелое настоящее и неизбежно славное будущее: *kháng chiến… gian khổ, nhưng tiền đồ… rất vẻ vang* ‘война тяжела, но будущее славно’, формируя перформативный дискурс, ориентированный на мобилизацию.

На лексико-семантическом уровне наблюдается высокая степень идеологизации. В табл. 1 приведены наиболее частотные лексемы этого периода и частотность их употреблений:

Таблица 1. **Частотность основных идеологем в текстах 1945–1975 гг.**
Table 1. **Frequency of key ideologemes in the 1945–1975 corpus**

Лексема / Lexeme	**Перевод / English gloss**	**Частотность / Frequency**
nhân dân	народ, народный / people	290
thắng lợi	победа / victory	134
Đảng	партия (правящая) / (ruling) party	118
kháng chiến	сопротивление, война сопротивления / resistance, war of resistance	110
độc lập	независимость / independence	109
sản xuất	производить / production	106
đồng bào	соотечественник / compatriot	92
đoàn kết	единство, сплоченность / unity, cohesion	84
giải phóng	освобождение / liberation	74

Источник: исследование автора В.А. Сербина.
Source: compiled by Vladimir A. Serbin.

Эти ключевые понятия функционировали как маркеры, оправдывающие мобилизацию и подчинение индивидуального коллективному. Оценочная лексика имела резко поляризованный характер: «свои» характеризовались как *anh dũng* ‘героический’, *vinh quang* ‘славный’, а враг демонизировался через *tàn bạo* ‘жестокий’, *xảo quyệt* ‘коварный’. Концептуальные метафоры репрезентировали борьбу как войну (*giặc đói* ‘враг–голод’), а нацию как семью (*đồng bào* ‘букв. дети из одного мешка’), что объясняло мобилизацию через образы угрозы и родства.

На дискурсивно-прагматическом уровне доминировала мобилизующая стратегия.

— Бинарное противопоставление «образа героя» и «образа врага» с моральным возвышением «нашего» дела и дегуманизацией противника.

— Апелляция к патриотизму и коллективным ценностям через обращение к национальной истории, памяти предков и использование родственной метафоры.

— Обещание неизбежной победы и светлого будущего (как награды за жертвы).

— Призывы к абсолютному единству и дисциплине под руководством партии, что легитимировало ее власть.

Поддерживающими механизмами служат постоянная фиксация достижений и интернационализация борьбы через отсылки к соцлагерю.

Риторика Хо Ши Мина систематически опиралась на народные пословицы и разговорные обороты *Có công mài sắt, có ngày nên kim* ‘Если усердно точить железо — однажды получится игла’, вписывая призывы к мобилизации в привычную фольклорную логику. Параллельно официальные тексты выстраивали вертикаль цитирования: изречения Хо Ши Мина функционировали как сакральный канон, «освящавший» текущую повестку, а апелляции к национальному эпосу — сестрам Чынг, королям Хунгам придавали событиям ореол исторической предопределенности.

Институционально-медиатизированный диалог (1986–2010)

С началом экономических реформ «Дой Мой» ‘обновление’ на VI съезде КПВ (1986) монологическая пропаганда трансформируется в прагматичный язык реформ. В терминологии медиатизации это переход ко второй-третьей фазам: медиа — прежде всего телевидение — становятся не просто каналом, а контекстом, который модифицирует формат и ритм политического послания [23]. Коллективный голос партийно-государственных институтов сменяет харизматического лидера, однако образ Хо Ши Мина продолжает функционировать как точка идеологической опоры, вписывающая реформы в логику уже признанного исторического пути. Адресат репрезентируется как обобщенное «национальное сообщество» (*toàn dân* ‘весь народ’, *nhân dân* ‘народ’). Появляются новые для прежней эпохи форматы (пресс-конференции, телеинтервью, трансляции сессий парламента).

На морфосинтаксическом уровне средняя длина предложения сокращается до 24–27 слов, что связано с доминированием телевидения как основного медиаканала. В устных телеобращениях преобладают фразы длиной 15–20 слов, тогда как письменные жанры (доклады съездов, программные документы) сохраняют канцелярскую развернутость (30 слов и более). Характерно широкое использование обезличенных и пассивных конструкций, придающих высказыванию объективированный и институциональный тон (*Cần phải thực hiện đồng bộ các biện pháp…* ‘Необходимо комплексно осуществить меры…’). Директивность речи сохраняется за счет модальных конструкций с *phải* ‘следует’ и *cần* ‘необходимо’, однако они встраиваются в развернутые предложения с целевым обоснованием, что смягчает императив.

Анафора и повтор по-прежнему применяются, но главным образом для логической организации текста, а не для эмоционального эффекта. Типичными становятся нумерованные перечисления *Thứ nhất…, Thứ hai…, Thứ ba…* ‘Во-первых,… Во-вторых,… В-третьих,…’, адаптированные к телевизионному формату. Антитеза смещается от эпического противопоставления «своих» и «врагов» к соотнесению трудностей и достижений развития *Tuy …*

nhưng ... ‘Хотя …, но …’, признающему проблемы, но одновременно нейтрализующему их через акцент на позитивной динамике.

Характерна прагматическая переориентация дискурса VI съезда. Доклад генсека Чыонг Тиня (1986) эксплицитно включает формулу самокритики: *phải nhìn thẳng vào sự thật* ‘нужно смотреть правде в глаза’, впервые позиционируя трудности не как результат внешней агрессии, а как следствие собственных ошибок. Этот прием — признание ошибок при сохранении руководящей роли КПВ — становится устойчивой прагматической фигурой всего периода.

На лексико-семантическом уровне фиксируется обновление терминологии. Доля идеологически маркированной лексики снижается до 10–13 % (против 23–26 % в Периоде 1), удельный вес технократической и экономической лексики возрастает до 17–21 %. В центре риторики закрепляются термины *đổi mới* ‘обновление’, *phát triển* ‘развитие’, *công nghiệp hóa* ‘индустриализация’, *hiện đại hóa* ‘модернизация’, *hội nhập* ‘интеграция’.

Ключевой формулой периода становится *kinh tế thị trường định hướng xã hội chủ nghĩa* ‘рыночная экономика с социалистической ориентацией’, впервые официально закрепленной в документах IX съезда (2001) и последовательно расширявшейся в материалах последующих съездов. XII съезд (2016) дополнил формулировку, добавив определения *hiện đại* ‘современная’ и *hội nhập quốc tế* ‘интегрированная в мировую экономику’. Эволюция формулы демонстрирует характерный для всего периода прием лексико-семантического приращения: идеологическая рамка сохраняется, но наполняется новым содержанием через последовательное добавление определений.

Частотность основных идеологем периода представлена в табл. 2.

Таблица 2. **Частотность основных идеологем в текстах 1986–2010 гг.**
Table 2. **Frequency of key ideologemes in the 1986–2010 corpus**

Лексема / Lexeme	Перевод / English gloss	Частотность / Frequency
nhân dân	народ, народный / people	401
Đảng	партия / party	393
phát triển	развитие / development	385
xây dựng	строительство / construction, building	310
kinh tế	экономика / economy	295
đổi mới	обновление, реформы / renewal, reforms	282
thị trường	рынок / market	198
hội nhập	интеграция (в мировое сообщество) / integration (into the global community)	102

Источник: исследование автора В.А. Сербина.
Source: compiled by Vladimir A. Serbin.

Эмоционально-оценочная лексика активизируется (до 5–7 %). Позитивная оценка используется для фиксации результатов реформ (*thành tựu to lớn* ‘крупные достижения’, *kết quả đáng khích lệ* ‘воодушевляющие результаты’), тогда как негативная семантика смещается с образа внешнего врага на

внутренние проблемы (*tệ nạn xã hội* ‘социальные пороки’, *tham nhũng* ‘коррупция’). Метафоры пути, строительства и управляемого механизма (*con đường đổi mới* ‘путь обновления’, *xây dựng đất nước* ‘строительство страны’) интерпретируют курс «Дой Мой» ‘обновление’ как контролируемый и рациональный процесс.

На дискурсивно-прагматическом уровне доминирующей становится стратегия обоснования власти через результаты, при которой право партии на власть обосновывается через демонстрацию экономических и социальных достижений. Параллельно реализуется стабилизационная стратегия (*không để xảy ra diễn biến hòa bình* ‘не допустить мирной трансформации’). Мобилизационная линия переходит из военного в экономическое русло. Лозунг *«đại đoàn kết toàn dân tộc»* ‘великое всенародное единство’ — апеллирует уже не к отпору агрессору, а к модернизации.

Важным компонентом стратегий остается апелляция к национальной гордости и исторической преемственности. Руководители партии и государства постоянно напоминают, что нынешние экономические достижения являются продолжением героических побед прошлого. Например, вступление Вьетнама в международные организации подается как логичное продолжение исторического пути к полноценному суверенитету и международному признанию страны.

Интертекстуальность обеспечивает идеологическую преемственность: несмотря на смену форматов и задач, реформаторский курс представляется не разрывом с прошлым, а его продолжением. В публичных выступлениях 1990–2000-х гг. систематически воспроизводятся цитаты и формулы Хо Ши Мина как высшего морально-идеологического авторитета: экономические реформы обосновываются через апелляцию к его завету *xây dựng đất nước ta đàng hoàng hơn, to đẹp hơn* ‘построить нашу страну более прекрасной и величественной’, что позволяет вписать политику «Дой Мой» в логику уже признанного исторического пути. Исторические аллюзии смещаются из повседневной риторики в юбилейно-ритуальные контексты, однако в обоих случаях выполняют одну функцию: обосновывают новые политические решения через апелляцию к признанным авторитетам без изменения базовых политико-ценностных ориентиров.

Платформенно-алгоритмический дискурс (2010—2025)

С начала 2010-х гг. политическая коммуникация Вьетнама вступает в цифровой этап, соответствующий четвертой фазе медиатизации по Штрембеку, на которой политические акторы полностью учитывают логику медиа при конструировании высказывания [23]. К середине 2020-х гг. Вьетнам входил в число наиболее активных потребителей платформенного контента в Юго-Восточной Азии[3], что создало устойчивую инфраструктурную основу для системной интеграции власти в платформенную среду.

[3] Режим доступа: https://datareportal.com/reports/digital-2025-vietnam (дата обращения: 19.01.2026).

Коммуникационная модель данного этапа — гибридно-диалогическая, в которой обращения от первого лица, вопросы к подписчикам и формулы благодарности имитируют диалог при сохранении институционально ограниченной интерактивности. Важной чертой становится поликодовость медиатекста — соединение вербальных, визуальных, аудиальных и кинетических семиотических кодов в единое коммуникативное целое [30], принципиально отличающее данный этап от линейной телевизионной риторики. Ключевое инфраструктурное отличие от предшествующих этапов состоит в том, что видимость сообщения определяется не редакторским решением, а алгоритмами платформ, вознаграждающими эмоциональные реакции, репосты и досмотр видео [26]. Это создает для политического высказывания прагматические ограничения, которые проявляются на всех трех анализируемых уровнях.

На морфосинтаксическом уровне наблюдается синтаксическая редукция в платформенном регистре и одновременно сохранение развернутого синтаксиса в доктринальном регистре, что в совокупности означает усиление жанрового расслоения. Данные табл. 4 фиксируют: средняя длина высказывания в платформенном/устном регистре сокращается до 12–15 слов (против 15–20 в Периоде 2), тогда как доктринальные жанры (резолюции, программные доклады) сохраняют показатель 30–40 слов. Платформенная редукция проявляется в постах, подписях к видео, визуальных карточках и хэштегах — жанрах, которых не существовало в предшествующие периоды. Принципиально новым является то, что доктринальные тексты теперь функционируют как источники цитатных модулей — кратких, синтаксически завершенных фрагментов, целенаправленно выделяемых для вторичной циркуляции в цифровой среде.

Механизм такой фрагментации прослеживается в речи генсека Нгуен Фу Чонга на конференции органов внутренних дел (сентябрь 2021 г.), где развернутое высказывание включало четырехчленную формулу *«Tiền hô hậu ủng», «Nhất hô bá ứng», «Trên dưới đồng lòng», «Dọc ngang thông suốt»* ‘Впереди зовут — сзади поддерживают; один призыв — сотня откликов; верх и низ едины; вертикаль и горизонталь слаженны’, где каждый сегмент из 4 слов автономно циркулировал как цитатный модуль в заголовках *Nhân Dân* и визуальных карточках Facebook-аккаунтов.

В отдельный подтип выделяются метафорические формулы Нгуен Фу Чонга, сочетающие народно-разговорную образность с модульной компрессией. Фраза *Lò đã nóng lên rồi thì củi tươi vào đây cũng phải cháy* ‘Если печь раскалена, то и сырые дрова в ней загорятся’ — развернутая антикоррупционная метафора, которая была редуцирована до двухсловного модуля *đốt lò* ‘топить печь’, ставшего хэштегом, заголовком и визуальным маркером целой политической кампании. Переход от 15-словного высказывания к двухсловному модулю наглядно демонстрирует синтаксическую редукцию платформенного регистра.

На лексико-семантическом уровне устойчивое идеологическое ядро (Партия, государство, народ, Отечество) сочетается с расширением экономического и цифрового словаря. Обобщенные данные по частотности ключевых идеологем в корпусе 2010–2025 гг. представлены в табл. 3.

Таблица 3. **Частотность основных идеологем в текстах 2010–2025 гг.**
Table 3. **Frequency of key ideologemes in the 2010–2025 corpus**

Лексема / Lexeme	Перевод / English gloss	Частотность / Frequency
phát triển	развитие / development	530
Chủ tịch	Президент / President	350
nhân dân	народ / people	349
xây dựng	строительство / construction, building	310
doanh nghiệp	предприятие / enterprise	306
công nghệ	технологии / technology	155
thị trường	рынок / market	155
khó khăn	сложности / difficulties	102
thế giới	Мир (географич.) / world (geographical)	100

Источник: исследование автора В.А. Сербина.
Source: compiled by Vladimir A. Serbin.

Таким образом, традиционная связка «партия — лидер — народ» дополняется устойчивыми рядами «развитие — предприятие — технологии — рынок», что отражает экономический и цифровой поворот дискурса.

Усиливается пласт эмоционально-оценочной лексики (до 14%). Формулы благодарности и признательности (*bày tỏ lòng biết ơn sâu sắc* ‘выражать глубокую признательность’, *tri ân chân thành* ‘сердечная благодарность’) вызывают эмоциональные реакции с высоким алгоритмическим весом и легко превращаются в автономные цитатные модули. Формулы мобилизации синхронно тиражируются через сеть официальных акторов: формула COVID-19 *«mỗi người dân là một chiến sĩ trên mặt trận chống dịch»* ‘каждый гражданин — боец на фронте борьбы с эпидемией’, 2020) воспроизводила структурную схему Периода 1 (каждый гражданин = боец), адаптируя ее к формату платформенного цитатного модуля.

На дискурсивно-прагматическом уровне обращение к историческому опыту сохраняется, однако право партии на власть все чаще обосновывается через достигнутые результаты (*kết quả đạt được* ‘достигнутые результаты’, *hiệu quả quản lý* ‘эффективность управления’) и конкретные обещания на будущее (*mục tiêu đến năm 2030* ‘цели до 2030 г.’; *phát triển nhanh và bền vững* ‘быстрое и устойчивое развитие’). Характерной особенностью этапа становится переход от монологической демонстрации власти к согласованному участию разных официальных голосов: политическое высказывание производится сетью акторов от Генерального секретаря и Председателя Национального собрания до министерств и местных партийных организаций, которые варьируют стиль и тональность (деловую, сочувственную, призывную), но сохраняют единое идеологическое поле. Важной стратегией становится

демонстрация заботы и сочувствия со стороны власти: в кризисных ситуациях и торжественных жанрах систематически используется лексика заботы и благодарности, создающая образ государства, «разделяющего трудности» адресата; в платформенной среде это повышает эмоциональную вовлеченность и вероятность алгоритмического распространения. Отсылки к истории и идеологии приобретают служебный характер: они используются уже не для развернутого повествования, а как краткие символические маркеры, легко включаемые в цифровые форматы (цитатные модули, визуальные карточки, хэштеги).

Таблица 4. **Динамика ключевых параметров политической медиариторики Вьетнама (1945–2025)**
Table 4. **Dynamics of key parameters of Vietnam's political media rhetoric (1945–2025)**

Indicator / Показатель	1945–1975	1986–2010	2010–2025
Средняя длина предложения, слов / Average sentence length (oral), words	10–20	15–20	12–15
Средняя длина предложения, слов / Average sentence length (written), words	35–40	30+	30–40
Доля идеологической лексики, % / Share of ideological lexis, %	23–26	10–13	6–9
Доля технократической лексики / Share of technocratic lexis, %	< 5	17–21	24–30
Эмоционально-оценочная лексика, % / Emotive-evaluative lexis	≤ 2	5–7	до 14 / up to 14
Тип коммуникации / Communication type	монолог / monologue	институциональный диалог / institutional dialogue	платформенно-сегментированный формат / platform-segmented format
Основные жанры / Main genres	воззвание, речь / appeal, speech	отчет, телеобращение / report, televised address	пост, видео, мем, инфографика / post, video, meme, infographic

Источник: исследование автора В.А. Сербина.
Source: compiled by Vladimir A. Serbin.

Формат короткого вертикального видео (TikTok, YouTube Shorts, Reels) доводит синтаксическую редукцию до предела. Ролики официальных аккаунтов министерств и ведомств длятся 15–30 секунд и строятся по устойчивой схеме: привлекающий внимание кадр (1–3 сек.) → короткая формула крупным шрифтом (*bảo vệ Tổ quốc* 'защищать Отечество', *vì nhân dân phục vụ* 'служить народу', *kiên quyết, kiên trì* 'решительно и неуклонно') → призыв

к репосту. Вербальная составляющая такого ролика сводится к одной-двум формулам, которые подчинены не синтаксической логике предложения, а ритму монтажного кадра: текст должен быть прочитан за 2–3 секунды, иначе зритель пролистает его. Это предельный случай платформенной редукции: идеологическое высказывание, которое в доктринальном регистре занимает абзац, здесь сжимается до двух–четырех слов, сохраняя при этом основное идеологическое содержание.

Суммируя характеристики этапов, можно проследить четкую динамику ключевых параметров (см. табл. 4).

Заключение

Результаты исследования подтверждают специфику вьетнамской модели политической медиариторики как системы адаптивной идеологической устойчивости, в которой трансформации языковых и медиатехнических параметров не ведут к размыванию смыслового ядра, а напротив обеспечивают его гибкую репрезентацию в меняющихся условиях. Данная модель характерна для политических систем с устойчивым режимом, однако вопрос о ее типологической универсальности требует отдельного сравнительного исследования.

Революционно-мобилизационный этап (1945–1975) сформировал канон: монолог, директивный синтаксис, сакрализованные идеологемы и интертекстуальность, легитимирующую власть через культ Хо Ши Мина и национально-исторические мифы.

Институционально-медиатизированный этап (1986–2010) адаптировал данный канон к задачам экономических реформ. Риторика сместилась от мобилизации к убеждению и демонстрации результативности, синтаксис частично упростился под влиянием телевидения, а ядро дискурса стало воспроизводиться через обоснование курса «Дой Мой».

Платформенно-алгоритмический этап (2010–2025) показывает, как цифровая среда усиливает тактический уровень адаптации. Фрагментация, модульность, «алгоритмический идиолект» функционируют как механизмы управляемого вовлечения и повторяемости. Операционализация понятия через четыре верифицируемых критерия позволила продемонстрировать его системное проявление в корпусном материале.

Перспективы дальнейших исследований видятся в двух направлениях: 1) углубленный анализ аудиовизуальных компонентов цифрового дискурса, где идеологическое ядро все чаще репрезентируется через эмоциональные и визуальные коды; 2) сравнительное изучение вьетнамской модели с другими политическими системами, прошедшими через сходные медиатехнологические переходы (Китай, Лаос), для выявления универсальных и специфических механизмов обеспечения идеологической устойчивости официального политического дискурса в цифровую эпоху.

Список литературы

1. *Ахренова Н.А.* Эволюция политической медиариторики: от монолога к цифровой фрагментации // Медиалингвистика. 2025. Т. 12. № S. С. 168–172.
2. *Ахренова Н.А.*, *Тер-Минасова С.Г.* Эволюция политической медиариторики в Великобритании // Вестник Российского университета дружбы народов. Серия: Теория языка. Семиотика. Семантика. 2025. Т. 16. № 4. С. 1090–1120. https://doi.org10.22363/2313-2299-2025-16-4-1090-1101 EDN: MXKGIU
3. *Добросклонская Т.Г., Логинова П.Г.* Политическая медиариторика: системный подход к исследованию речевых практик политического медиадискурса // Научные исследования и разработки. Современная коммуникативистика. 2023. Т. 12. № 4. С. 47–54. https://doi.org/10.12737/2587-9103-2023-12-4-47-54 EDN: VVCMAF
4. *Логинова П.Г.* Генезис политической медиариторики: французский вектор. М. : Директ-Медиа, 2024. EDN: VVCMAF
5. *Логинова П.Г.* Политическая медиариторика: новый подход к анализу языка политической медиакоммуникации (на примере Франции). М. : Директ-Медиа, 2025.
6. *Логинова П.Г.* Метафорический компонент как предмет исследования политической медиариторики (на материале традиционных французских СМИ) // Вестник МГПУ. Серия: Филология. Теория языка. Языковое образование. 2025. № 3(59). С. 76–91. https://doi.org/10.24412/2076-913X-2025-359-76-91 EDN: EGKTGF
7. *Франчук Ю.А.* Оценка и оценочность как свойство американской политической медиариторики // Медийно-информационная грамотность как составляющая национальной коммуникативной культуры: материалы Международной научно-практической конференции, Липецк, 9–10 ноября 2023 г. Липецк : Липецкий государственный педагогический университет имени П.П. Семенова-Тян-Шанского, 2023. С. 276–280.
8. *DeCaro P.A.* Rhetoric of Revolt: Ho Chi Minh's Discourse for Revolution. Westport : Praeger, 2003.
9. *Nguyen T.H., Tran T.H.A.* Bàn về diễn ngôn chính trị [On political discourse] // Tạp chí Ngôn ngữ và Đời sống. 2014. Iss. 12(230). Tr. 20–26.
10. *Nguyen Thi Hong Minh, Valdez M.L.A.* Ideologies Reflected in the Selected Speeches of Ho Chi Minh // International Journal on Studies in English Language and Literature. 2019. Vol. 7. Iss. 4. P. 17–28. https://doi.org/10.20431/2347-3134.0704002
11. *Стародубова О.Ю.* Концептуальная метафора во вьетнамском политическом дискурсе как механизм конструирования социальной модели // Филологический аспект. 2021. № 3(71). С. 100–109.
12. *Tran T.N.* Social Media, Political Propaganda and Herding: A Case Study of Vietnam in the Coronavirus Crisis // ESSACHESS — Journal for Communication Studies. 2021. Vol. 14. Iss. 2(28). P. 171–202. https://doi.org/10.21409/V6S2-X722
13. *Nguyen H.P.* Conflict Recovery in Socialist Utopia: A Critical Analysis of Vietnam's Postwar Reconstruction Policies Through a Visual Interpretation of 1976—1986 Propaganda Posters // Journal of Public Policy & Marketing. 2025. Vol. 44. № 1. P. 160–178. https://doi.org/10.1177/07439156241246243
14. *Nguyen T.M.N.* Strategies of Direct and Indirect Verbal Expression of Negative Evaluation in Russian and Vietnamese Political Discourse // Proceedings of the Southwest State University. Series: Linguistics and Pedagogics. 2024. Vol. 14. Iss. 3. P. 95–108. https://doi.org/10.21869/2223-151x-2024-14-3-94-108 EDN: PNZDAZ
15. *Le M., Nicolaisen F.S.* Narrative and Framing of a Pandemic: Public Health Communication in the Vietnamese Public Sphere // Southeast Asian Studies. 2024. Vol. 13. Iss. 1. P. 35–71. https://doi.org/10.20495/seas.13.1_35
16. *Mai N.T.T., Luyen N.T.K., Hien Ph.* A Comparative Analysis of American and Vietnamese Presidents' Speeches: A Systemic Functional Grammar Perspective // Theory and Practice in Language Studies. 2024. Vol. 14. Iss. 11. P. 3630–3640. https://doi.org/10.17507/tpls.1411.33 EDN: ROQBNQ

17. *Shohet M.* How to Speak to the Masses: Hồ Chí Minh's Instructions to Cadres and the Dynamics of Register Formation in 20th-Century Vietnam // Journal of Linguistic Anthropology. 2024. Vol. 34. Iss. 1. P. 1–25. https://doi.org/10.1111/jola.12412
18. *Kumaat J.A.W.* Representing Vietnam's Political Leadership in International Media: A Critical Discourse Analysis of Coverage from 2018 to 2024 // International Review of Humanities Studies. 2025. Vol. 10. Iss. 2. Art. 4. https://doi.org/10.7454/irhs.v10i2.1410
19. *Thư N.* Phân tích diễn ngôn bài phát biểu về văn hóa Việt Nam của cố Tổng bí thư Nguyễn Phú Trọng // Tạp chí Khoa học Trường Đại học Mở Hà Nội. 2025. Iss. 1. P. 272–279. https://doi.org/10.59266/houjs.2025.643 EDN: TCVBFP
20. *Txu M.H.H.* Вербальные способы выражения несогласия в политическом дискурсе (на материале вьетнамского парламентского дискурса) // Вестник Московского государственного лингвистического университета. Гуманитарные науки. 2022. № 6(861). С. 99–106. https://doi.org/10.52070/2542-2197_2022_6_861_99 EDN: QXMFXE
21. *Vu C.T., Nguyen Q.T.* When Communist Propaganda Meets Western Public Relations: Vietnam's COVID-19 Communication // Political Communication. 2024. Vol. 12. Iss. 1. С. 33–67. https://doi.org/10.1177/2046147X231218310 EDN: LTKZVC
22. *Нгуен Тхи Минь Нгует.* Вьетнамский политический дискурс: лингвистическая перспектива // Вьетнамские исследования. 2023. Т. 7. № 3–2. С. 78–88. EDN: DDCSWX
23. *Strömbäck J. Four Phases of Mediatization:* An Analysis of the Mediatization of Politics // The International Journal of Press/Politics. 2008. Vol. 13. Iss. 3. P. 228–246. https://doi.org/10.1177/1940161208319097
24. *Strömbäck J., Esser F.* Mediatization of Politics: Towards a Theoretical Framework // Mediatization of Politics: Understanding the Transformation of Western Democracies / eds. F. Esser, J. Strömbäck. London: Palgrave Macmillan, 2014. P. 3–28. https://doi.org/10.1057/9781137305374_1
25. *van Dijck J., Poell T.* Understanding Social Media Logic // Media and Communication. 2013. Vol. 1. Iss. 1. P. 2–14. https://doi.org/10.17645/mac.v1i1.70
26. *Chadwick A.* The Hybrid Media System: Politics and Power. Oxford : Oxford University Press, 2013.
27. *Nieborg D.B., Poell T.* The Platformization of Cultural Production: Theorizing the Contingent Cultural Commodity // New Media & Society. 2018. Vol. 20. Iss. 11. P. 4275–4292. https://doi.org/10.1177/1461444818769694
28. *Boulianne S.* Social Media Use and Participation: A Meta-Analysis of Current Research // Information, Communication & Society. 2015. Vol. 18. Iss. 5. P. 524–538. https://doi.org/10.1080/1369118X.2015.1008542
29. *van Leeuwen T.* Discourse and Practice: New Tools for Critical Discourse Analysis. Oxford : Oxford University Press, 2008.
30. *Чернявская В.Е.* Медиальный поворот в лингвистике: поликодовые и гибридные тексты // Вестник Иркутского государственного лингвистического университета. 2013. № 2(23). С. 43–49. EDN: RBWOWN

References

1. Akhrenova, N.A. (2025). Evolution of Political Media Rhetoric: From Monologue to Digital Fragmentation. *Medialingvistika*, *12*(S), 168–172. (In Russ.).
2. Akhrenova, N.A., & Ter-Minasova, S.G. (2025). Evolution of Political Media Rhetoric in Great Britain. *RUDN Journal of Language Studies, Semiotics and Semantics*, *16*(4), 1090–1120. (In Russ.). https://doi.org10.22363/2313-2299-2025-16-4-1090-1101 EDN: MXKGIU
3. Dobrosklonskaya, T.G., & Loginova, P.G. (2023). Political Media Rhetoric: a Systematic Approach to the Study of Speech Practices of Political Media Discourse. *Nauchnyye issledovaniya i razrabotki. Sovremennaya kommunikativistika, 12*(4), 47–54. (In Russ.). https://doi.org/10.12737/2587-9103-2023-12-4-47-54 EDN: VVCMAF

4. Loginova, P.G. (2024). *Genesis of Political Media Rhetoric: the French Vector*. Moscow: Direkt-Media publ. (In Russ.). EDN: VVCMAF
5. Loginova, P.G. (2025). *Political Media Rhetoric: a New Approach to the Analysis of Political Media Communication Language (the Case of France)*. Moscow: Direkt-Media publ. (In Russ.).
6. Loginova, P.G. (2025). The Metaphorical Component as a Subject of Political Media Rhetoric Research (Based on Traditional French Media). *MCU Journal of Philology. Theory of Linguistics. Linguistic Education, 3*(59), 76–91. (In Russ.). https://doi.org/10.24412/2076-913X-2025-359-76-91 EDN: EGKTGF
7. Franchuk, Yu.A. (2023). Evaluation and Evaluativity as a Property of American Political Media Rhetoric. In *Mediyno-Informatsionnaya Gramotnost kak Sostavlyayushchaya Natsionalnoy Kommunikativnoy Kultury: Materialy Mezhdunarodnoy Nauchno-Prakticheskoy Konferentsii* (pp. 276–280). Lipetsk: Lipetskiy gosudarstvennyy pedagogicheskiy universitet imeni P.P. Semenova-Tyan-Shanskogo publ. (In Russ.).
8. DeCaro, P.A. (2003). *Rhetoric of Revolt: Ho Chi Minh's Discourse for Revolution*. Westport: Praeger.
9. Nguyen, T.H., & Tran, T.H.A. (2014). Bàn về diễn ngôn chính trị [On Political Discourse]. *Tạp chí Ngôn ngữ và Đời sống*, *12*(230), 20–26. (In Vietnamese).
10. Nguyen, T.H.M., & Valdez, M.L.A. (2019). Ideologies Reflected in the Selected Sspeeches of Ho Chi Minh. *International Journal on Studies in English Language and Literature*, *7*(4), 17–28. https://doi.org/10.20431/2347-3134.0704002
11. Starodubova, O.Yu. (2021). Conceptual Metaphor in Vietnamese Political Discourse as a Mechanism for Constructing a Social Model. *Filologicheskiy Aspekt*, *3*(71), 100–109. (In Russ.).
12. Tran, T.N. (2021). Social Media, Political Propaganda and Herding: a Case Study of Vietnam in the Coronavirus Crisis. *ESSACHESS – Journal for Communication Studies*, *14*(2(28)), 171–202. https://doi.org/10.21409/V6S2-X722
13. Nguyen, H.P. (2025). Conflict Recovery in Socialist Utopia: a Critical Analysis of Vietnam's Postwar Reconstruction Policies Through a Visual Interpretation of 1976–1986 Propaganda Posters. *Journal of Public Policy & Marketing*, *44*(1), 160–178. https://doi.org/10.1177/07439156241246243
14. Nguyen, T.M.N. (2024). Strategies of Direct and Indirect Verbal Expression of Negative Evaluation in Russian and Vietnamese Political Discourse. *Proceedings of the Southwest State University. Series: Linguistics and Pedagogics*, *14*(3), 95–108. https://doi.org/10.21869/2223-151x-2024-14-3-94-108 EDN: PNZDAZ
15. Le, M., & Nicolaisen, F.S. (2024). Narrative and Framing of a Pandemic: Public Health Communication in the Vietnamese Public Sphere. *Southeast Asian Studies*, *13*(1), 35–71. https://doi.org/10.20495/seas.13.1_35
16. Mai, N.T.T., Luyen, N.T.K., & Hien, Ph. (2024). A Comparative Analysis of American and Vietnamese Presidents' Speeches: a Systemic Functional Grammar Perspective. *Theory and Practice in Language Studies*, *14*(11), 3630–3640. https://doi.org/10.17507/tpls.1411.33 EDN: ROQBNQ
17. Shohet, M. (2024). How to Speak to the Masses: Hồ Chí Minh's Instructions to Cadres and the Dynamics of Register Formation in 20th-century Vietnam. *Journal of Linguistic Anthropology, 34*(1), 1–25. https://doi.org/10.1111/jola.12412
18. Kumaat, J.A.W. (2025). Representing Vietnam's Political Leadership in International Media: a Critical Discourse Analysis of Coverage from 2018 to 2024. *International Review of Humanities Studies*, *10*(2), 4. https://doi.org/10.7454/irhs.v10i2.1410
19. Thư, N. (2025). Discourse Analysis of the Speech on Vietnamese Culture by the Late General Secretary Nguyen Phu Trong. *Tạp chí Khoa học Trường Đại học Mở Hà Nội*, 1, 272—279. (In Vietnamese). https://doi.org/10.59266/houjs.2025.643 EDN: TCVBFP

20. Thi, M.N.N. (2022). Verbal Means of Expressing Disagreement in Political Discourse (Based on Vietnamese Parliamentary Discourse). *Vestnik of Moscow State Linguistic University. Humanities*, 6(861), 99–106. (In Russ.). https://doi.org/10.52070/2542-2197_2022_6_861_99 EDN: QXMFXE
21. Vu, C.T., & Nguyen, Q.T. (2024). When Communist Propaganda Meets Western Public Relations: Vietnam's COVID-19 Communication. *Political Communication*, *12*(1). https://doi.org/10.1177/2046147X231218310
22. Nguyen, Th.M.Ng. (2023). Vietnamese Political Discourse: A Linguistic Perspective. *Russian Journal of Vietnamese Studies, 7*(3–2), 78–88. (In Russ.). EDN: DDCSWX
23. Strömbäck, J. (2008). Four Phases of Mediatization: An Analysis of the Mediatization of Politics. *The International Journal of Press/Politics*, *13*(3), 228–246. https://doi.org/10.1177/1940161208319097
24. Strömbäck, J., & Esser, F. (2014). Mediatization of Politics: Towards a Theoretical Framework. In: *Mediatization of politics: understanding the transformation of Western democracies* (pp. 3–28). London: Palgrave Macmillan. https://doi.org/10.1057/9781137305374_1
25. van Dijck, J., & Poell, T. (2013). Understanding Social Media Logic. *Media and Communication*, *1*(1), 2–14. https://doi.org/10.17645/mac.v1i1.70
26. Chadwick, A. (2013). *The Hybrid Media System: Politics and Power*. Oxford: Oxford University Press.
27. Nieborg, D.B., & Poell, T. (2018). The Platformization of Cultural Production: Theorizing the Contingent Cultural Commodity. *New Media & Society*, *20*(11), 4275–4292. https://doi.org/10.1177/1461444818769694
28. Boulianne, S. (2015). Social Media Use and Participation: A Meta-Analysis of Current Research. *Information*, *Communication & Society*, *18*(5), 524–538. https://doi.org/10.1080/1369118X.2015.1008542
29. van Leeuwen, T. (2008). *Discourse and Practice: New Tools for Critical Discourse Analysis*. Oxford: Oxford University Press.
30. Chernyavskaya, V.E. (2013). The Medial Turn in Linguistics: Polycode and Hybrid Texts. *Vestnik Irkutskogo gosudarstvennogo lingvisticheskogo universiteta*, 2(23), 43–49. (In Russ.). EDN: RBWOWN

Сведения об авторе:
Сербин Владимир Алексеевич, ведущий научный сотрудник кафедры теории и практики иностранных языков института иностранных языков, Российский университет дружбы народов (117198, Российская Федерация, г. Москва, ул. Миклухо-Маклая, д. 6); доцент Военного учебного центра, Московский государственный лингвистический университет (119034, Российская Федерация, г. Москва, ул. Остоженка, д. 38 с. 1); *сфера научных интересов:* медиалингвистика, алгоритмическая медиалингвистика, лингвопрагматика медиатекста, терминоведение, вьетнамский язык, трансмедийный медианарратив; *e-mail:* vietvladimir@gmail.com
ORCID: 0000-0002-8066-2519; SPIN-код: 2346-8449; AuthorID: 1064447.

Information about the author:
Vladimir A. Serbin, PhD in Philology, Leading Researcher at the Institute of Foreign Languages, RUDN University (6 Miklukho-Maklaya St., Moscow, Russian Federation, 117198); Associate Professor at the Military Training Centre, Moscow State Linguistic University (38 build 1 Ostozhenka St., Moscow, Russian Federation, 119034); *Research interests:* media linguistics, algorithmic media linguistics, linguopragmatics of media text, terminology studies, Vietnamese language, transmedia narrative; *e-mail* vietvladimir@gmail.com
ORCID: 0000-0002-8066-2519; SPIN-code: 2346-8449; AuthorID: 1064447.

RUDN Journal of Language Studies, Semiotics and Semantics
2313-2299 (Print), ISSN 2411-1236 (Online)
Вестник Российского университета дружбы народов.
Серия: ТЕОРИЯ ЯЗЫКА. СЕМИОТИКА. СЕМАНТИКА

2026 Vol. 17 No. 2 796–812
http://journals.rudn.ru/semiotics-semantics

LITERARY TEXT LINGUISTICS

ЛИНГВИСТИКА ХУДОЖЕСТВЕННОГО ТЕКСТА

DOI: 10.22363/2313-2299-2026-17-2-796-812
EDN: NHNQGK

Research Article / Научная статья

Non-Verbal Behavior as a Marker of Altered States of Consciousness in a Literary Piece

Vera V. Ganina[1] ✉, **Faina I. Kartashkova**[2]

[1]Ivanovo State University of Chemistry and Technology, *Ivanovo, Russian Federation*
[2]Ivanovo State University, *Ivanovo, Russian Federation*
✉vvganina@mail.ru

Abstract. The problem of altered states of consciousness from linguistic angle is relevant. Attention is focused on the ways in which altered states of consciousness are reflected in a literary piece. The research is delivered on the basis of Dostoevsky's novel "Crime and Punishment". Different types of non-verbal behavior of the main character and the language correlates are discussed. It is demonstrated that in the situations under consideration, typical controlled non-verbal behavior includes phonation, mimics, eye contact, and pantomimics which have corresponding language correlates, the most frequent of them are adverbial and attributive phrases. The main types of uncontrolled non-verbal behavior (or psychophysiological reactions) in cases of altered states of consciousness are trembling and heartbeat, expressed by different linguistic units, adverbial phrases and comparative constructions being predominant. Specific non-verbal actions that mark the character's negative emotional state and his growing inadequacy often go undeciphered by those around him. The description of how the main character's emotions are displayed and concealed contributes to the creation of his portrayal.

Keywords: altered states of consciousness, emotions, communicative behavior, non-verbal behavior, linguistic correlations, comparative approach

Authors' contribution: The authors contributed equally to this article and preparation of the article text.

Conflict of interests: The authors stated that there were no conflicts of interest.

Article history: Received: 23.12.2025; Accepted: 15.02.2026.

© Ganina V.V., Kartashkova F.I., 2026

This work is licensed under a Creative Commons Attribution-NonCommercial 4.0 International License
https://creativecommons.org/licenses/by-nc/4.0/legalcode

For citation: *Ganina, V.V., & Kartashkova, F.I.* (2026). Non-Verbal Behavior as a Marker of Altered States of Consciousness in a Literary Piece. *RUDN Journal of Language Studies, Semiotics and Semantics, 17*(2), 796–812. https://doi.org/10.22363/2313-2299-2026-17-2-796-812 EDN NHNQGK

Невербальное поведение как признак изменённых состояний сознания в литературном произведении

В.В. Ганина[1], Ф.И. Карташкова[2]

[1]Ивановский государственный химический и технологический университет, *Иваново, Российская Федерация*
[2]Ивановский государственный университет, *Иваново, Российская Федерация*
✉vvganina@mail.ru

Аннотация. Изучение изменения состояний сознания с лингвистической точки зрения является актуальным. Внимание сосредоточено на том, как измененные состояния сознания отражаются в литературном произведении. Исследование проводится на основе романа Достоевского «Преступление и наказание». Обсуждаются различные типы невербального поведения главного героя и языковых коррелятов. Показано, что в рассматриваемых ситуациях типичное контролируемое невербальное поведение включает фонацию, мимику, зрительный контакт и пантомимику, которые имеют соответствующие языковые корреляции, наиболее частыми из которых являются нарекальные и присваивающие фразы. Основными типами неконтролируемого невербального поведения (или психофизиологических реакций) в случаях измененных состояний сознания являются дрожь и сердцебиение, выраженные разными языковыми единицами, преобладающими являются наречия и сравнительные конструкции. Конкретные невербальные действия отмечают негативное эмоциональное состояние персонажа, и его растущая неадекватность не может быть расшифрована другими людьми. Описание того, как эмоции главного героя отображаются и скрываются, способствует созданию его изображения.

Ключевые слова: изменение состояния сознания, эмоции, коммуникативное поведение, невербальное поведение, языковые корреляции, сравнительный подход

Вклад авторов: вклад авторов равнозначен на всех этапах исследования и подготовки текста статьи.

Конфликт интересов: авторы заявили об отсутствии конфликтов интересов.

История статьи: дата поступления: 23.12.2025; дата приема в печать: 15.02.2026.

Для цитирования: *Ganina V.V., Kartashkova F.I.* Non-Verbal Behavior as a Marker of Altered States of Consciousness in a Literary Piece // Вестник Российского университета дружбы народов. Серия: Теория языка. Семиотика. Семантика. 2026. Т. 17. № 2. С. 796–812. https://doi.org/10.22363/2313-2299-2026-17-2-796-812 EDN NHNQGK

Introduction

Theoretical Framework

An altered state of consciousness (ASC) as described by A. Ludwig, is a mental state that occurs due to various factors causing significant deviations in subjective experiences during mental processes [1. P. 9]. The term was first used in

the international anthology titled *Altered States of Consciousness* [1]. The study of ASC is a priority area of interdisciplinary research.

According to Bourguignon [2], ASC refers to states characterized by modifications in sensation, perception, emotion, and the cognitive sphere. Such deviations involve a greater focus on inner feelings or psychedelic experiences, modifications in the formal aspects of thought, and distortions in the perception of reality.

ASC are subdivided into the following types:

1) spontaneously occurring under normal (unstressed) or unusual (extreme) conditions;

2) artificially induced by exposure to psychoactive substances or procedures;

3) caused by psychotechnical techniques [3].

Exploring the structure and dynamics of the mechanisms responsible for the emergence of ASC has provided human sciences with numerous noteworthy findings.

Different models describing ASC have been shown in contemporary psychology [4–6].

Russian psychology provides quite a detailed description of ASC [7–9]. The goal of psychological research is to explore ASC by analyzing the factors that link their origins and mechanisms of manifestation, potentially influencing the personal transformation of an individual [10–13].

The study of personality through linguistic manifestations is also of great interest for literary studies. The monograph “Psychological Literary Studies. The text as a reflection of inner worlds of author and reader” by V.P. Belyanin reveals the possibility of identifying typological personality traits through psycholinguistic methods [14].

Despite the detailed development of modules describing ASC, the module of language ability and speech activity had not been found in the studies of prominent ASC theorists. D.L. Spivak suggested incorporating language and speech module into ASC theory on a large scale [15]. The main objective was to identify linguistic markers for shifts in the functional state of human mind and to uncover fundamental ways of how language supports altered functional mental states [16]. D.L. Spivak defined the core principles of linguistic support for modified human functional states. The author created a classification of linguistic indicators of main types of changes in the mental state of an individual. He also developed a method of psycholinguistic assessment of ASC followed by statistic-linguistic analysis and revealed alterations of speech production mechanisms in ASC [17].

Linguistic representation of human ASC in an exogenous state (when exposed to extreme external conditions) is characterized by several features [18].

Summing up, we can say that ASC linguistics has been formed. Investigations of linguistic features have been conducted at various levels, including phonology, lexis, morphology, syntax, and semantics. Structural forms of speech in ASC have been outlined, characteristics of syntactic changes have been given, a method for detecting basic modifications in affective syntax has been devised, and a

classification of primary structural and semantic features of emotional speech has been offered [16; 18–19].

However, a complex description of human communicative behavior in ASC was missing. As communication tasks are accomplished through a blend of verbal and non-verbal communication channels, the study of non-verbal behavior is of vital importance. In recent decades, Russian scholars have been studying non-verbal communication from linguistic perspective [20–22]. Hence, the need to study non-verbal behavior in ASC research is evident.

As we are going to analyze the peculiarities of non-verbal behavior of the characters from Dostoevsky's novel "Crime and Punishment", we should say a few words about the great author. As we know, Dostoevsky's realism is the sort of realism based not on cognition but on "penetration" [24]. In this regard, Dostoevsky's researcher V. Kirpotin wrote that "Dostoevsky had the seeming capacity to visualize directly someone else's psyche. He looked into someone else's soul as if equipped with a magnifying glass that permitted him to detect the subtlest nuances, to follow the most inconspicuous modulations and transitions in the inner life of a man. Dostoevsky, as if passing over the external barriers, observes directly the psychological processes taking place in a man, and fixes them on paper..." (V. Kirpotin, cited from [23. P. 22]).

The aim of this article is to find out what types of non-verbal behavior are peculiar for a person who is in ASC and what their language correlates are. The classification includes controlled non-verbal components of communication, uncontrolled ones related to human vegetative activity (psychophysiological reactions), and specific non-verbal actions.

Methods and Procedure

The study uses a comparative approach that starts by identifying non-verbal behavior in the source text, for which we use the above-mentioned classification. The analysis focuses only on the protagonist's non-verbal behavior, considering three categories: controlled and uncontrolled non-verbal behavior, and specific non-verbal actions. These encompass gestures, postures, facial expressions, eye movements, paralanguage, haptics (touch as a communicative act), proxemics (the use of personal space), and psychophysiological reactions.

After identifying the non-verbal cues in the Russian source text, they were compared with their equivalents in the English translation. When discussing the examples, the description of the non-verbal behavior in the Russian source text goes first, followed by its matching sentence in English.

Data Source

This research is focused on exploring how non-verbal human behavior in ASC cases is reflected in a literary text. The primary source for the study is the novel *Crime and Punishment* by F.M. Dostoyevsky and its English translation.

Controlled Non-Verbal Behavior and Its Language Correlates

The main character of the novel, R. Raskolnikov (during the time when he is planning, executing, and reflecting on his criminal action) is in an altered state of consciousness. According to experts in the field of psychiatry, Raskolnikov has a character trait known as schizoid accentuation. It occupies an intermediate position between the usual characteristics of the average, normal personality and schizoid psychopathy, which is a disease marked by episodes of exacerbation leading to social maladaptation and relatively calm periods [24].

When Raskolnikov finds himself fully absorbed in his theory, significant changes appear in his mental state: *"…но с некоторого времени он был в раздражительном и напряженном состоянии, похожем на ипохондрию. Он до того углубился в себя и уединился от всех, что боялся даже всякой встречи, не только встречи с хозяйкой. Он был задавлен бедностью; но даже стесненное положение перестало в последнее время тяготить его. Насущными делами своими он совсем перестал и не хотел заниматься"* (ПиН)[1]. In the English version: "*but for some time past he had been in an overstrained irritable condition, verging on hypochondria. He had become so completely absorbed in himself and isolated from everyone else that he dreaded meeting not only his landlady, but anyone at all. He was crushed by poverty, but even the anxieties of his position had recently ceased to weigh upon him. He had given up attending to matters of practical importance; he had lost all desire to do so*" (CaP)[2].

According to our research, certain types of controlled non-verbal behavior are characteristic of Raskolnikov. Let's consider phonation first of all: *"Или отказаться от жизни совсем! — **вскричал он вдруг в исступлении**, — послушно принять судьбу, как она есть, раз навсегда, и задушить в себе все, отказавшись от всякого права действовать, жить и любить!"* (ПиН)[3] In the English version: "*Or throw up life altogether!" **he cried suddenly, in a frenzy** — accept one's lot humbly as it is, once for all, and stifle everything in oneself, giving up all claim to activity, live, and love!*" (CaP)[4] In this part of the text, Raskolnikov experiences the emotion of despair. He is extremely irritated, which is explicitly expressed by the meaning of the adverbial phrase "*v isstuplenii" (in a frenzy)*. Exclamatory sentences serve as syntactic markers that indicate the level of emotion.

In certain contexts, non-verbal cues can indicate the gradation of negative emotions: "*Э, черта еще не доставало — **бормотал он, скрыпя зубами**, — нет, это мне теперь… некстати… **Дура она**, — прибавил он громко. —*

[1] Dostoevskii F.M. Prestuplenie i nakazanie. M. : Martin, 2021 (Достоевский Ф.М. Преступление и наказание. М. : Мартин, 2021).
[2] Dostoevsky, F. (2000). *Crime and Punishment*. Ware: Wordsworth Editions Limited.
[3]Dostoevskii F.M. Prestuplenie i nakazanie. M. : Martin, 2021 (Достоевский Ф.М. Преступление и наказание. М. : Мартин, 2021).
[4] Dostoevsky, F. (2000). *Crime and Punishment*. Ware: Wordsworth Editions Limited.

Я сегодня к ней зайду, поговорю" (ПиН)[5]. In the English version: "*The devil, that's the last straw,"* ***he muttered, grinding his teeth****, "no, that would not suit me… just now.* ***She is a fool****," he added aloud. "I'll go and talk to her today*" (CaP)[6]. Here Raskolnikov's verbal and non-verbal behavior conveys his annoyance. Phonation is represented by the past tense of the phonosemantic verb *bormotat' (muttered)*. Further, this emotion is transformed into the emotion of anger, which the author conveys with the help of the invective — "*Dura ona*" (*She is a fool*), and the commissive speech act "*Ya segodnya… pogovoryu*" (*I'll… talk to her*). There is also an increase in phonation, which is described by the verb-adverb phrase "*pribavil gromko*" (*added aloud*).

Facial expressions are commonly observed, with smiles being the most prevalent example that conveys both positive and negative emotions. Such situations are particularly interesting because they are often incongruent. "*…но когда он кончил, оно [лицо] было бледно,* ***искривлено судорогой****, и* ***тяжелая, желчная, злая улыбка змеилась по его губам***" (ПиН)[7]. In the English version: *"…but when he finished it, his face was pale and* ***distorted*** *and* ***a bitter, wrathful and malignant smile was on his lips***" (CaP)[8]. Raskolnikov's negative reaction is determined by the meaning of the adjectives: "*tyazhelaya, zhelchnaya, zlaya*" (***wrathful and malignant***). Both metaphoric use of the verb "*zmeit'sya*" (which is not found in the English version) and the participial construction "*iskrivleno sudorogoi*" (***distorted***) contribute to the portrayal of the character's negative emotional state via non-verbal signals.

Raskolnikov's pantomimic non-verbal cues also mark his mental state, which is characteristic of people with altered states of consciousness. "*Чай-то от хозяйки, что ль? — спросил он,* ***медленно и с болезненным видом приподнимаясь на софе***" (ПиН)[9]. In the English version: "*From the landlady, eh?" he asked, slowly and with a sickly face sitting up on the sofa* (CaP)[10]. The change in Raskolnikov's posture is described here with the help of the verb *pripodnimat'sya* (sit up), indicating changes in the position of the body and the adverbs indicating the mode of action: *medlenno, s boleznennym vidom* (*slowly, with a sickly face*). The semantic center of the adverbial phrase is the adjective *boleznennyi*, which reveals the character's altered state of consciousness.

Eye behavior is undoubtedly crucial in depicting an altered state of consciousness, as the visual channel often reveals disruptions in cognitive activity: "*Молодой человек несколько раз припоминал потом это первое впечатление*

[5] Dostoevskii F.M. Prestuplenie i nakazanie. M. : Martin, 2021 (Достоевский Ф.М. Преступление и наказание. М. : Мартин, 2021).
[6] Dostoevsky, F. (2000). *Crime and Punishment*. Ware: Wordsworth Editions Limited.
[7] Dostoevskii F.M. Prestuplenie i nakazanie. M. : Martin, 2021 (Достоевский Ф.М. Преступление и наказание. М. : Мартин, 2021).
[8] Dostoevsky, F. (2000). *Crime and Punishment*. Ware: Wordsworth Editions Limited.
[9] Dostoevskii F.M. Prestuplenie i nakazanie. M. : Martin, 2021 (Достоевский Ф.М. Преступление и наказание. М. : Мартин, 2021).
[10] Dostoevsky, F. (2000). *Crime and Punishment*. Ware: Wordsworth Editions Limited.

и даже приписывал его предчувствию. Он ***беспрерывно взглядывал на чиновника****, конечно, и потому еще, что и* ***сам тот упорно смотрел на него****, и видно было, что тому очень хотелось начать разговор. На остальных же, бывших в распивочной, не исключая и хозяина, чиновник смотрел как-то привычно и даже со скукой, а вместе с тем и с оттенком некоторого высокомерного пренебрежения, как бы на людей низшего положения и развития, с которыми нечего ему говорить*" (ПиН)[11]. In the English version: "*The young man often recalled this impression afterwards, and even ascribed it to presentiment. He* ***looked repeatedly at the clerk****, partly no doubt because the latter* ***was staring persistently at him****, obviously anxious to enter into conversation. At the other persons in the room, including the tavern-keeper, the clerk looked as though he were used to their company, and weary of it, showing a shade of condescending contempt for them as persons of station and culture inferior to his own, with whom it would be useless for him to converse*" (CaP)[12]. Indication of the mode of action expressed by the verbal phrase *vzglyadyvat' bespreryvno (persistently)* provides a clue to the state of Raskolnikov's inner world. It becomes clear that he is experiencing the emotion of anxiety. It should also be noted that, if one of the communicators uses eye-related non-verbal cues denoting gaze, then in the other communicator's reactive behavior there is an explicit expression of the emotional state (excitement), which is also transmitted via the visual channel: [*Porfirii Petrovich] i sam tot uporno smotrel na nego → [Raskol'nikov] bespreryvno vzglyadyval na chinovnika (*[*Porfirii Petrovich]* ***was staring persistently at him*** *→ [Raskol'nikov] looked repeatedly at the clerk)*. In such cases there is cause-and-effect relationship: one communicator's eye-related non-verbal behavior provokes emotional state of the other communicator.

Raskolnikov's breathing rhythm disorder is indicated by respiratory non-verbal cues: *"Слава богу, это только сон! — сказал он, садясь поддеревом и* ***глубоко переводя дыхание****. — Но что это? Уж не горячка ли во мне начинается: такой безобразный сон!"* (ПиН)[13] In the English version: "*Thank God, that was only a dream," he said, sitting down under a tree and* ***drawing deep breaths****. "But what is it? Is it some fever coming on? Such a hideous dream!*" (CaP)[14] In the above example, respiratory non-verbal cues are expressed by an adverbial construction "*gluboko perevodya dykhanie" (drawing deep breaths)*.

Uncontrolled Non-Verbal Behavior and Its Language Correlates

Uncontrolled non-verbal cues, also known as psychophysiological reactions, make an important part of Raskolnikov's behavior. People in altered state of

[11] Dostoevskii F.M. Prestuplenie i nakazanie. M. : Martin, 2021 (Достоевский Ф.М. Преступление и наказание. М. : Мартин, 2021).
[12] Dostoevsky, F. (2000). *Crime and Punishment*. Ware: Wordsworth Editions Limited.
[13] Dostoevskii F.M. Prestuplenie i nakazanie. M. : Martin, 2021 (Достоевский Ф.М. Преступление и наказание. М. : Мартин, 2021).
[14] Dostoevsky, F. (2000). *Crime and Punishment*. Ware: Wordsworth Editions Limited.

consciousness exhibit different types of psychophysiological reactions: tremor, tear shedding, etc. (for more details on classification of psychophysiological reactions see [22]).

Since the number of psychophysiological reactions in the novel "Crime and Punishment" significantly exceeds the number of other types of non-verbal behavior displayed by Raskolnikov, it is essential to analyze this specific form of behavior in detail. Psychophysiological reactions are interpreted in medicine as a human body's vegetative response to intense emotions [25]. We are going to investigate the evolution of Raskolnikov's emotions via his psychophysiological reactions. We assume that the starting point of this process is a person's thought that is being formed in his consciousness (cognitive level).

In *Crime and Punishment* Raskolnikov's idea centers on his famous theory according to which people can be divided into two categories: "inferior people", that is, ordinary, mediocre people (Raskolnikov defines them as "trembling creatures") and people "who have the gift or the talent to utter a new word", i.e., outstanding people, "Napoleons" who make the world go round. Before committing the murder of an old pawnbroker and her sister, Raskolnikov hopes that he belongs to the second category, to extraordinary individuals. He believes that he is a person who can say a "new word" in the world around him. Raskolnikov's idea takes shape when he writes and publishes an article on the subject in a journal, which draws people's attention.

Even before Raskolnikov commits the crime (according to his plan to murder the old pawnbroker) he is experiencing anxiety, which is evident in his general state and is marked by psychophysiological reactions whose linguistic representation is to be found on the initial pages of the novel: "***С замиранием сердца и нервною дрожью*** *подошел он к преогромнейшему дому, выходившему одною стеной на канаву, а другою в-ю улицу*" (ПиН)[15]. In the English version: "***With a sinking heart and a nervous tremor****, he went up to a huge house which on one side looked on to the canal, and on the other into the street*" (CaP)[16]. In this fragment of the text we observe a psychophysiological reaction of heartbeat and trembling, described by the adverbial construction "*s zamiraniem serdca i nervnoju drozh'ju*".

The moment Raskolnikov decides to validate his theory through action, his transition to an altered state of consciousness begins, which is evident in the psychophysiological responses he experiences. Verbal descriptions of psychophysiological reactions provide a comprehensive reflection of cognitive processes of an individual in an altered state of consciousness: "*Он даже шел теперь делать пробу своему предприятию, и с каждым шагом* ***волнение его возрастало все сильнее и сильнее***" (ПиН)[17]. In the English version: "*He was*

[15] Dostoevskii F.M. Prestuplenie i nakazanie. M. : Martin, 2021 (Достоевский Ф.М. Преступление и наказание. М. : Мартин, 2021).
[16] Dostoevsky, F. (2000). *Crime and Punishment*. Ware: Wordsworth Editions Limited.
[17] Dostoevskii F.M. Prestuplenie i nakazanie. M. : Martin, 2021 (Достоевский Ф.М. Преступление и наказание. М. : Мартин, 2021).

positively going now for a "rehearsal" of his project, and at every step ***his excitement grew more and more violent***" (CaP)[18]. The language correlates of the psychological reaction show intensity of his emotions, as indicated by the meaning of the verb *vozrastalo* (*grew*) and the adverb *sil'nee* (*more and more violent*), used twice.

In the description of the next phase of Raskolnikov's altered state of consciousness, we find a very vivid language representation of the psychophysiological reaction of stupor: "*Сходя по лестнице, он* ***несколько раз даже останавливался, как будто чем-то пораженный***" (ПиН)[19]. In the English version: "*As he went down the stairs,* ***he even stopped short, two or three times, as though suddenly struck by some thought***" (CaP)[20]. Language correlates of this psychophysiological reaction include a comparative structure with a verb indicating the repeated stopping of movement: *несколько раз даже останавливался (he even stopped short, two or three times*).

The following example is also a convincing illustration of Raskolnikov's altered state of consciousness: "*Он шел по тротуару* ***как пьяный, не замечая прохожих и сталкиваясь с ними***" (ПиН)[21]. In the English version: "*He walked along the pavement* ***like a drunken man, regardless of the passers-by, and jostling against them***" (CaP)[22]. Linguistic markers of ASC here are: the comparative construction "*как пьяный*" (*like a drunken man*) and the participial construction (*ne zamechaja prohozhih i stalkivajas' s nimi* (*regardless of the passers-by, and jostling against them*)). According to experts, such behavior is a clear indication of mental automatism, a characteristics of Kandinsky-Clerambault syndrome, where individuals experience a sense of external control of all feelings, movements and thoughts [24].

Altered state of consciousness is revealed in Raskolnikov's reactive behavior: he begins to avoid people and even experiences extremely negative emotions when interacting with them: "*Он решительно ушел от всех, как черепаха в свою скорлупу, и даже лицо служанки ... возбуждало в нем* ***желчь и конвульсии****"* (ПиН)[23]. In the English version: "*He had got completely away from every one, like a tortoise in its shell, and even the sight of the servant girl … made him* ***writhe with nervous irritation***" (CaP)[24]. Raskolnikov's psychophysiological reaction is expressed by nominative phrase "*желчь и конвульсии"* (missing in the English version), which marks malfunctioning of Raskolnikov's vegetative system.

[18] Dostoevsky, F. (2000). *Crime and Punishment*. Ware: Wordsworth Editions Limited.
[19] Dostoevskii F.M. Prestuplenie i nakazanie. M. : Martin, 2021 (Достоевский Ф.М. Преступление и наказание. М. : Мартин, 2021).
[20] Dostoevsky, F. (2000). *Crime and Punishment*. Ware: Wordsworth Editions Limited.
[21] Dostoevskii F.M. Prestuplenie i nakazanie. M. : Martin, 2021 (Достоевский Ф.М. Преступление и наказание. М. : Мартин, 2021).
[22] Dostoevsky, F. (2000). *Crime and Punishment*. Ware: Wordsworth Editions Limited.
[23] Dostoevskii F.M. Prestuplenie i nakazanie. M. : Martin, 2021 (Достоевский Ф.М. Преступление и наказание. М. : Мартин, 2021).
[24] Dostoevsky, F. (2000). *Crime and Punishment*. Ware: Wordsworth Editions Limited.

Of special interest is the part of the text where the distinguished writer traces the transition of the character in ASC in a manner that was uniquely his own: idea — emotion — physical feelings. Raskolnikov's initial thought, which originated in his mind, that is on the cognitive level, triggers the emotion of fear, which is reflected in his physical state: *"Вдруг он вздрогнул: одна, тоже вчерашняя,* ***мысль*** *пронеслась в его голове... Ему* ***стукнуло в голову, и потемнело в глазах****"* (ПиН)[25]. In the English version: "*He gave a sudden start; another* ***thought****, that he had had yesterday, slipped back into his mind. ...* ***He felt a hammering in his head, and there was a darkness before his eyes***" (CaP)[26].

As the plot is developing, the number of language correlates of the key psychophysiological reactions reflecting ASC rises, which is shown in the examples given below: *"****Сердце его страшно билось****"* (ПиН)[27]. In the English version: "***His heart beat terribly***" (CaP)[28]. *"****Руки его тряслись****..."* (ПиН)[29]. In the English version: "***His hands shook***…" (CaP)[30]. *"...****Но сердце не переставало****. Напротив, как нарочно*, ***стучало сильней, сильней, сильней****..."* (ПиН)[31]. In the English version: "…*But* ***his heart did not leave off****. On the contrary, as though to spite him,* ***it throbbed more and more violently***" (CaP)[32].

Specific Non-Verbal Actions and Their Language Correlates

When discussing human non-verbal behavior in ASC situations, it is essential to consider specific non-verbal actions that elude interpretation by communicative partners, often resulting in communicative failures. Below is a bright illustration of the above-mentioned situation: "*Раскольников молча взял немецкие листки статьи, взял три рубля и, не сказав ни слова, вышел. Разумихин с удивлением поглядел ему вслед. Но дойдя уже до первой линии, Раскольников вдруг воротился, поднялся опять к Разумихину и, положив на стол и немецкие листы, и три рубля, опять-таки ни слова не говоря, пошел вон*" (ПиН)[33]. In the English version: "*Raskolnikov took the German sheets in silence, took the three roubles and without a word went out. Razumihin gazed after him in astonishment. But when Raskolnikov was in the next street, he turned back, mounted the stairs to*

[25] Dostoevskii F.M. Prestuplenie i nakazanie. M. : Martin, 2021 (Достоевский Ф.М. Преступление и наказание. М. : Мартин, 2021).
[26] Dostoevsky, F. (2000). *Crime and Punishment*. Ware: Wordsworth Editions Limited.
[27] Dostoevskii F.M. Prestuplenie i nakazanie. M. : Martin, 2021 (Достоевский Ф.М. Преступление и наказание. М. : Мартин, 2021).
[28] Dostoevsky, F. (2000). *Crime and Punishment*. Ware: Wordsworth Editions Limited.
[29] Dostoevskii F.M. Prestuplenie i nakazanie. M. : Martin, 2021 (Достоевский Ф.М. Преступление и наказание. М. : Мартин, 2021).
[30] Dostoevsky, F. (2000). *Crime and Punishment*. Ware: Wordsworth Editions Limited.
[31] Dostoevskii F.M. Prestuplenie i nakazanie. M. : Martin, 2021 (Достоевский Ф.М. Преступление и наказание. М. : Мартин, 2021).
[32] Dostoevsky, F. (2000). *Crime and Punishment*. Ware: Wordsworth Editions Limited.
[33] Dostoevskii F.M. Prestuplenie i nakazanie. M. : Martin, 2021 (Достоевский Ф.М. Преступление и наказание. М. : Мартин, 2021).

Razumihin's again and laying on the table the German article and the three roubles, went out again, still without uttering a word" (CaP)[34]. Raskolnikov agrees to do a translation offered to him by his friend for a fee, then leaves without thanking him or saying goodbye, which, in turn, causes confusion on the part of Razumikhin — "*Разумихин* ***с удивлением поглядел*** *ему вслед*" (*Razumihin* ***gazed after him in astonishment***.). Razumikhin's emotion of surprise is justified: Raskolnikov first agrees to translate the text in order to improve his financial situation but later returns and rejects Razumikhin's offer without any explanation.

Raskolnikov's Emotions in ASC and Their Linguistic Expression

The novel contains instances of direct linguistic representation of emotions, with words explicitly identifying various feelings, sometimes modified by an attribute: gloomy excitement, fear, wonder, terror. In description of the character's non-verbal behavior indirect linguistic emotionality prevails. Dostoevsky depicts various types of non-verbal behavior: controlled, uncontrolled (psychophysiological reactions), specific non-verbal actions. Raskolnikov's emotional spectrum is exclusively negative, with the main emotion being fear and the accompanying emotions of surprise and disgust. In indirect linguistic emotionality, emotions are represented implicitly, but the reader has no difficulty in deciphering them. For example, in the text of the novel, fear/horror at the thought of committing the crime is represented by the description of the psychophysiological reaction of heartbeat and changes in the respiratory system: *"…а сердце все билось, стукало так, что ему дышать стало тяжело*" (ПиН)[35]. In the English translation: "*and his heart kept beating and thumping so that he could hardly breathe*" (CaP)[36].

Controlled non-verbal expressions of fear, involving body movements, are also significant: "*…и вдруг* ***вскочил****, совершенно опомнившись,* ***как будто кто его сорвал с дивана***" (ПиН)[37]. In the English version: "…*suddenly* ***jumped up*** *wide awake* ***as though some one had pulled him off the sofa***" (CaP)[38].

Fear makes Raskolnikov move in a particular way: "*…и стал* ***сходить*** *вниз свои тринадцать ступеней,* ***осторожно, неслышно, как кошка***" (ПиН)[39]. In the English version: "…*and began to* ***descend*** *his thirteen steps* ***cautiously, noiselessly, like a cat***" (CaP)[40].

[34] Dostoevsky, F. (2000). *Crime and Punishment*. Ware: Wordsworth Editions Limited.
[35] Dostoevskii F.M. Prestuplenie i nakazanie. M. : Martin, 2021 (Достоевский Ф.М. Преступление и наказание. М. : Мартин, 2021).
[36] Dostoevsky, F. (2000). *Crime and Punishment*. Ware: Wordsworth Editions Limited.
[37] Dostoevskii F.M. Prestuplenie i nakazanie. M. : Martin, 2021 (Достоевский Ф.М. Преступление и наказание. М. : Мартин, 2021).
[38] Dostoevsky, F. (2000). *Crime and Punishment*. Ware: Wordsworth Editions Limited.
[39] Dostoevskii F.M. Prestuplenie i nakazanie. M. : Martin, 2021 (Достоевский Ф.М. Преступление и наказание. М. : Мартин, 2021).
[40] Dostoevsky, F. (2000). *Crime and Punishment*. Ware: Wordsworth Editions Limited.

Dostoevsky draws the reader's attention to the peculiarities of Raskolnikov's movements. Here is the description of his manner of walking to the old pawnbroker. "*Он **шел** дорогой **тихо и степенно**, не торопясь, чтобы не подать каких подозрений*" (ПиН)[41]. In the English version: "*He **walked** along **quietly and sedately**, without hurry, to avoid awakening suspicion*" (CaP)[42]. Thus Raskolnikov tries to conceal his emotions.

In analyzing Raskolnikov's mental processes, the author also notes his attempts to control his behavior: "*Переведя дух и прижав рукой стукавшее сердце, … он стал **осторожно и тихо** подниматься на лестницу, **поминутно прислушиваясь***" (ПиН)[43]. In the English version: "*Drawing a breath, pressing his hand against his throbbing heart, …he began **softly and cautiously** ascending the stairs, **listening every minute***" (CaP)[44]. *"Не подождать ли еще… **пока сердце перестанет**?.."* (ПиН)[45]. In the English version: "*Had I better wait a little longer… **till my heart leaves off thumping***?" (CaP)[46]

However, concealing of emotions is not always successful because psychophysiological reactions are uncontrollable and the recipient (the old pawnbroker) has no difficulty in deciphering the emotion: "*Да чтой-то вы какой **бледный**? Вот и **руки дрожат**!*" (ПиН)[47] In the English version: "*But how **pale** you are, to be sure … and your **hands are trembling too**?*" (CaP)[48]

Of no less importance is the type of non-verbal behavior that marks fear experienced by Raskolnikov — a specific non-verbal action that takes place in Raskolnikov's reactive behavior even in an ordinary situation, such as hearing a scream in the yard: "*Он **бросился к двери, прислушался, схватил шляпу**…*" (ПиН)[49]. In the English version: "***He rushed to the door, listened, caught up his hat***…" (CaP).

We assume that there exists cause-and-effect relationship between the character's ASC and its manifestations with the ASC being the cause, and the specific non-verbal action: *схватил шляпу* (*caught up his hat*) being the effect.

With the development of the plot, negative emotions escalate and become more intense: "*Страх охватывал его все больше и больше …*" (ПиН)[50]. In the English

[41] Dostoevskii F.M. Prestuplenie i nakazanie. M. : Martin, 2021 (Достоевский Ф.М. Преступление и наказание. М. : Мартин, 2021).
[42] Dostoevsky, F. (2000). *Crime and Punishment*. Ware: Wordsworth Editions Limited.
[43]Dostoevskii F.M. Prestuplenie i nakazanie. M. : Martin, 2021 (Достоевский Ф.М. Преступление и наказание. М. : Мартин, 2021).
[44] Dostoevsky, F. (2000). *Crime and Punishment*. Ware: Wordsworth Editions Limited.
[45]Dostoevskii F.M. Prestuplenie i nakazanie. M. : Martin, 2021 (Достоевский Ф.М. Преступление и наказание. М. : Мартин, 2021).
[46] Dostoevsky, F. (2000). *Crime and Punishment*. Ware: Wordsworth Editions Limited.
[47] Dostoevskii F.M. Prestuplenie i nakazanie. M. : Martin, 2021 (Достоевский Ф.М. Преступление и наказание. М. : Мартин, 2021).
[48] Dostoevsky, F. (2000). *Crime and Punishment*. Ware: Wordsworth Editions Limited.
[49] Dostoevskii F.M. Prestuplenie i nakazanie. M. : Martin, 2021 (Достоевский Ф.М. Преступление и наказание. М. : Мартин, 2021).
[50] Dostoevskii F.M. Prestuplenie i nakazanie. M. : Martin, 2021 (Достоевский Ф.М. Преступление и наказание. М. : Мартин, 2021).

version: "*Fear gained more and more mastery over him*" (CaP). We observe a blending of strong negative emotions (fear, horror, disgust, fright), the central one being fear, explicated by a phraseological unit which serves as a language correlate of the psychophysiological reaction of trembling: "*Он* ***дрожал*** *как лист…*" (ПиН)[51]. In the English version: "*He was* ***shaking*** *like a leaf…*" (CaP).

Another key element in interpreting Raskolnikov's state is the increasing emotion of disgust that he experiences: "***Отвращение*** *особенно* ***поднималось и росло*** *в нем с каждою минутою*" (ПиН)[52]. In the English version: "***The feeling of loathing*** *especially* ***surged up within him and grew stronger*** *every minute*" (CaP).

The dynamics of the ASC and, consequently, the dynamics of the emotional state of the character are especially noteworthy in this context: "*Он был в полном уме, затмений и головокружений уже не было, но* ***руки все еще дрожали***" (ПиН)[53]. In the English version: "*He was in full possession of his faculties, free from confusion or giddiness, but* ***his hands were still trembling***" (CaP). The closer Raskolnikov gets to committing the crime, the higher the degree of his ASC: "*…он* ***сумасшествует****…*" (ПиН)[54]. In the English version: "…***he was mad***" (CaP). "*Он был* ***точно в бреду***" (ПиН)[55]. In the English version: "*He was* ***in a sort of delirium***" (CaP). "*Он* ***плохо теперь помнил себя; чем дальше, тем хуже***" (ПиН)[56]. In the English version: "*He was* ***only dimly conscious of himself now, and the farther he went the worse it was***" (CaP).

The final fragment of the first part of the novel reflects a collapse in Raskolnikov's cognitive processes, leading to a breakdown in his emotional state. The author gives a detailed description of the malfunctioning of Raskolnikov's brain: "*Он не спал, но* ***был в забытьи****… Клочки и отрывки каких-то мыслей так и кишели в его голове; но он ни одной не мог схватить, ни на одной не мог остановится****, несмотря даже на усилия…*" (ПиН)[57]. In the English version: *"…he did not sleep, but* ***sank into blank forgetfulness****… Scraps and shreds of thoughts were simply swarming in his brain, but he could not catch at one, he could not rest on one****, in spite of all his efforts…*" (CaP)[58]. In these examples, we observe the verbal representation of Raskolnikov's altered state of consciousness as he reaches the peak of his anxiety.

[51] Dostoevskii F.M. Prestuplenie i nakazanie. M. : Martin, 2021 (Достоевский Ф.М. Преступление и наказание. М. : Мартин, 2021).
[52] Ibid.
[53] Ibid.
[54] Ibid.
[55] Ibid.
[56] Ibid.
[57] Ibid.
[58] Dostoevsky, F. (2000). *Crime and Punishment*. Ware: Wordsworth Editions Limited.

Conclusions

The identification of uncontrolled non-verbal cues reflecting human vegetative activity (psychophysiological reactions), as well as specific non-verbal actions (not fitting into the generally accepted classification of non-verbal cues) along with controlled non-verbal cues, offers valuable insights into the non-verbal interactions of communicators in various communication situations. Focusing on the language correlates of Raskolnikov's non-verbal actions in F.M. Dostoevsky's novel "Crime and Punishment" enables us to uncover the distinctiveness of non-verbal behavior in ASC contexts. The research identified the predominant controlled non-verbal cues, psychophysiological reactions, specific non-verbal actions and determined the language correlates used to represent human non-verbal behavior in ASC.

In the situations under consideration, typical controlled non-verbal behavior includes phonation, facial expressions, eye contact, pantomime, and proxemics. The main types of psychophysiological reactions in cases of ASC, according to our research are trembling and heartbeat, the intensity of which marks the increasing anxiety of the character, as well as specific non-verbal actions that express the emotional state of the negative vector.

The portrayal of Raskolnikov's transition to the ASC (following the sequence of idea — emotion — physical actions) is mainly achieved through non-verbal signals. Raskolnikov's emotions are conveyed through detailed descriptions of his non-verbal behavior.

References

1. Ludwig, A.M. (1966). Altered States of Consciousness. *Archives of General Psychiatry*, 15, 225–234. https://doi.org/10.1001/archpsyc.1966.01730150001001
2. Bourguignon, E. (2001). Mental Illness, Culture, and Altered States of Consciousness: An Interdisciplinary Field of Research. In: *Personality, Culture, Ethnic Group: Modern Psychological Anthropology* (pp. 405–461). Moscow: Smysl. (In Russ.).
3. Spivak, L.I., & Spivak, D.L. (1996). Altered States of Consciousness: Typology, Semiotics, Psychophysiology. *Consciousness and Physical Reality*, *1*(4), 48–55. (In Russ.). EDN: XASOWR
4. Tart, C.T. (1975). *States of Consciousness*. NewYork: E.P. Dutton, Plume.
5. Martindale, C. (1981). *Cognition and Consciousness*. Homewood, Ill.: Dorsey Press.
6. Dittrich, A. (1979). Studies on Altered States of Consciousness in Normals. In: *Psychiatrische Universitaetsklinik Zuerich, Buighoelzli. Forschungsdirektion. Zehnjahrebericht 1969–1979* (pp. 188–197). Zuerich: Forschungsdirektion.
7. Kyugler, P. (2005). *Alchemy of Discourse: Image, Sound, and the Psyche*. Moscow: PER SE. (In Russ.)
8. Rossokhin, A.V., & Izmagurova, V.L. (2004). *Personality in Altered States of Consciousness in Psychoanalysis and Psychotherapy*. Moscow: Smysl. (In Russ.). EDN: QXKLCH
9. Rudnev, V. (2007). *Philosophy of Language and Semiotics of Madness: Selected works.* Moscow: Territoriya budushchego. (In Russ.). EDN: RAYWCJ
10. Kardash, S. (1998). *Altered States of Consciousness.* Donetsk: Stalker. (In Russ.). EDN: QXKLCH
11. Kucherenko, V.V., Petrenko, V.F., & Rossokhin, A.V. (1998). Altered States of Consciousness: Psychological Analysis. *Voprosy Psychologii*, 4, 70–77. (In Russ.).

12. Gordeeva, O.V. (2012). Main Problems and Directions of Theoretical Research in the Psychology of Altered States of Consciousness. In: *Altered States of Consciousness* (pp. 10–35). Moscow: Kogito-Tsentr. (In Russ.).
13. Kozlov, V.V. (2005). *Psychotechnologies of Altered States of Consciousness: Methods and Techniques.* Moscow: Izdatel'stvo Instituta psikhoterapii. (In Russ.). EDN: QXLUSB
14. Belyanin, V.P. (2016). *Psychological Literary Studies. Text as a Reflection of the Inner Worlds of the Author and the Reader.* Moscow: Genezis. (In Russ.). EDN: ZGJNWJ
15. Spivak, D.L. (1986). *Linguistics of Altered States of Consciousness*. Leningrad: Nauka. (In Russ.). EDN: QXLUSB
16. Spivak, D.L. (2000). *Altered States of Consciousness: Psychology and Linguistics*. Saint Petersburg: Yuventa; Filologicheskii f-t SPbGU. (In Russ.). EDN: ZGJNWJ
17. Spivak, D.L. (1998). *Linguistics of Altered States of Consciousness* [PhD thesis]. Saint Petersburg. (In Russ.). EDN: UNAICR
18. Sukhikh, S.A. (2006). Linguistic Representation of Altered States of Consciousness. *Language, Communication, and Social Environment* (Iss. 4. pp. 78–93). Voronezh: Voronezh State University. (In Russ.).
19. Sineokova, T.N. (2010). Linguistics of Altered States of Consciousness. *Yurislingvistika*, 10(10), 288–301. (In Russ.). EDN: WCLRXJ
20. Kreidlin, G.E (2002). *Nonverbal Semiotics: Body Language and Natural Language*. Moscow: Novoe literaturnoe obozrenie. (In Russ.). EDN: STOOBN
21. Romanov, A.A., & Sorokin, Yu.A. (2004). *Somaticon: Aspects of Nonverbal Semiotics*. Moscow: IYa RAN, TvGU. (In Russ.). EDN: JYBDBB
22. Kartashkova, F.I. (2015). *Psychophysiological Reactions of Humans and Their Reflection in the English Language*. Ivanovo: Ivanovo State University. (In Russ.). EDN: VZHCOF
23. Bakhtin, M.M. (1963). *Problems of Dostoevsky's Poetics*. Moscow: Sovetskii pisatel'. (In Russ.). EDN: VQMUQJ
24. Petrushin, A. (2018). Analysis of the Patient's Medical History Rodion Romanovich Raskolnikov. *Mundo Eslavo*, 17, 165–181. (In Russ.). EDN: VZHCOF
25. Bogolepov, N.K. (1971). *Clinical Lectures on Neuropathology.* Moscow: Meditsina. (In Russ.). EDN: VQMUQJ

Список литературы

1. *Ludwig A.M.* Altered States of Consciousness // Archives of General Psychiatry. 1966–1967. Vol. 15–16. Iss. Cumulative Index. P. 225–234 https://doi.org/10.1001/archpsyc.1966.01730150001001
2. *Бургиньон Э.* Психические заболевания, культура и измененные состояния сознания: междисциплинарное поле исследования // Личность, культура, этнос: современная психологическая антропология / под ред. А.А Белик. М. : Смысл, 2001. С. 405–461.
3. *Спивак Л.И., Спивак Д.Л.* Изменённые состояния сознания: типология, семиотика, психофизиология // Сознание и физическая реальность. 1996. Т. 1. № 4. С. 48–55. EDN: XASOWR
4. *Tart C.T.* States of Consciousness. N.Y. : E.P. Dutton, Plume, 1975.
5. *Martindale C.* Cognition and Consciousness. Homewood, Ill.: Dorsey Press, 1981.
6. *Dittrich A*. Studies on Altered States of Consciousness in Normals // Psychiatrische Universitaetsklinik Zuerich, Buighoelzli. Forschungsdirektion. Zehnjahrebericht 1969–1979. Zuerich : Forschungsdirektion, 1979. P. 188–197.
7. *Кюглер П.* Алхимия дискурса: образ, звук и психическое. М. : ПЕР СЭ, 2005.
8. *Россохин А.В., Измагурова В.Л.* Личность в измененных состояниях сознания в психоанализе и психотерапии. М. : Смысл, 2004. EDN: QXKLCH
9. *Руднев В.* Философия языка и семиотика безумия: Избранные работы. М. : Территория будущего, 2007. EDN: RAYWCJ

10. *Кардаш С.* Измененные состояния сознания. Донецк : Сталкер, 1998. EDN: QXKLCH
11. *Кучеренко В.В., Петренко В.Ф., Россохин А.В.* Измененные состояния сознания: психологический анализ // Вопросы психологии. 1998. № 3. С. 70–77. EDN: RAYWCJ
12. *Гордеева О.В.* Основные проблемы и направления теоретических исследований в психологии измененных состояний сознания // Измененные состояния сознания / под ред. О.В. Гордеевой. М. : Когито-Центр, 2012. С. 10–35.
13. *Козлов В.В.* Психотехнологии измененных состояний сознания: методы и техники. М. : Изд-во Ин-та психотерапии, 2005. EDN: QXLUSB
14. *Белянин В.П.* Психологическое литературоведение. Текст как отражение внутренних миров автора и читателя. М. : Генезис, 2016. EDN: ZGJNWJ
15. *Спивак Д.Л.* Лингвистика измененных состояний сознания. Л. : Наука, 1986. EDN: QXLUSB
16. *Спивак Д.Л.* Измененные состояния сознания: психология и лингвистика. СПб. : Ювента; Филол. ф-т СПбГУ, 2000. EDN: ZGJNWJ
17. *Спивак Д.Л.* Лингвистика измененных состояний сознания : дисс. … д-ра филол. наук. СПб., 1998. EDN: UNAICR
18. *Сухих С.А.* Языковая репрезентация измененных состояний сознания // Язык, коммуникация и социальная среда: сбор. научных трудов. Вып. 4. Воронеж: Воронежский гос. университет, 2006. С. 78–93.
19. *Синеокова Т.Н.* Лингвистика измененных состояний сознания // Юрислингвистика. 2010. № 10 (10). С. 288—301. EDN: WCLRXJ
20. *Крейдлин Г.Е.* Невербальная семиотика: Язык тела и естественный язык. М. : Новое литературное обозрение, 2002. EDN: STOOBN
21. *Романов А.А., Сорокин Ю.А.* Соматикон: Аспекты невербальной семиотики. М. : ИЯ РАН, ТвГУ, 2004. EDN: JYBDBB
22. *Карташкова Ф.И.* Психофизиологические реакции человека и их отражение в английском языке. Иваново : Ивановский государственный университет, 2015. EDN: VZHCOF
23. *Бахтин М.М.* Проблемы поэтики Достоевского. М. : Советский писатель, 1963. EDN: VQMUQJ
24. *Петрушин А.* Анализ истории болезни пациента Р.Р. Раскольникова // Mundo Eslavo. 2018. Vol. 17. P. 165–181.
25. *Боголепов Н.К.* Клинические лекции по невропатологии. М. : Медицина, 1971.

Information about the authors:

Faina I. Kartashkova, Dr. Sc. (Philology), Professor, Professor of the Department of Foreign Philology Institute of Humanities, Honorary Worker of Higher Professional Education of the Russian Federation, Head of NOC “Laboratory of Communicative Human Behavior”, Ivanovo State University (39 Ermaka Str., Ivanovo, Russian Federation, 153025); *Research interests:* theory of speech nomination, theory of speech activity, cognitive linguistics, non-verbal communication, communicative pragmatics, semantics; *e-mail:* kartashkova@rambler.ru
ORCID: 0000-0001-7095-9143; SPIN-code: 1031-6512; Scopus Author ID: 57204361085.
Vera V. Ganina, PhD in Philology, Associate Professor of the Department of Foreign Languages and Linguistics, Ivanovo State University of Chemistry and Technology (27 Sheremetev Av., Ivanovo, Russian Federation, 153000); *Research interests*: Linguistics, Germanic languages, psycholinguistics; *e-mail*: vvganina@mail.ru
ORCID: 0000-0002-2263-5627; SPIN-code: 9717-6971; AuthorID: 783556.

Сведения об авторах:
Карташкова Фаина Иосифовна, доктор филологических наук, профессор, профессор кафедры зарубежной филологии Института гуманитарных наук, почетный работник высшего профессионального образования РФ, руководитель НОЦ «Лаборатория коммуникативного поведения человека», Ивановский государственный университет (153025, Российская Федерация, Иваново, ул. Ермака, д. 39); *сфера научных интересов*: теория речевой номинации, теория речевой деятельности, когнитивная лингвистика, невербальная коммуникация, коммуникативная прагматика, семантика; *e-mail*: kartashkova@rambler.ru
ORCID: 0000-0001-7095-9143; SPIN-код: 1031-6512; Scopus Author ID: 57204361085.
Ганина Вера Владимировна, кандидат филологических наук, доцент кафедры иностранных языков и лингвистики Гуманитарного института, Ивановский государственный химико-технологический университет (153000, Российская Федерация, г. Иваново, пр. Шереметевский, д. 27); *сфера научных интересов*: языкознание, германские языки, психолингвистика; *e-mail*: vvganina@mail.ru
ORCID: 0000-0002-2263-5627; SPIN-код: 9717-6971; AuthorID: 783556.

RUDN Journal of Language Studies, Semiotics and Semantics
2313-2299 (Print), ISSN 2411-1236 (Online)
Вестник Российского университета дружбы народов.
Серия: ТЕОРИЯ ЯЗЫКА. СЕМИОТИКА. СЕМАНТИКА

2026 Vol. 17 No. 2 813–835
http://journals.rudn.ru/semiotics-semantics

DOI: 10.22363/2313-2299-2026-17-2-813-835
EDN: NHXXYL

Научная статья / Research Article

Трансформации прецедентных феноменов в дискурсе современной сетевой литературы

В.Е. Анисимов[1] ✉, **Е.Д. Анисимова**[2], **Р.А. Савастенко**[3]

[1]Московский государственный институт международных отношений (университет) МИД РФ, *Москва, Российская Федерация*
[2]Независимый исследователь, *Москва, Российская Федерация*
[3]Российский университет дружбы народов, *Москва, Российская Федерация*
✉anisimov.vladislav.95@mail.ru

Аннотация. Сетевая литература представляет собой многогранный и малоизученный феномен, особый тип дискурса, в рамках которого прослеживаются черты как литературно-художественного дискурса, так и современного интернет-дискурса. Как и в традиционных художественных текстах, в текстах сетевой литературы присутствуют прецедентные феномены, функцией которых становится моделирование вторичной авторской реальности произведения. В связи с жанровыми и стилистическими особенностями произведений современной сетевой литературы, наблюдаемые в них прецедентные феномены претерпевают определенные трансформационные изменения. Целью данной статьи становится выявление, анализ и описание стратегий трансформации прецедентных феноменов в произведениях современной сетевой литературы жанра фэнтези/фантастика, а также смешанных жанров. Материалами исследования стали 150 произведений сетевой литературы, принадлежащие 21 циклу книг и написанные в жанрах фэнтези/фантастика. На основе анализа составленного авторами корпуса текстов были выявлены ключевые стратегии трансформации прецедентных феноменов, а также описаны особенности их применения относительно трансформации прецедентных феноменов-онимов и прецедентных высказываний как наиболее распространенных типов прецедентных феноменов. Основными методами анализа стали корпусный анализ текста, контент-анализ, лингвистический анализ текста, метод сплошной выборки, методы математической статистики и метод интертекстуального анализа. Результатами исследования стало выявление и описание 5 основных стратегий трансформации прецедентных феноменов-онимов в произведениях сетевой литературы, а также особенности их применения в комплексных прецедентных феноменах-высказываниях, и частотность применения данных стратегий. В качестве основной функции прецедентных феноменов в текстах сетевой литературы определена функция формирования категории вторичной (авторской) реальности.

Ключевые слова: категория реальности, вторичная реальность, авторский неологизм, трансформация

Вклад авторов: равнозначен на всех этапах исследования и подготовки статьи к публикации.

© Анисимов В.Е., Анисимова Е.Д., Савастенко Р.А., 2026
This work is licensed under a Creative Commons Attribution-NonCommercial 4.0 International License
https://creativecommons.org/licenses/by-nc/4.0/legalcode

Конфликт интересов: авторы заявили об отсутствии конфликтов интересов.

История статьи: дата поступления: 20.01.2026; дата приема в печать: 10.02.2026.

Для цитирования: *Анисимов В.Е.*, *Анисимова Е.Д., Савастенко Р.А.* Трансформации прецедентных феноменов в дискурсе современной сетевой литературы // Вестник Российского университета дружбы народов. Серия: Теория языка. Семиотика. Семантика. 2026. Т. 17. № 2. С. 813–835. https://doi.org/10.22363/2313-2299-2026-17-2-813-835 EDN NHXXYL

Transformation Strategies of Precedent Phenomena in the Discours of Modern Web-Literature

Vladislav E. Anisimov[1], Ekatherina D. Anisimova[2], Roman A. Savastenko[3]

[1]MGIMO University, *Moscow*, *Russian Federation*
[2] Independent researcher, *Moscow*, *Russian Federation*
[3]RUDN University, *Moscow, Russian Federation*
✉anisimov.vladislav.95@mail.ru

Abstract. Web-literature is a multifaceted and underinvestigatedp henomenon, a distinct discourse type that exhibits features of both literary discourse and contemporary internet discourse. As in traditional literary texts, web-literature texts contains precedent phenomena whose function is to model the work's secondary (authorial) reality. Due to the genre and stylistic characteristics of contemporary web-literature, the precedent phenomena observed within them undergo certain transformational changes. The purpose of this article is to identify, analyze, and describe strategies for transforming precedent phenomena in contemporary web-literature works of fantasy/science fiction, as well as hybrid genres. The study's materials include 150 works of web-literature, belonging to 21 book series and written in the fantasy/science fiction genres. Based on an analysis of a text corpus compiled, we identified key strategies for transforming precedent phenomena and discribed the specifics of their application to the transformation of precedent onyms and precedent statements, the most common types precedent phenomena. The primary analytical methods used included corpus analysis, content analysis, linguistic text analysis, continuous sampling, mathematical statistics, and intertextual analysis. The main results of the study were the describing of five key strategies for transforming precedent-based onymic phenomena in online literature, as well as the specifics of their application in complex precedent statements and the frequency of their use. The primary function of precedent phenomena in web-literature is the formation of a category of secondary (authorial) reality.

Keywords: web-literature, category of reality, secondary reality, author's neologism, precedent phenomenon, transformation

Authors' contribution: The authors contributed equally to this article and preparation of the article text**.**

Conflict of interests: The authors stated that there were no conflicts of interest.

Article history: Received: 20.01.2026; Accepted: 10.02.2026.

For citation: Anisimov, V.E., Anisimova, E.D., & Savastenko, R.A. (2026). Transformation Strategies of Precedent Phenomena in the Discours of Modern Web-Literature. *RUDN Journal of Language Studies, Semiotics and Semantics, 17*(2), 813–835. (In Russ.). https://doi.org/10.22363/2313-2299-2026-17-2-813-835 EDN NHXXYL

Введение

Одной из наиболее актуальных тенденций современного общества становится активное развитие сети интернет за счет возникновения, становления и функционирования в ней новых феноменов. Часть таких феноменов является уникальной для сетевого пространства, тогда как другая представляет собой трансформацию уже существующих в реальном мире явлений других дискурсов или отраслей жизни общества. В свою очередь, одним из феноменов, существующих в современном виртуальном пространстве, становится феномен сетевой литературы.

На сегодняшний день в работах современных отечественных и зарубежных ученых нет единого подхода к определению феномена сетевой литературы, а также возможности отнесения к ней определенных художественных текстов. Так, Ю.Л. Проект определяет сетевую литературу или *сетературу* как художественные произведения, создаваемые и размещаемые на виртуальных площадках в сети Интернет на основе интерактивного взаимодействия автора и других пользователей площадки размещения. Исследователь выделяет такие черты сетевой литературы, как гипертекстовость, включенность читателей в процесс создания художественного текста, низкие требования к грамматическому, стилистическому и сюжетному оформлению текста, возможность соавторства создателя текста и читателей, а также широкий коммуникативный спектр взаимодействия читателей с автором в процессе создания произведения [1].

И.В. Гурулев относит сетевую литературу к более широкому феномену современного сетевого искусства «нет-арт». По мнению исследователя, все виды сетевого искусства обладают рядом схожих характеристик, таких как гипертекстовость, динамичность, повторяемость и реинтерпретируемость, а также принципиальная процессуальность и возможность бесконечной коммуникации [2].

К.Н. Урбан и О.В. Токарь говорят о сетевой литературе в широком смысле как совокупности всех текстов, размещенных в сети Интернет и обладающих признаками художественного текста [3]. При этом в узком смысле сетевая литература понимается как совокупность художественных текстов, появившихся благодаря Интернету и не способных существовать вне его. Аналогичного мнения придерживаются и А.В. Андреев и О.И. Бычкова [4; 5].

Среди иностранных исследователей также нет единого мнения относительно феномена сетевой литературы. Так, в различных работах можно встретить наименования *network fiction*, *internet literature*, *cyber literature*, а также *online literature*. При этом к сетевой литературе различные исследователи относят тексты интернет-дневников, сетевых блогов, расширенных комментариев или публикаций в социальных сетях, содержащих элементы художественного текста, а также художественные произведения, публикуемые на различных онлайн-платформах, и так называемые *stream novels* — особый вид сетевой литературы, являющийся способом коллективного соавторства писателя и пользователей сайта на основе направленного развития сюжета. Для

описания данного явления нами предлагается использовать термин *литературная трансляция* в силу его семантической близости с английским *stream novel* [6–9].

Изучением сетевой литературы как особого вида современных художественных текстов занимались К.Н. Урбан, О.В. Токарь и С.А. Петренко. Авторы отмечают, что тексты сетевой литературы, обладая признаками гипертекста, тем не менее по форме содержания повторяют традиционные литературные произведения [3; 10]. К схожим выводам приходит О.И. Бычкова, исследуя стилистические особенности произведений современной сетевой литературы [5]. Отмечая среди специфических черт упрощенную синтактико-грамматическую структуру текста, исследователи приходят к выводу, что произведения сетевой литературы зачастую представляют собой феномен симулякра [11]. В свою очередь, анализ когнитивных особенностей текстов сетевой литературы позволяет делать выводы о высокой степени взаимодействия автора и читателя в формате коллективного авторства произведения, что становится все более актуальным с развитием новых, более коммуникативно-ориентированных цифровых платформ [10].

Таким образом, феномен современной сетевой литературы представляет собой многогранное и динамически развивающееся явление, связанное с активным развитием и становлением собственно сети интернет, в рамках которого функционирует данный феномен. Значительный интерес для современных исследователей представляют различные характеристики и особенности их проявления в рамках дискурса сетевой литературы в различных аспектах: социальном, когнитивном, языковом, стилистическом и литературном.

В данном исследовании под сетевой литературой понимаются собственно произведения, создающиеся или впервые опубликованные на различных сетевых платформах: работающих по принципу самиздата («Литнет», «Литрес», «АвторТудей», «ЛитГород», «КнигоГрад», «Литмаркет»), личные блоги и платформы авторов-литераторов (Boosty, LiveLib, WattPad), в качестве постов в социальных сетях, в том числе в авторских группах или на личных страницах авторов (ВКонтакте, Telegram, RuTube) и др. При этом можно отметить, что такие тексты могут быть как *малоформатными* (до 1 п.л. — в случае литературного текста), так и *полноформатными* (свыше 1 п.л.), но по структуре должны соответствовать классическому литературному произведению: содержать завязку, развитие и развязку сюжета, обладать как минимум одной сюжетной линией, а также содержать законченную мысль [3]. В отличие от авторских постов, публикуемых в социальных сетях, произведения сетевой литературы обладают всеми классическими признаками художественного текста: образность, связность, смысловое и стилевое единство, членимость, эмотивность, экспрессивность, оценочность, образность, эстетичность, интенциональность, интерпретируемость и реинтерпретируемость, антропоцентричность, социологичность, диалогичность [12–15].

При этом произведения сетевой литературы также обладают отличиями от классических литературных произведений, к которым относятся:

первоначальное отсутствие физического носителя, свободная публикация авторских материалов, публикация текста в авторской редакции, возможность последовательной публикации элементов текста, возможность редакции текста в процессе работы, прямое взаимодействие автора и читателя посредством инструментов комментирования, ориентация на массового потребителя в моменте, высокая авторская продуктивность, регулярная публикация материалов, серийность публикаций, преимущественное функционирование в сети Интернет, распространение текста в гипертекстовом виде с включенными метаданными.

На основании специфических признаков, дискурс сетевой литературы может быть выделен в *отдельный тип дискурса*, *обладающий одновременно признаками литературно-художественного и интернет-дискурсов*, что, в свою очередь, требует применения смешанного методологического аппарата на стыке лингвистики, литературоведения, дискурс- и контент-анализа.

Прецедентные тексты в сетевой литературе

Одним из свойств как классического литературного текста, так и текстов сетевой литературы является интертекстуальность, «текстуальная интеракция, которая происходит внутри отдельного текста» [16. С. 5], «устройство, с помощью которого один текст перезаписывает другой текст» [17. С. 48], «свойство и способность текстов различных дискурсов или иных текстов самого дискурса к взаимодействию между собой, воплощенное в рамках конкретного текста конкретного дискурса» [18. С. 116]. В рамках литературных текстов интертекстуальность может быть реализована за счет прецедентных феноменов культуры, поскольку «феномен прецедентности имеет интертекстуальную природу <…> ибо в обоих случаях речь идет о единицах, отсылающих к явлению, которое осознается как артефакт и принадлежит контексту, находящемуся за пределами данного текста» [19. С. 55]. Апелляция к прецедентным феноменам культуры призвана создать у читателя ощущение реалистичности происходящего, моделируя категорию реальности [20–22], что особенно важно для текстов реалистического, авто- или биографического форматов. По справедливому замечанию исследователей, с прагматической точки зрения прецедентные имена отличаются яркой образностью и выразительностью [23].

Особым случаем моделирования категории реальности становится формирование так называемого «вторичного мира», «фантазийного мира» или «фантастической реальности» как категории вымышленной реальности, функционирующей в произведениях фантастической и фэнтезийной литературы [24–26]. Данные термины описывают вымышленную систему вымышленных реалий, вторичную по отношению к объективной, реально существующей. Так, Е.Ю. Ваулина анализирует языковую картину нереального мира применительно к литературным произведениям различных жанров, в первую очередь — фантастического [25]. О.В. Лекарева, рассматривая различные

термины, применяемые к описанию вторичной реальности вымышленного или литературного мира, говорит о квазизнаниях, которые формируют языковую картину нереального мира [27]. Важность обращения к категории реальности и, соответственно, вторичной реальности, связана с тем, что в любом литературно-художественном произведении, в том числе относящемся к дискурсу сетевой литературы, формируется собственный, авторский мир, который, в зависимости от жанровой принадлежности, может иметь больше или меньше пересечений с объективной реальностью. Таким образом, язык и собственно лингвистические способы репрезентации различных социокультурных реалий будут играть основную роль в способах репрезентации объектов нереального авторского мира в художественных произведениях.

Представляется возможным отметить, что на сегодняшний день возрастает популярность и социальная востребованность новой сетевой литературы, которая, в своем большинстве, относится к различным фантастическим и фэнтезийным жанрам, а также к жанрам смешанного типа с элементами фантастики. Современное технологическое обеспечение, в частности, доступность платформ для публикации авторских произведений, обеспечивает увеличение количества произведений, публикуемых авторами, что также обеспечивает беспрецедентный рост количества литературных произведений, публикуемых в новом интернет-формате.

По данным информационно-аналитического журнала «Университетская книга», на сайтах сетевой литературы, таких как «Литнет», по совокупности произведений 2–3-е место занимают фэнтезийные и фантастические произведения, тогда как на 1-е и 4-е местах находится жанр любовного романа и ужасов, при этом следует отметить, что более 80 % таких произведений имеют вторичную пометку о принадлежности к фэнтези или фантастике, что позволяет отнести их к произведениям смешанных жанров[1]. В свою очередь, согласно статистическим данным таких платформ как AuthorToday и «Литрес» фантастика и фэнтези также входят в первые 5 жанров как по количеству публикаций произведений авторами, так и по популярности (4-е и 1-е место по количеству публикаций, 2-е и 1-е по популярности за 2024 г. на платформе AuthorToday, 3-е и 2-е место по количеству публикаций, 4-е и 1-е по популярности за 2024 г. на платформе «Литрес»)[2]. При этом стоит отметить, что произведения, обладающие множественными пометками метаданных, то есть смешанным жанром, отображаются в статистике только по первому указанному в метаданных жанру, что может косвенным образом влиять на учет произведений, принадлежащих определенному жанру.

[1] «Литнет» и итоги 2024 года: рейтинг самых пишущих городов и самых популярных среди писателей жанров // Информационно-аналитический журнал «Университетская книга». URL: https://www.unkniga.ru/company-news/litnet-i-itogi-2024-goda-reiting-samykh-pishushchikh-gorodov-i-samykh-populiarnykh-sredi-pisatelei-zhanrov.html (дата обращения: 10.11.2025).

[2] Жанры книг // AuthorToday. URL: https://author.today/work/genres (дата обращения: 15.11.2025); Жанры книг // Литрес. URL: https://www.litres.ru/pages/new_genres/ (дата обращения: 11.11.2025).

В свою очередь, большое количество произведений, публикующихся авторами на рассматриваемых платформах, может быть связано с такими наблюдаемыми нами характеристиками современной сетевой литературы, как высокая скорость публикации произведений; публикация произведений, находящихся в процессе написания, а также большой процент серийных произведений. Последняя из наблюдаемых особенностей связана с развитием многопланового сюжета при создании произведения, что позволяет автору формировать вторичную реальность произведения или авторский мир. Вместе с тем публикация произведения сетевой литературы в процессе написания, феномен, который можно обозначить как *процессионная публикация произведения*, значительно осложняет автору проработку собственно внутреннего прецедентного содержания и позволяет ему с большей вероятностью обращаться к созданию трансформированных прецедентных феноменов с опорой на реально существующие (исходные) прецеденты.

Исследование элементов, формирующих языковую картину нереального мира в фантастической и фэнтезийной литературе, является новым направлением исследований. Ряд авторов изучают как особенности перевода таких элементов на примере классических фантастических произведений, таких как работы Дж. Р.Р. Толкиена, так и собственно специфика создания и функционирования определенных квазизнаний, таких как вымышленные онимы различной конфигурации, окказионализмы и авторские неологизмы [28–31].

Прецедентные феномены в языковой картине нереального мира: способы формирования и функционирования

В исследовании предполагается обратиться к особенностям формирования и функционирования квазиэлементов, таких как прецедентные феномены вторичной реальности в современной сетевой литературе фантастического, фэнтезийного и смешанного жанров, имеющих в метаданных пометку о вторичной принадлежности к ним же. Под прецедентными феноменами мы понимаем *прецедентные онимы* (антропонимы, топонимы, хрематонимы, урбанонимы, ойконимы и др.), а также *прецедентные высказывания* (цитаты из литературных, научных и религиозных произведений, крылатых выражений, текстов песен и элементов полимодальных дискурсов — фильмов, сериалов, мультфильмов и пр.).

Формирование вторичных прецедентных феноменов подобного рода может происходить за счет заимствования реальных прецедентов из объективной реальности для формирования у читателя знакомых ассоциативных образов и воссоздания элементов реального быта, а также составления композиции принадлежности к определенному временному периоду (что актуально для фантастических жанров, таких как «альтернативная история», «путешествия во времени», «попаданцы» и т.д., а также смешанных жанров фантастического детектива, фантастического любовного романа и пр.), что в целом направлено на формирование образа реальности произведения сетевой литературы. Иным способом формирования прецедентных феноменов в

текстах современной сетевой литературы становятся авторские неологизмы. К ним относятся прецедентные феномены, не имеющие аналогов в реальном мире и принадлежащие вторичному нереальному миру конкретного произведения или серии произведений, что формирует пласт нереальных социокультурных феноменов, формирующих авторский вымышленный мир [32–37].

При исследовании особенностей формирования авторских неологизмов в современной сетевой литературе, необходимо подробно остановиться на таком способе их образования, как языковая трансформация прецедентных феноменов реального мира в рамках их функционирования во вторичной реальности сетевого литературного произведения. Известно, что трансформация различных прецедентных феноменов может происходить и в рамках объективной реальности: примерами таких изменений могут быть пословичные трансформы [38; 39], локальные мемы [40; 41], а также трансформации мемов как нового типа прецедентных феноменов, функционирующих в сети Интернет [42; 43]. В частности, П.И. Трофимчук выделяет следующие типы трансформации прецедентных феноменов на примере анализа интернет-мемов: замена компонента прецедентной фразы, образование симулякра и ассоциативная замена фрагмента [42]. Следуя аналогичной схеме и анализируя тексты современной сетевой литературы фэнтезийного и фантастического жанров, нами было проведено исследование и составлена классификация трансформационных стратегий прецедентных феноменов как в рамках онимов различной природы, так и прецедентных высказываний, которая будет подробно рассмотрена в рамках представленной работы.

Авторские неологизмы, репрезентирующие прецедентные феномены вторичной реальности литературного произведения, могут быть соотнесенными с реальными прецедентами различной природы за счет трансформации, а также несоотнесенными, т.е. созданными автором для функционирования исключительно в рамках конкретного произведения / серии произведений. Использование прецедентных феноменов ― авторских неологизмов, трансформирующих реально существующие прецеденты, позволяет создавать для потенциального читателя аллюзии к знакомой ему объективной реальности, тем самым моделируя категорию реальности и делая сетевой литературный текст более привлекательным за счет упоминания знакомых и культурно-значимых элементов.

В рамках представленного исследования мы будем оперировать термином *исходный прецедентный феномен (далее — ИПФ)* для реально существующего прецедентного феномена (прецедентного имени, прецедентного высказывания и др.) и *трансформированный прецедентный феномен (далее — ТПФ)* как преобразованный автором существующий реальный прецедентный феномен, в том числе посредством создания неологизма, для создания вторичной реальности сетевого художественного произведения. Обращение к термину *прецедентный феномен* как зонтичному производится нами вслед за В.В. Красных ввиду его гиперонимичности по отношению к другим существующим терминам [44].

Материалы и методы

Эмпирической базой исследования послужили 150 произведений сетевой литературы фантастического, фэнтезийного или смешанного с элементами фантастического или фэнтезийного элементов жанров, являющиеся элементами 21 цикла произведений, опубликованными на специальных платформах для размещения литературного авторского контента “AuthorToday”, «Литрес» и «Литнет». Выбор платформ обусловлен их наибольшей популярностью среди русскоязычных пользователей. Выбор произведений и циклов произведений, в свою очередь, связан с их жанровой направленностью и популярностью среди пользователей платформ. Дополнительным критерием изучения конкретных циклов произведений стало включение не менее 5 книг в каждую из рассматриваемых серий для изучения вторичной (вымышленной) реальности, что полноценно раскрывается на примерах серийных произведений сетевой литературы.

Для анализа полноформатных текстов произведений использовался автоматизированный авторский корпус на базе программного обеспечения AntCoreGen, составленный на основе загруженных полноформатных произведений, разделенный по циклам. Основными методами исследования являлись корпусный анализ текста, контент-анализ, лингвистический анализ текста, метод сплошной выборки, методы математической статистики и метод интертекстуального анализа для выявления способов трансформации прецедентных феноменов.

Трансформации прецедентных феноменов в текстах сетевой литературы как метод создания вторичной (авторской) реальности

В ходе проведения исследования современной сетевой литературы нами выявлено, что трансформации прецедентных феноменов — онимов различного генеза осуществляются при помощи 5 основных стратегий: а) трансформация по созвучию; б) трансформация по аналогии; в) трансформация по принципу синонимично-антонимичной замены; г) трансформация по анаграмматическому принципу; д) стратегия доместикации/форенизации.

Трансформация прецедентных феноменов — имен собственных по созвучию происходит за счет фонетической общности исходного и трансформированного высказывания. Применение данной стратегии может быть осложнено в рамках восприятия потенциальным читателем за счет отсутствия аудиальной модальности в текстах современной сетевой литературы, так как произведение представлено в виде текста. Таким образом, прецедентные феномены, трансформированные по фонетическому принципу, могут не всегда распознаваться читателем, в особенности при трансформировании автором менее известных исходных прецедентных феноменов. В таком случае следует подчеркнуть, что при применении фонетического типа трансформации авторам следует использовать наиболее известные прецеденты для верной интерпретации аллюзий потенциальным читателем.

Примерами таких трансформаций могут стать следующие имена собственные:

«*...ведущим программы был известный умник* ***Алессандр Груздь*** *...*» (MeXXanik Гоблин, Каин. *Морозов. Книга 3*. 2022).

В данном примере можно наблюдать фонетическую трансформацию прецедентного антропонима: по созвучию изменено прецедентное имя Александра Друзя, известного участника интеллектуальной игры «Что? Где? Когда?», на фонетически созвучное. При этом можно отметить как выпадение оглушенной согласной с последующим удвоением последующего согласного звука во избежание его озвончения в имени, так и замену одного звонкого согласного на другой в слабой позиции в начале фамилии. Следует отметить, что максимальную фонетическую близость исходного и трансформированного вариантов обеспечивает нахождение их конечных фонем в слабой (неартикулированной) позиции.

Помимо собственно схожей фонетической формы верно интерпретировать ТПФ читателю помогает описательная стратегия, применяемая автором: словосочетание *известный умник*, характеризующее главного героя, представляет собой аллюзию на интеллектуальную телепередачу «Что? Где? Когда?».

Иным вариантом трансформации прецедентного феномена по созвучию можно назвать следующий вариант:

«*...оглянувшись по сторонам, он заметил ресторанчик популярной едальни* ***Вкусно точно***» (MeXXanik Гоблин, Каин. *Морозов. Книга 1*. 2022).

В данном случае, прецедентное имя — эргоним-название сети ресторанов быстрого питания «Вкусно и Точка» — также претерпевает трансформацию по фонетическому типу. Частичная замена фонетической формы эргонима позволяет, с одной стороны, оставить прецедентный феномен узнаваемым для потенциального читателя, а с другой — изменить его в достаточной степени во избежание конфликта интересов и учета авторских прав. В данном примере можно также отметить описательную стратегию, применяемую автором для верной интерпретации прецедентного феномена (*ресторанчик популярной едальни* — указание на сеть общественного питания);

«*...ты — моя* ***Бузиновая палочка****, а я — местный* ***ВолонМуморд***» (Джон Голд. *Калибр личности. Книга 1*. 2024).

В рассматриваемом примере изменению подвергаются несколько онимов: Волан-де-Морт — имя главного антагониста книг о Гарри Поттере, а сама фонетическая трансформация достигается за счет восстановления полнозвучия в составной части имени (Волан — Волон), а также в силу изменения искажения фонетической формы составной части сложного имени в неартикулируемой позиции и графической замены парной согласной с глухого

инварианта на озвонченный с сохранением ее произношения (де-Морт — Муморд); ИПФ-хрематоним *Бузинная палочка*, изменен согласно стратегии синонимично-антонимичной замене на ТПФ *Бузиновая палочка*.

Иной стратегией становится *трансформация путем синонимично-антонимичной замены*. Использование данной стратегии позволяет использовать при трансформации синоним или антоним (также возможно применение обоих вариантов) для изменения формы прецедентного феномена. В таком случае также возможна как частичная, так и полная замена прецедентного феномена, чему можно привести следующие примеры:

«*…по телевизору шло отвратительное и скандальное шоу* ***Изба-2***» (MeXXanik Гоблин, Каин. *Дело Васильевской лавки*. 2023);
«*…для вылета назначили международный аэропорт* ***Домобабово***» (MeXXanik Гоблин, Каин. *Морозов. Книга 4*. 2024);
«*…наши призраки не пришли смотреть* ***«Несчастливы порознь»***…» (MeXXanik Гоблин, Каин. *Грехи отцов*. 2025).

В приведенных примерах использована синонимическая замена (замена ИПФ *«Дом-2»* на синоним с более узуальным значением ТПФ *«Изба-2»* относительно названия телешоу), антонимическая частичная замена (ИПФ *аэропорт Домодедово* на ТПФ *аэропорт Домобабово* за счет замены части урбанонима на антонимичный корень), а также антонимичная лексическая замена с синонимичным совокупным значением (ИПФ *«Счастливы вместе»* — ТПФ *«Несчастливы порознь»* — антонимичная замена каждого из членов прецедентного названия телесериала, что в совокупности позволяет сохранить синонимичное значение). Таким образом, подчеркнем, что применение трансформационной стратегии синонимично-антонимичной замены позволяет сохранить аллюзию на исходный прецедентный феномен за счет использования синонимов или антонимов различного узуса, что, в свою очередь, способствует созданию и поддержанию аутентичной языковой картины нереального мира за счет употребления лексических единиц со специфической коннотацией, а также создать определенный необходимый автору произведения эмотивный эффект, например, комический или саркастический.

Следующей стратегией, применяемой для трансформации прецедентных феноменов — имен собственных — является *замена имени или части имени по аналогии или ассоциативно*. В таком случае изменение прецедентного феномена происходит по принципу замены онима или его части на ассоциируемый с ним прецедент. В подобном случае можно говорить о двойном прецеденте: как о прецедентности собственно имени, так и о связанной с ним ассоциации или аналогии.

Примером может стать использование следующего прецедентного имени:

«…в гостиной, закинув ногу на ногу, сидела скандальная журналистка ***Ксения Лошадчак****»* (MeXXanik Гоблин. *Морозов. Книга 2*. 2023)

В данном случае трансформации по ассоциативному принципу подвергается часть антропонима — фамилия. Обращение к известному медийному и интернет-прозвищу[3] журналистки и политической активистки позволяет избежать прямой номинации (как собственно фамилии, так и прозвища) и при этом создать узнаваемый современным читателем образ, что также поддерживается прецедентным описанием — *скандальная журналистка*. При этом в трансформированном антропониме сохраняется суффиксальная часть фамилии для сохранения созвучия, а также заменяется корневая его часть на ассоциативный прецедент. Применение такой стратегии позволяет одновременно избежать прямого упоминания известной личности и создать узнаваемый потенциальным читателем образ. Данная стратегия имеет дальнейшее раскрытие в произведении за счет упоминания онима-названия блога К.А. Собчак, также трансформированного согласно данной стратегии: ИПФ *«Осторожно, Собчак!»* — ТПФ *«Осторожно, Лошадчак!»*.

Аналогичным примером может стать изменение фамилий иных медийных и известных в различных контекстах личностей, в том числе политических деятелей: примером подобной трансформации может служить антропоним *Усков* (Новиков Н.Н. *Господин наследник, вы уже не приемный*. Т. 3. 2023) как преобразование имени российского государственного деятеля и пресс-секретаря Президента Российской Федерации Дмитрия Сергеевича Пескова, построенный на основе отличительной черты его внешности — усов, а также подкрепленный знаменитым интернет-мемом «усы Пескова»[4]. В качестве дополнительного примера приведем трансформацию антропонима ИПФ: Филипп Бедросович (Киркоров) на ТПФ: Филипп Примадонович (Новиков Н.Н. *Наномашины, звереныш!* Т. 5. 2024) на основе ассоциации с экс-супругой певца Аллой Пугачевой, известной также как Примадонна за счет исполняемой ею одноименной песни.

При этом стоит отметить, что используемые варианты ассоциаций могут носить как позитивные, так и негативные коннотации, за счет чего автором достигается комический или саркастический эффект. В отличие от изменения прецедентного феномена по созвучию данная трансформационная стратегия также несет в себе дополнительную эмотивную функцию в создании образа персонажа.

Четвертой выявленной нами трансформационной стратегией становится *анаграмматическая трансформация*. В случае применения подобного рода трансформации, в исходном варианте прецедентного феномена меняется порядок букв в словах, частично или полностью, за счет чего достигается эффект преобразования имени собственного. Как и в случае с фонетической

[3] «Людям приятно написать „Собчак — лошадь“»: почему Ксения не страдает от неприязни в свой адрес // Gazeta.ru. URL: https://www.gazeta.ru/culture/news/2025/11/25/27264343.shtml?utm_auth=false (дата обращения: 20.12.2025).

[4] Усы Пескова // Neolurk. URL: https://neolurk.org/wiki/%D0%A3%D1%81%D1%8B_%D0%9F%D0%B5%D1%81%D0%BA%D0%BE%D0%B2%D0%B0 (дата обращения: 21.11.2025).

трансформацией, данный способ изменения прецедентных феноменов может быть сложным для восприятия потенциальным читателем и не всегда приводить к верному пониманию аллюзии на оригинальный социокультурный прецедентный феномен. Приведем ряд примеров, иллюстрирующих применение авторами сетевых художественных произведений означенной стратегии:

> «*…здесь снимают кино с самой* ***Боней Витой****!*» (MeXXanik Гоблин Каин. *Карамазов. Книга 2*. 2022);
> «*…на сцене, весь в перьях, появился король эстрады Российской империи —* ***Кирилл Фаркопов****…*» (Капба Е.А. *Великий и ужасный 2.* 2023);
> «*…как всегда, по вопросам внешней политики резко высказался скандально известный* ***Данжей Юда****, но его уже никто не слушал…*» (Григорий В. *Возрождение Фениса*. Т. 02. 2022).

В случае применения анаграмматических трансформаций изменению могут подвергаться:

- собственно лексемы имени и фамилии исходного прецедентного феномена-антропонима: ТПФ *Бона Вита* — ИПФ *Виктория Боня* (исходным прецедентным феноменом является имя российской телеведущей Виктории Бони);
- звуко-буквенная последовательность в одной из частей антропонима: ТПФ *Данжей Юда* — ИПФ *Анджей Юда* — трансформация первых букв имени и фамилии экс-президента Республики Польша 2015—2025 гг.;
- звуко-буквенная последовательность в обеих частях (имя и фамилия) антропонима: ТПФ *Кирилл Фаркопов* — ИПФ *Филипп Киркоров*, где осуществлена взаимозамена первых букв имени и фамилии, изменение звуко-буквенного порядка в паре имя-фамилия с сохранением схожего образа.

При этом стоит отметить, что анаграмматическая стратегия трансформации прецедентных феноменов также представляется достаточно сложной, так как в случае с реальными онимами для образования авторского неологизма следует подбирать не только благозвучный, но и адекватный с точки зрения представления читателю вариант (реально существующий вариант имени и т.д.), то есть тот, который не будет смотреться инородно относительно иных языковых единиц и других употребляемых в произведении онимов. Данный вопрос в меньшей степени актуален для фэнтезийных и фантастических произведений за счет возможности введения авторских онимов, однако в случае моделирования вторичной реальности с опорой на реальные события (в таких жанрах, как альтернативная история, путешествия во времени и т.д.) следует учитывать данный аспект.

Наконец, последней выявленной нами стратегией трансформации прецедентных феноменов в текстах сетевой литературы становится стратегия доместикации/форенизации, которая подразумевает дословный перевод и/или трансформацию ИПФ согласно стилистическим, грамматическим и/или фонетическим нормам иного языка. Так, формат доместикации подразумевает изменение иноязычных прецедентных феноменов-онимов согласно нормам

языка культуры, для носителей которой пишется произведение, тогда как формат форенизации позволяет изменять прецедентные феномены-онимы родной лингвокультуры согласно правилам иного лингвокультурного сообщества.

> «*На северном материке стальной рукой ковал коммунизм* ***Джозеф Стил****, абориген, кровавым путем пришедший к власти и получивший признание цивилизованного мира…*» (Кронос А. *Возвышение Меркурия*. Т. 1. 2023–2025).

Данный пример демонстрирует яркий образец трансформации по принципу форенизации, когда ИПФ-антропоним *Иосиф Сталин* претерпевает трансформацию в формате англицизма: имя Иосиф заменяется на англоязычный аналог Джозеф (трансформация по принципу англоязычного инварианта имени), фамилия Сталин претерпевает трансформацию по принципу дословного перевода (Сталин — Стил (от англ. steel ― сталь)).

> *«…пора было наведаться к моему старому приятелю,* ***Антохе Простоватому****… давеча он горел прожектом изобресть железный костюм, чтобы «человек мог летать так, как птица»…»* (Горбов А. «К». *Человек государев*. Т. 1. 2025).

В приведенном примере можно увидеть применение стратегии доместикации, когда изменениям подвергается антропоним известного персонажа американских комиксов и фильмов о супергероях Тони Старка (ИПФ), который в ходе стилистико-переводческих трансформаций (Тони — Антоха ― изменение англоязычного варианта имени на русскоязычный вариант и стилистическая трансформация по принципу просторечной формы антропонима (Антон — Антоха); Старк — Простоватый (от англ. stark ― empty, simple, or obvious, especially without decoration or anything that is not necessary (пустой, простой, очевидный, без приукрашений и лишних деталей)[5] — по принципу дословного перевода и стилистической аутентичности) он приобретает черты русскоязычного антропонима с признаками просторечия.

Следует отметить, что ТПФ вторичной реальности литературного произведения не всегда могут быть верно интерпретированы читателем. Для решения данного вопроса авторы используют прием комментирования или описания упоминаемого прецедента с опорой на его содержание, например, добавление характеристики к трансформированному прецедентному феномену *король эстрады Кирилл Фаркопов*, где описательное пояснение «король эстрады» позволяет соотнести ТПФ с ИПФ за счет апелляции к его содержанию. Таким образом, описательные единицы играют большую роль в распознавании и верной интерпретации трансформированных прецедентных единиц, так как функционально позволяют сохранить и актуализировать их прецедентное содержание.

[5] Stark: dictionary entry // Cambridge Dictionary Online. URL: https://dictionary.cambridge.org/dictionary/english/stark (accessed: 28.12.2025).

Таким образом, на основе проанализированных нами произведений можно сделать следующие статистические наблюдения: процент авторских неологизмов относительно онимов иной природы составляет в среднем 15 %, что позволяет, с одной стороны, не перенасыщать текст отсылками к объективной реальности, и, с другой стороны создать для потенциального читателя ряд аллюзий на реальную жизнь и/или актуальные вопросы современности. Применение различных стратегий транформации прецедентных феноменов может быть продиктовано идиостилем автора, жанровой составляющей произведения, а также жанрово-стилистическими особенностями конкретного произведения. Согласно проведенному анализу трансформированные прецедентные феномены чаще всего встречаются в таких поджанровых категориях сетевой литературы, как «Бояръ-Аниме», «Бытовое фэнтези», «Городское фэнтези», «Альтернативная история», «Социальная фантастика», «Попаданцы» и «Славянское фэнтези». Выявленная тенденция может быть связана со спецификой обозначенных жанров, когда автору для формирования образа вторичной реальности приходится обращаться к историческим или современным событиям путем упоминания определенных прецедентов. Так, в период 2020–2025 гг. в произведениях сетевой литературы актуальным представляется обращение к пандемии COVID-19 посредством ТПФ ― авторских неологизмов, таких как «коронавирус — манавирус», «атипичная пневмония — стеклянная пневмония», «Ковид — Кьювид» и др. (циклы: «Лекарь империи». А. Лиманский, С. Карелин 2024–2025; «Спасите меня, Кацураги-сан!». В. Молотов, А. Аржанов 2023–2025; «Калибр личности», Дж. Голд 2024–2025), на примере которых можно наблюдать применение различными авторами разных стратегий трансформации одних и тех же исходных прецедентных феноменов.

Согласно проведенному статистическому анализу составленного корпуса наиболее частотной стратегией трансформации прецедентных феноменов становится синонимично-антонимичная трансформация (33 %), а наименее частотной — анаграмматическая трансформация (10 %). Такая тенденция может быть связана со сложностью подбора лексических единиц с аналогичным звукобуквенным составом, что позволило бы сохранить и аллюзию к ИПФ, и адекватную форму ТПФ.

Таким образом, стратегии трансформации прецедентных феноменов в произведениях современной сетевой литературы (на русском языке) может быть представлены как:

— *трансформация по созвучию* (19 %), ИПФ (Иисус), ТПФ (Иссу): Голд Дж. Цикл «Калибр личности». 2024–2025);

— *трансформация по аналогии* (21 %), ИПФ (Москва-Сити), ТПФ (Снежинск-Град): МеXXanik Гоблин, Каин. Цикл «Морозов». 2022–2024);

— *синонимично-антонимичная замена* (33 %), ИПФ (Тони Старк «*Железный человек*»), ТПФ (Тони Карк «*Золотой человек*»): Голд Дж. Цикл «Калибр личности». Т. 7. 2024–2025);

— *анаграмматическая трансформация* (10 %), ИПФ (София Августа Фредерика *Ангальт-Цербстская* (Екатерина II)), ТПФ (София Августа Фредерика *Цегбальт-Анверрская*): Горбов А.«К». Цикл «Дядя самых честных правил». 2022–2025);

— *доместикация / форенизация* (17 %), ИПФ (Илья Егорович Маскин), ТПФ (Илон Макс, отец Эррол Маск): Еслер А. Цикл «Соприкосновение миров. Книга 4». 2023–2025);

Таким образом, можно говорить о том, что различные стратегии трансформации прецедентных феноменов — онимов могут взаимодополнять друг друга как в рамках одного произведения, так и в рамках изменения одного прецедентного феномена (ТПФ *Маргарита Федоровна Тэтчер по прозвищу «Стальная сударыня»* — ИПФ *Маргарет Тэтчер «Железная леди»* (отчество трансформировано от имени отца Альфред — Фёдор) (Капба Е.А. *На золотом крыльце*. Т. 1. 2025). Применение различных стратегий трансформации прецедентных феноменов позволяет им выполнить ряд функций во вторичной реальности произведения сетевой литературы: придать эмоциональную окраску, репрезентировать национальную аутентичность, привлечь внимание читателя, моделировать знакомый читателю образ и т.д.).

Особенности трансформаций прецедентных высказываний в произведениях сетевой литературы

Помимо трансформации собственно прецедентных феноменов онимичной природы, следует отметить, что изменению могут подвергаться также и целостные высказывания ― прецедентные фразы, заимствованные из других дискурсов, чаще всего дискурсов креативной природы (кинодискурс, песенный дискурс, интернет-дискурс и т.д.). В случае с изменением прецедентной фразы можно отметить применение метода частичной замены лексем с сохранением общей синтаксической структуры высказывания.

«*…на каждый вопрос есть четкий ответ, — вдруг продекламировал Заболоцкий, — у нас есть картечница, у них ее нет.*» (Горбов А.«К». «Дядя самых честных правил». Книга 11. 2025) — трансформация прецедентного высказывания Х. Беллока: «*На каждый вопрос есть четкий ответ: у нас есть «максим», у них — его нет*» (Беллок Х. *Современный путешественник*. 1989);

«*…а теперь вы, барон, приходите и говорите: ваша светлость, мне нужна ваша помощь. Но вы просите без уважения, не предлагая дружбу, даже не думая обратиться ко мне «мессер»…*» (Горбов А.«К». *Дядя самых честных правил. Книга 8.* 2024) — трансформация прецедентного высказывания из кинофильма «Крестный отец» Ф.Ф. Копполы, 1972: «*Ты пришел и говоришь: Дон Корлеоне, мне нужна справедливость. Но ты просишь без уважения, ты не предлагаешь дружбу, ты даже не назвал меня крестным отцом*».

На данных примерах можно проследить тенденцию сохранения ключевых элементов прецедентных высказываний грамматико-синтаксической организации предложений, смысловых лексем, с заменой собственно прецедентных имен в высказываниях (*мессер — крестный отец*, *картечница* — «*максим*» (станковый пулемет Максима), *ваша светлость — Дон Корлеоне* и др.).

Таким образом, для актуализация прецедентных высказываний под нужды автора для создания вторичной языковой картины нереального мира используется тот же принцип, что и при трансформации прецедентных феноменов-онимов: сохранение синтактико-грамматической структуры с заменой прецедентного имени-онима. При этом упоминаемое прецедентное имя подвергается замене, а не авторской трансформации, что может быть связано со сложностью проработки комплексного включения сложносоставных элементов в канву произведения сетевой литературы за счет краткого времени написания и публикации произведения, а также быстрого распространения посредством специализированных интернет-платформ.

Заключение

Подводя итог проведенному анализу и описанию стратегий трансформации прецедентных феноменов в произведениях современной российской сетевой литературы, можно сделать вывод о функции формирования категории реальности как одной из основных прагматических функций прецедентных феноменов в текстах сетевой литературы в целом, и в текстах фантастической сетевой литературы в частности. Функция формирования категории реальности реализуется за счет создания лингвистических маркеров языковой картины нереального мира, то есть создания прецедентных феноменов, характерных для вторичной (авторской) реальности произведения. Наиболее отчетливо феномен вторичной реальности прослеживается в серийных произведениях сетевой литературы, тенденция к которым является актуальной для современного состояния дискурса сетевой литературы, вследствие многопланового и последовательного развития сюжета и описания авторского мира.

Прецедентные феномены вторичной (авторской) реальности произведений сетевой литературы могут быть заимствованы из реального мира, подвергаться трансформации, имея в качестве своей основы реально существующие прецедентные феномены (при этом формируются авторские неологизмы) или создаваться без опоры на реально существующие прецедентные феномены (в таком случае прецедентный феномен является авторским неологизмом). Трансформация реально существующих прецедентных феноменов в текстах сетевой литературы позволяет авторам создавать знакомые читателю аллюзии и опираться на знакомые и популярные образы из объективной реальности, формируя таким образом категорию реальности в произведении, которая является вторичной, и приближая ее к действительности (делая более «реальной» для читателя), а также оперировать существующими в коллективном сознании образами, а не создавать их с нуля.

В ходе анализа выявлено, что трансформации подвергаются отдельные единицы-прецедентные феномены, преимущественно онимичной природы различных типов, тогда как целостные единицы-прецедентные высказывания подвергаются только частичной трансформации, сохраняя свой облик на грамматическом (синтаксическом) и семантическом уровнях. В прецедентных высказываниях трансформации или замене также подвергаются прецедентные феномены онимичной природы.

Нами выделено пять основных стратегий трансформации прецедентных феноменов в текстах сетевой литературы: фонетическая (по созвучию), анаграматическая, синонимично-антонимичная и ассоциативная (по аналогии), стратегия доместикации/форенизации. Применение каждой из стратегий может быть обосновано требованиями жанра, идиостилем конкретного автора, а также жанрово-стилистическими особенностями конкретного произведения. Наиболее частотной стратегией трансформации прецедентных феноменов является синонимично-антонимичная.

Для узнаваемости трансформированных прецедентных феноменов в тексте сетевой литературы часто встречаются описательные или уточняющие элементы, позволяющие читателю соотнести трансформированную словоформу с исходным прецедентом за счет сопровождающего его описания.

Дальнейшим направлением исследования может являться уточнение выявленных стратегий трансформации прецедентных феноменов применительно к конкретным жанрам и произведениям сетевой литературы, а также обращение к особенностям трансформации прецедентных феноменов в различных дискурсах, в первую очередь дискурсах креативной природы.

Список литературы

1. *Проект Ю.Л.* Сетевая литература // Universum: Вестник Герценовского университета. 2012. № 2. С. 208–209. EDN: QIRJZT
2. *Гурулев И.В.* Типология и художественные стратегии сетевого искусства: 1994–2004 : дис. … канд. иск. М., 2005. EDN: NNTSTH
3. *Урбан К.Н., Токарь О.В.* Особенности сетевой литературы, отличия и сходства с традиционной литературой // Труды БГТУ. Серия 4: Принт- и медиатехнологии. 2019. № 2 (225). С. 95—100. EDN: YZMGGQ
4. *Андреев А.В.* CETERAтура как ее NET: От эстетики Хэйана до клеточного автомата — и обратно // Сетевая словесность. URL: https://www.netslova.ru/andreev/setnet/ (дата обращения: 12.12.2025).
5. Бычкова О.И. Сетевая литература: вопросы формы и стиля // Наследие веков. 2015. № 2. С. 55—59. EDN: VPJHYD
6. *Ciccoricco D.* Reading Network Fiction. Tuscaloosa: The University of Alabama Press, 2007.
7. *Fathu R.* Cyber Literature: A Reader-Writer Interactivity // International Journal of Social Sciences and Educational Studies. 2017. № 3(4). P. 156–164. https://doi.org/10.23918/ijsses.v3i4p156
8. *Jamirul I.* A Reciprocal Intercommunication between Reader and Writer: A Critical Study of Cyber Literature // European journal of volunteering and community-based projects. 2022. Vol. 1. № 1. P. 26—37.

9. *Skains L.* The Shifting Author-Reader Dynamic: Online Novel Communities as a Bridge from Print to Digital Literature Convergence // The International Journal of Research into New Media Technologies. 2016. Vol. 16. № 1. P. 95–111. https://doi.org/10.1177/1354856509347713
10. *Петренко С.А.* Коллективное и авторское в жанровых матрицах сетевой словесности // Всероссийский конгресс фольклористов : сб. науч. статей. 2020. № 3 (3). С. 267–272.
11. Корнев С. «Сетевая литература» и завершение постмодерна: Интернет как место обитания литературы // Новое литературное обозрение. 1998. № 32. С. 29–47.
12. *Бахтин М.М*. Литературно-критические статьи. М. : Художественная литература, 1986. EDN: VQMUGT
13. *Новиков А.И.* Текст и его смысловые доминанты. М. : Институт языкознания РАН, 2007.
14. *Лотман Ю.М.* Структура художественного текста. Анализ поэтического текста: монография. СПб. : Азбука-Аттикус, 2016.
15. *Арнольд И.В.* Семантика. Стилистика. Интертекстуальность / под ред. П.Е. Бухаркина. СПб. : Издательство Санкт-Петербургского университета, 1999.
16. *Кристева Ю.* Семиотика. М. : Издательство МГУ, 1970.
17. *Пьеге-Гро Н.* Введение в теорию интертекстуальности / пер. с фр., общ. ред. и вступ. ст. Г.К. Косикова. М. : ЛКИ, 2008.
18. *Анисимов В.Е.* Интертекстуальность малоформатных текстов кинодискурса: внутренние связи и внешние отношения: монография. М. : РУСАЙНС, 2025. EDN: ABOYIM
19. *Дымарский М.Я.* Прецедентность и художественность // Феномен прецедентности и преемственность культур. под общ. ред.: Л.И. Гришаевой, М. К Поповой, В.Т. Титова. Воронеж: Воронежский государственный университет, 2004. С. 51–62. EDN: SAXUQH
20. *Красных В.В., Гудков Д.Б., Багаева Д.В.* O tempora, o mores! Новые структуры русской когнитивной базы // Лингвостилистические и лингводидактические проблемы коммуникации. М., 1996. С. 107–120.
21. Язык и национальное сознание: вопросы теории и методологии: коллективная монография / под ред. И.А. Стернина. Воронеж : Истоки, 2002.
22. *Новиков С.А.* Явление прецедентности в языке и современная когнитивная теория // Альманах современной науки и образования. 2013. № 10 (77). С. 124–127. EDN: RDPANB
23. *Новоспасская Н.В., Перфильева Н.В., Лазарева О.В., Никашина Н.В.* Эргонимы-названия свадебных салонов на материале русского и английского языков // Филологические науки. Вопросы теории и практики. 2017. № 2–2(68). С. 158–162. EDN: XWRSWR
24. *Божко Е.М.* Квазиреалии мира фэнтези, их классификация и роль в воздействии текста перевода на получателя // Научные ведомости Санкт-Петербургского государственного политехнического университета. Гуманитарные и общественные науки. 2011. № 3 (131). С. 188–191. EDN: OYONUJ
25. *Ваулина Е.Ю.* Языковая картина нереального мира в современном толковом словаре // Филологические и социокультурные вопросы науки и образования : сб. материалов IV Международной научно-практической очно-заочной конференции. Краснодар : КубГТУ, 2019. С. 133–141.
26. *Мзареулов К.Д.* Фантастика. Общий курс. Севастополь : Шико-Севастополь, 2018.
27. *Лекарева О.В.* О фантастической реальности и фантастических реалиях: формирование терминологии исследований по языку литературы фэнтези // Вестник Адыгейского государственного университета. Серия 2: Филология и искусствоведение. 2020. № 4 (267). С. 68–73. EDN: TZTTJZ
28. *Скрыльник В.В.* Способы образования авторских неологизмов в английском языке на материале книг Дж. К. Роулинг // Филологические науки. Вопросы теории и практики. 2017. № 7–1 (73). С. 154–157. EDN: YQEELF
29. *Кулакова Ю.В.* Механизм образования псевдореферентных номинаций в произведениях жанра фэнтези и особенности их перевода на русский язык // Вестник Адыгейского государственного университета. Серия: Филология и искусствоведение. 2008. № 1. С. 132–137. EDN: JBTIJJ

30. *Иванов А.В., Гундина О.С.* Специфика перевода квазитопонимов в жанре фэнтези (на материале произведения Дж.Р.Р. Толкина «Властелин колец» и его русских переводов) // Проблемы романо-германской филологии, педагогики и методики преподавания иностранных языков. 2019. № 15. С. 9–13. EDN: KDRXYE
31. *Дьяченкова Е.С., Богатская Е.Ю., Савела С.В.* Передача окказионализмов с английского языка на русский (на материале произведения Терри Пратчета «Мор, ученик Смерти») // Актуальные проблемы лингвистики и лингводидактики : материалы VII научно-практической Интернет-конференции с международным участием, посвященной 90-летию Донского государственного технического университета. Ростов-на-Дону : ДГТУ, 2020. С. 110–116.
32. *Елисеев А.В.* Фантастика и реальность в системе культуры и социальных связей // Мир науки. Социология, филология, культурология. 2019. № 4. Режим доступа: https://sfk-mn.ru/PDF/03KLSK419.pdf (дата обращения 23.11.2025). EDN: VOTTZQ
33. *Юречко О.Н.* Фантастическая реальность культуры как текст и дискурс: особенности семиотического анализа // Aspectus. 2015. № 3. С. 46–49. EDN: UMRKSR
34. *Saler M.* As If: Modern Enchantment and the Literary Prehistory of Virtual Reality. Oxford, UK : Oxford University Press. 2012.
35. *Fazio L.K., Marsh E.J.* Slowing presentation time increases, rather than decreases, errors learned from fictional stories // Psychonomic Bulletin and Review. 2008. № 15. P. 180–185. https://doi.org/10.3758/pbr.15.1.180 EDN: XLDZJG
36. *Ochs E., Capps L.* Living Narrative: Creating Lives in Everyday Storytelling. Cambridge, MA : Harvard University Press, 2009.
37. *Kidd D., Castano E.* Reading Literary Fiction Improves Theory of Mind // Science. 2013. Vol. 342. Iss. 6156. P. 377–380.
38. *Вальтер Х., Мокиенко В.М.* Антипословицы русского народа. СПб. : Нева, 2005.
39. *Мокиенко В.М.* Паремиологический минимум и паремиологические максимы современной русской жизни // Слова. Концепты. Мифы : сб. статетй к 60-летию А. Ф. Журавлева. М. : Индрик, 2011. С. 218–231. EDN: QJPXKO
40. *Вешнякова А.В.* Лингвокреативный аспект интернет-мемов // Актуальные проблемы гуманитарных и естественных наук. 2016. № 6–4. С. 34–40. EDN: WCYKXX
41. *Щурина Ю.В.* Интернет-мемы как феномен интернет-коммуникации // Научный диалог. 2012. № 3. С. 161–173. EDN: OWQFTD
42. *Трофимчук П.И.* Трансформация прецедентных кинофраз в комических креолизованных интернет-мемах (на примере использования в качестве источника происхождения мемов кинотекста «Иван Васильевич меняет профессию») // Филологические науки. Вопросы теории и практики. 2024. № 17 (4). С. 1207–1216. https://doi.org/10.30853/phil20240175 EDN: YAFIXU
43. *Заболоцкая М.И.* Основные тенденции трансформации интернет-мема // Этнопсихолингвистика. 2021. № 2 (5). С. 61–75. https://doi.org/10.31249/epl/2021.02.04 EDN: SEAJRP
44. *Красных В.В.* Система прецедентных феноменов в контексте современных исследований // Язык, сознание, коммуникация : сб. статей / под ред. В.В. Красных, А.И. Изотова. М. : Филология, 1997. № 2. С. 5–12.

References

1. Proekt, Ju.L. (2012). Web Literature. *Universum: Vestnik Gercenovskogo universiteta*, 2, 208–209. (In Russ.). EDN: QIRJZT
2. Gurulev, I.V. (2005). *Typology and Artistic Strategies of Web Art: 1994–2004* [PhD Thesis]. Moscow. (In Russ.). EDN: NNTSTH
3. Urban, K.N. & Tokar, O.V. (2019). Features of Online Literature, Differences and Similarities with Traditional Literature. *Proceedings of BSTU. Issue 4: Print- and Media Technologies*, *2*(225), 95–100. (In Russ.). EDN: YZMGGQ

4. Andreev, A.V. CETERAture as NET: From Heian Elegance to Cellular Automata — and Back. *Setevaja Slovesnost*. URL: https://www.netslova.ru/andreev/setnet/ (accessed: 12.12.2025). (In Russ.).
5. Bychkova, O.I. (2015). Network Literature: Issues of Form and Style. *Heritage Magazine*, 2, 55–59. (In Russ.). EDN: VPJHYD
6. Ciccoricco, D. (2007). *Reading Network Fiction.* Tuscaloosa: The University of Alabama Press. 256.
7. Fathu, R. (2017). Cyber Literature: A Reader-Writer Interactivity. *International Journal of Social Sciences and Educational* Studies, 3(4), 156–164 https://doi.org/10.23918/ijsses.v3i4p156
8. Jamirul, I. (2022). A Reciprocal Intercommunication between Reader and Writer: A Critical Study of Cyber Literature. *European Journal of Vlunteering and Community-Based Projects*, *1*(1), 26–37.
9. Skains, L. (2016). The Shifting Author-Reader Dynamic: Online Novel Communities as a Bridge from Print to Digital Literature Convergence. *The International Journal of Research into New Media Technologies*, *16*(1), 95–111 https://doi.org/10.1177/1354856509347713
10. Petrenko, S.A. (2020). The Collective and the Authorial in the Genre Matrices of Digital Literature. *Collection of Scientific Articles* "*All-Russian Congress of Folklorists*", *3*(3), 267–272. (In Russ.).
11. Kornev, S. (1998). "Digital Literature" and the End of Postmodernism: The Internet as a Habitat for Literature. *Novoe Literaturnoe Obozrenie*, 32, 29–47. (In Russ.).
12. Bahtin, M.M. (1986). *Literary and Critical Articles*. Moscow: Hudozhestvennaja Literatura. EDN: VQMUGT
13. Novikov, A.I. (2007). *Text and Its Semantic Dominants*. Moscow: Institut Jazykoznanija RAN. (In Russ.).
14. Lotman, Ju.M. (2016). *The Structure of a Literary Text. Analysis of a Poetic Text*. Saint Peterburg: Azbuka-Attikus. (In Russ.).
15. Arnold I.V. (1999). *Semantics. Stylistics. Intertextuality*. SPb.: Izdatel'stvo Sankt-Peterburgskogo Universiteta. (In Russ.).
16. Kristeva, Ju. (1970). *Semiotics*. Moscow: MGU publ. (In Russ.).
17. Piegay-Gros, N. (2008). *Introduction in Intertextuality*, G.K. Kosikov (Ed.). Moscow: LKI. (In Russ.).
18. Anisimov, V.E. (2025). *Intertextuality of Small-Format Texts in Cinema Discourse: Internal Connections and External Relations*. Moscow: RUSIENCE. (In Russ.). EDN: ABOYIM
19. Dymarskij, M.Ja. (2004). Precedentness and Artistry. In: *The Phenomenon of Precedentness and Cultural Continuity* (pp. 51–62). Voronezh: Voronezhskij gosudarstvennyj universitet. (In Russ.). EDN: SAXUQH
20. Krasnyh, V.V., Gudkov, D.B., & Bagaeva, D.V. (1996). O Tempora, o Mores! New Structures of the Russian Cognitive Base. In: *Linguistic and Linguodidactic Problems of Communication* (pp. 107–120). Moscow. (In Russ.).
21. Sternin, I.A. (Ed.). (2002). *Language and National Consciousness: Issues of Theory and Methodology*: A Collective Monograph. Voronezh: Istoki. (In Russ.).
22. Novikov, S.A. (2013). Precedence Phenomenon in Language and Modern Cognitive Theory. *Almanac of Modern Science and Education*, *10*(77), 124–127. (In Russ.). EDN: RDPANB
23. Novospasskaya, N.V., Perfilieva, N.V., Lazareva, O.V., & Nikashina, N.V. (2017). Ergonyms-Names of Wedding Salons by the Material of Russian and English. *Philology. Theory & Practice*, 2–2(68), 158–162. (In Russ.). EDN: XWRSWR
24. Bozhko, E.M. (2011). Quasi-Realities of the Fantasy World: Their Classification and Role in the Impact of a Translated Text on the Recipient. *Nauchnye vedomosti Sankt-Peterburgskogo Gosudarstvennogo Politehnicheskogo Universiteta. Gumanitarnye i Obshhestvennye Nauki*, 3(131), 188–191. (In Russ.). EDN: OYONUJ

25. Vaulina, E.Ju. (2019). The Linguistic Picture of an Unreal World in a Modern Explanatory Dictionary. In: *Philological and Socio-Cultural Issues of Science and Education: Proceedings of the IV International Scientific and Practical Conference* (pp. 133–141.). Krasnodar: KubGTU. (In Russ.).
26. Mzareulov, K.D. (2018). *Fantasy Fiction*. Sevastopol: Shiko-Sevastopol. (In Russ.).
27. Lekareva, O.V. (2020). On Fantastic Reality and Fantastic Realities: Formation of Terminology for Research on the Fantasy Fiction Language. *The Bulletin of Adyghe State University: Internet Scientific Journal*, *4*(267), 68–73. (In Russ.). EDN: TZTTJZ
28. Skryl'nik, A.V. (2017). The Ways of Formation of Author's Neologisms in the English Language by the Material of J.K. Rowling's Books. *Philology. Theory & Practice*, 7–1(73), 154–157. (In Russ.). EDN: YQEELF
29. Kulakova, Ju.V. (2008). The Mechanism of Formation of Pseudo-Referential Nominations in Fantasy Genre Works and the Specifics of Their Translation into Russian. *The Bulletin of Adyghe State University: Internet Scientific Journal*, 1, 132–137. (In Russ.).
30. Ivanov, A.V., & Gundina, O.S. (2019). Specifics of Translation of Quasitoponyms in the Genre of Fantasy (Exemplified by J.R.R. Tolkien's Novel "Lord of the Rings" and its Russian Translations). *Problemy Romano-Germanskoj Filologii, Pedagogiki i Metodiki Prepodavanija Inostrannyh Jazykov*, 15, 9–13. (In Russ.). EDN: KDRXYE
31. D'achenkova, E.S., Bogatskaya, E.Ju. & Savela, S.V. (2020). Translation of Occasionalisms from English into Russian (Based on Terry Pratchett's "Mort: The Discworld Novel"). In: *Current Issues in Linguistics and Linguodidactics: Proceedings of the VII Scientific and Practical Online Conference with International Participation, Dedicated to the 90th Anniversary of Don State Technical University* (pp. 110–116). Rostov-na-Donu: DGTU. (In Russ.).
32. Eliseev, A.V. (2019). Fiction and Reality in the System of Culture and Social Relations. *World of Science. Series: Sociology, Philology, Cultural Studies*. 4. URL: https://sfk-mn.ru/PDF/03KLSK419.pdf (accessed 23.11.2025). (In Russ.). EDN: VOTTZQ
33. Jurechko, O.N. (2015). The Fantastic Reality of Culture as Text and Discourse: Specifics of Semiotic Analysis. *Aspectus*, 3, 46–49. (In Russ.). EDN: UMRKSR
34. Saler, M. (2012). *As If: Modern Enchantment and the Literary Prehistory of Virtual Reality.* Oxford, UK: Oxford University Press.
35. Fazio, L.K., & Marsh, E.J. (2008). Slowing Presentation Time Increases, Rather Than Decreases, Errors Learned from Fictional Stories. *Psychonomic Bulletin and Review*, 15, 180–185. https://doi.org/10.3758/pbr.15.1.180 EDN: XLDZJG
36. Ochs, E., & Capps, L. (2009). *Living Narrative: Creating Lives in Everyday Storytelling.* Cambridge, MA: Harvard University Press.
37. Kidd, D., & Castano, E. (2013). Reading Literary Fiction Improves Theory of Mind. *Science*, 342(6156), 377–380.
38. Valter, H., & Mokienko, V.M. (2005). *Anti-Proverbs of the Russian People.* Saint Peterburg: Neva. (In Russ.).
39. Mokienko, V.M. (2011). The Paremiological Minimum and Paremiological Maxims of Modern Russian Life. In: *Words. Concepts. Myths: A Collection of Articles for the 60th Anniversary of A.F. Zhuravlev* (pp. 218–231). Moscow: Indrik. (In Russ.). EDN: QJPXKO
40. Veshnjakova, A.V. (2016). The Linguo-Creative Aspect of Internet Memes. *Aktual'nye Problemy Gumanitarnyh I Estestvennyh Nauk*, 6–4, 34–40. (In Russ.). EDN: WCYKXX
41. Shhurina, Ju.V. (2012). Internet Meme as a Phenomenon of Internet Communication. *Nauchnyj dialog*, 3, 161–173. EDN: OWQFTD
42. Trofimchuk, P.I. (2024). Transformation of Precedent Movie Phrases in Comic Creolized Internet Memes (Using the Movie "Ivan Vasilyevich Changes His Profession" As an Example of the Source of Meme Origin). *Philology. Theory & Practice*, *17*(4), 1207–1216. (In Russ.). https://doi.org/10.30853/phil20240175 EDN: YAFIXU

43. Zabolockaya, M.I. (2021). The Main Trends in the Transformation of Internet-Memes. *Ethnopsycholinguistics*, *2*(5), 61–75. (In Russ.). https://doi.org/10.31249/epl/2021.02.04 EDN: SEAJRP
44. Krasnyh, V.V. (1997). The System of Precedent Phenomena in the Context of Contemporary Research. In: *Language, Consciousness, Communication: A Collection of Articles* (Vol. 2, pp. 5–12). Moscow: Filologija. (In Russ.).

Сведения об авторах:
Анисимов Владислав Евгеньевич, доктор филологических наук, доцент, доцент кафедры французского языка, Московский государственный институт международных отношений (университет) МИД РФ (119454, Российская Федерация, г. Москва, проспект Вернадского, д. 76); *сфера научных интересов*: семиотика французского кинодискурса, межкультурная прагматика, дискурс-анализ, переводоведение; *e-mail*: anisimov.vladislav.95@mail.ru
ORCID: 0000-0002-6006-3965; SPIN-код: 3521-9813; AuthorID: 1016736.
Анисимова Екатерина Дмитриевна, независимый исследователь; *сфера научных интересов*: семиотика, переводоведение, рекламный дискурс, математическая статистика; *e-mail*: kate.dk01@gmail.com
ORCID: 0000-0002-9854-3921; Author ID: 973406.
Савастенко Роман Андреевич, кандидат филологических наук, доцент кафедры массовых коммуникаций филологического факультета, Российский университет дружбы народов (117198, Российская Федерация, г. Москва, ул. Миклухо-Маклая, д. 6); *сфера научных интересов*: массмедийный текст; *e-mail*: savastenko-ra@rudn.ru
ORCID: 0000-0002-3312-3790, SPIN-код: 3523-7627, AuthorI D: 621164.

Information about the authors:
Vladislav E. Anisimov, PhD in Philology, Associate Professor at the Department of French Language, MGIMO University (76, Vernadskogo av., Moscow, Russian Federation, 119454); *Research interests*: embrace semiotics of the French film discourse, intercultural pragmatics, discourse analysis, translation studies; *e-mail*: anisimov.vladislav.95@mail.ru
ORCID: 0000-0002-6006-3965; SPIN-code: 3521-9813; AuthorID: 1016736.
Ekatherina D. Anisimova, Independent Researcher; *Research interests*: semiotics, translation studies, advertising discourse, mathematical statistics; *e-mail*: kate.dk01@gmail.com
ORCID: 0000-0002-9854-3921; Scopus ID: 973406.
Roman A. Savastenko, PhD in Philology, Associate Professor of Mass Communications Department of Faculty of Philology, RUDN University (6 Miklukho-Maklaya st., Moscow, Russian Federation, 117198); *Research interests*: mass media; *e-mail*: savastenko-ra@rudn.ru
ORCID: 0000-0002-3312-3790, SPIN-code: 3523-7627, Author ID: 621164.

RUDN Journal of Language Studies, Semiotics and Semantics
2313-2299 (Print), ISSN 2411-1236 (Online)
Вестник Российского университета дружбы народов.
Серия: ТЕОРИЯ ЯЗЫКА. СЕМИОТИКА. СЕМАНТИКА

2026 Vol. 17 No. 2 836–848
http://journals.rudn.ru/semiotics-semantics

DOI: 10.22363/2313-2299-2026-17-2-836-848
EDN: NIKYCO

Research Article / Научная статья

The Birth of a Myth in the Literary Works by S. Sanbaev: the Archetype of Cosmogony

Sholpan A. Ismailova[1] ✉, Zhadyra A. Bayanbaeva[2], Samalbai D. Daribaev[2]

[1]Asfendiyarov Kazakh National Medical University, *Almaty, the Republic of Kazakhstan*
[2]Al-Farabi Kazakh National University, *Almaty, the Republic of Kazakhstan*
✉sholpan_is@mail.ru

Abstract. Mythopoetics constitutes a significant system of elements within contemporary Kazakh literary discourse. The connection between myth and literature can be traced at all levels of the artistic whole. The myth-making of Kazakh writers shapes a worldview that underpins the nation's specific consciousness. While Soviet literature viewed the mythological nature of literary discourse as an inherent characteristic of "young written literature," for Kazakh writers, it was an act of introspection — an understanding of their own soul as the soul of a representative of an authentic ethnic group. This was determined by several extralinguistic factors: the shift away from a nomadic economy contributed to the disintegration of national identity, and writers were compelled to assume the role of mythmaker, conveying spiritual wisdom in their work. The purpose of this study is to analyze S. Sanbayev's story "White Aruana" from a mythopoetic perspective. It has been established that the camel myth is one of the most significant in Central Asian literature. Camel mythology derives from nomadic notions of cosmogony and the primordial indivisibility of life and death. Aruana mythology, as a mythopoetic image, is linked to the archetype of the home — a sacred spatial code that can be analyzed in two aspects: the toponymic and the existential.
Key words: mythopoetics, archetype, contemporary Kazakh prose, Kazakh literature

Authors' contribution: The authors contributed equally to this article and preparation of the article text.

Conflict of interests: The authors declare no conflicts of interest.

Article history: Received: 23.12.2025; Accepted: 15.0212.2025.

For citation: Ismailova, Sh.A., Bayanbaeva, Zh.A., & Daribaev, S.D. (2026). The Birth of a Myth in the Literary Works by S. Sanbaev: the Archetype of Cosmogony. *RUDN Journal of Language Studies, Semiotics and Semantics, 17*(2), 836–848. https://doi.org/10.22363/2313-2299-2026-17-2-836-848 EDN NIKYCO

© Ismailova Sh.A., Bayanbaev, Zh.A., Daribaev S.D., 2026
This work is licensed under a Creative Commons Attribution-NonCommercial 4.0 International License
https://creativecommons.org/licenses/by-nc/4.0/legalcode

Рождение мифа в творчестве С. Санбаева: архетипика космогонии

Ш.А. Исмаилова[1], Ж.А. Баянбаева[2], С.Д. Дарибаев[2]

[1]Казахский государственный медицинский университет имени С.Д. Асфендиярова, *Алматы, Республика Казахстан*
[2]Казахский национальный университет им. аль-Фараби, *Алматы, Республика Казахстан*
✉sholpan_is@mail.ru

Аннотация. Мифопоэтика — значимая система элементов современного казахстанского литературного дискурса. Связь мифа с литературой прослеживается на всех уровнях художественного целого. Через мифотворчество казахских писателей происходит то структурирование мира, на котором базируется само национальное сознание народа. Если советская литература рассматривала мифологичность художественного дискурса как изначальное свойство «младописьменной литературы», то для казахского писателя это был акт интроспекции, осмысления своей собственной души как души представителя аутентичного этноса. Это детерминировано рядом экстралингвистических причин: отход от кочевого типа хозяйствования способствовал распаду национальной идентичности, и писатель должен был взять на себя роль Мифотворца, несущего в своем творчестве духовную информацию. Цель настоящего исследования — проанализировать повесть С. Санбаева «Белая Аруана» в мифопоэтическом аспекте. Установлено, что миф о верблюде — один из наиболее значимых в центрально-азиатской литературе. Мифологема верблюда восходит к представлениям кочевников о космогонии, изначальной неделимости жизни и смерти. Мифологема аруаны как мифопоэтического образа связана с архетипом дома — сакрального пространственного кода, который закономерно рассматривать в двух аспектах: собственно топонимическом и экзистенциальном.

Ключевые слова: мифопоэтика, современная казахстанская повесть, казахская литература

Вклад авторов: вклад авторов равнозначен на всех этапах исследования и подготовки текста статьи.

Конфликт интересов: авторы заявили об отсутствии конфликтов интересов.

История статьи: дата поступления: 23.12.2025; дата приема в печать: 15.02.2026.

Для цитирования: *Ismailova Sh.A., Bayanbaev, Zh.A., Daribaev S.D.* The Birth of a Myth in the Literary Works by S. Sanbaev: the Archetype of Cosmogony // Вестник Российского университета дружбы народов. Серия: Теория языка. Семиотика. Семантика. 2026. Т. 17. № 2. С. 836–848. https://doi.org/10.22363/2313-2299-2026-17-2-836-848 EDN NIKYCO

Introduction

20th-century literature is distinguished by the unique nature of artistic quest, particularly the synthesis of realism and neo-mythologism. Mythologism often becomes a figurative embodiment of an intellectual principle that does not contradict the tenets of a realistic portrayal of reality. Mythology remains for literature the basis of motifs, images, and archetypes. Many works pose pressing moral questions to their readers.

The relationship between myth and literature is a problem that still requires scientific understanding. Myth is interpreted as a narrative of a special type, born of spontaneous artistry and having aesthetic significance. The direction of the myth is always harmonizing: it is a victory over the forces of Chaos, as noted by E. Meletinsky. By accepting conventional forms of life, myth contributes to the

creation of some higher reality, which is perceived by the bearers of a given tradition as a source of the ideal. Mythological thinking is affective, concrete, and sensuous: these very qualities ensure the poetics of myth. By virtue of its syncretism, myth cannot be reduced to strictly artistic phenomena; nevertheless, its artistic origin is undeniable. According to E. Meletinsky, world literature has gone from universal symbolism to realism, having overcome three stages: mythologism, demythologization, and remythologization [1].

Remythologization contrasts the poetry of myth with the prose of everyday life; it is in myth that views that are universal are embodied.

What is myth? It is a collective fantasy capable of reflecting reality in the form of a system of concretely sensual, tangible personifications, conceivable to the bearer of archaic consciousness as reality itself. In terms of this system, the world is not only described but also known. Literature, science, and religion differentiated themselves from mythology at later stages of civilization. When included in new structures, mythological models are reinterpreted. Relics of mythological thinking are residually preserved in the consciousness of modern society. Myth is a symbolic expression of events that took place in the history of people during the archaic stages of their development. Even Socrates said that a poet, to be a true poet, must create myths, not arguments. Literature as a system of symbols does not appeal to any particular myth, but to Myth as a generic concept; in a sense, it itself is Myth. Art in the modern world replaces the mythology of primitive society: literature can rise to a comprehensive reality.

American researcher D. Vickery offers four principles of mythological literary criticism. According to the first, the creation of myths is inherent in human thinking itself. It meets the basic needs of human consciousness. The second principle says that myth is the fundamental principle from which literature arose historically. The third principle suggests that myth not only initiates artistic creation but also provides images for its interpretation. The fourth principle is that the ability of literature to move people is due to its mythological qualities. Thus, the role of literature lies in its expression of the essence of myth — namely, creating for a person a sense of the meaning of his life.

Through the myth-making of Kazakh writers, a structuring of the world took place, on which the very national consciousness of the people is based. If Soviet literature considered the mythological nature of artistic discourse as the original property of “young written literature”, then for the Kazakh writer it was an act of introspection, understanding his own soul as the soul of a representative of an authentic ethnic group. This was determined by several extralinguistic factors: the departure from the nomadic type of management contributed to the disintegration of national identity, and the writer had to take on the role of the Myth-Creator, carrying spiritual information in his work.

According to Carl Jung, myth-making is a real function of a civilized person [2. P. 81]. National identity is becoming an important task of our time: its function is to unify the country’s population, united by common interests and ideas about life. In this context, mythology acts in this case as a “unifying principle”.

The revival of the historical experience of the past, national traditions, and customs is involved in the reconstruction of the national worldview, which acts as a mechanism for the spiritual consolidation of the ethnic community. It is through mythology that the individual plunges into the past, which creatively transforms him and returns him to reality mentally whole and protected. According to Jung, the essence of immersion in mythology is a return to one's origins.

Methodological basis

The first attempts to homogenize the Kazakh nation belong to the writers of the 20th century: it is they who sought to create harmony out of chaos, to bring an existential sound into literature.

The connection between Kazakh literature and mythology is genetic: art originates from the depths of traditional mythology, its motifs and archetypal images, including cosmogony as the birth of a new world, the eternal confrontation between the forces of good and evil, descent into the underworld, wandering into labyrinth, the motif of an orphan child, a journey to the end of the world in search of sacred knowledge or finding a treasure as an attribute of the sacred. Art grows from the very depths of the human spirit: myth opens a path for the individual to create a new world.

Kazakh literature of the second half of the 20th century is anthropocentric, but the person is not considered in it as an immanent, self-contained being; he is included in the history of mankind, his existence being is universal and archetypal. The Kazakh writer is a thinker who is aware of the power of retransmitted images.

Kazakhstan has developed its own mythological school. It goes back to the folkloric works of Ch. Valikhanov, A. Baitursynov, M. Auezov, S. Kaskabasov, and E. Tursunov. Thus, in the works of S. Kaskabasov we find mythological and ritual structures of folklore. In the work "Kazakh non-fairytale prose" the researcher uses such terms as myth, mythical consciousness, mythical time, and mythological thinking [3].

According to S. Kondybay, the Kazakh equivalent of the term myth is *"ertegi",* in literal translation — a legend, a fairy tale. This word goes back to the proto-form ert, "action", "order": a body of knowledge of an ideological nature in ancient times meant certain rules governing social behavior. In Old Kazakh, this signified "myth". Subsequently, when myth ceased to be realized and lived as a reality, the term acquired modern actual semantics — "fairy tale". It gave impetus to the emergence of such formulas as "a long time ago, in time immemorial" (ertede)[1].

Researcher E. Tursunov introduces the concept of secondary mythologization into scientific circulation. He sees the ideological core of Kazakh mythology in the cult of ancestors (worship of aruakhs), noting that it is typical for Kazakhs to comprehend themselves in the world in the aspect of mythological time. The present is perceived in unity with the past, and the past is perceived as the present. This is

[1] Kondybai, S. (2005). Kazakh Mythology: A Brief Dictionary. Almaty: Nurly Alem.

manifested, in particular, in ideas about the totem: the totem is not only the ancestor of humans (for example, the Wolf), but also their contemporary. The man, the bearer of the mythological consciousness, is convinced that he descends from the Wolf, but he himself is the Wolf. Mythological motifs, according to Tursunov, appear not only in folklore but also in literature [4].

Sh. Yeleukenov believes that myth is the cornerstone of the Kazakh epic — it contains the rudiments of the people's worldview. Through the work of akyns and zhyrau, the myth entered the "flesh of Kazakh prose" [5].

The achievement of the Kazakh mythological school was the fundamental work of S. Kondybay "Mifke oralu" ("Return to the myth"). According to scientists, the myth is the focus of the cultural memory of mankind, it is the possibility of creating a single field of the Kazakh scientific paradigm. He believes that the reconstruction of Kazakh mythology should take place at the intersection of such humanities disciplines as linguistics, literary criticism, folklore, philosophy, psychology, ethnography, archeology, cultural studies. Understanding the Turkic world itself is possible through the comprehension of the Kazakh language itself. Mythology, according to S. Kondybai, is "the brain of any independent nation"[2]. National mythology, according to M.M. Auezov, allows a people to answer the etiological and teleological questions ("Who are we?", "Where are we going from?", "What are we striving for?") and is the foundation for Kazakhstan's cultural self-sufficiency.

Literature has always been the material that "nourishes" the theory of myth. Myth is born in the depths of human consciousness, it is supraindividual, therefore its archetypal core is universal. Differences in mythology are due to the life situation of ethnic groups. The retransmission of archetypal motifs leads us to the idea of the existence of a single primary source—the Creator.

"The idea of some divine essence, the power of which surpasses any human conception, was always widespread everywhere and manifested itself… unconsciously, since it is precisely this that is the archetype. …It seems that we have no choice but to consciously recognize the idea of God" [2. P. 35]. This study does not adhere to a single mythological school — ritual, linguistic, structuralist or archetypal. However, we will turn to the methodological basis of each of them during our analysis to solve a specific task — to identify the mythological substratum in the work of S. Sanbaev "White Aruana".

If the Kazakh writer expresses the national spirit in the myth, it is worth noting that the "national" is not an immanent spiritual system, it is expressed in the manifestation of the existential in a person. Myths make up the sum of knowledge that people use: only an appeal to accumulated experience allows a person to continue his individual existence. Forgetting the contents of the collective memory is tantamount to a regression to a state of nature, sin, or catastrophe [6].

The return to myth, as demonstrated in Kazakh literature, is an awareness of another, divine world or the world of ancestors, since literature is a continuation of

[2] Kondybai, S. (2005). Kazakh Mythology: A Brief Dictionary. Almaty: Nurly Alem.

the mythological narrative: going beyond time, carried out with the help of reading, is what brings together the functions of myth and literature.

The mythological way of mastering the world has become for Kazakh literature a way of creative transformation of reality. The Kazakh worldview is characterized by such features as a sense of freedom, an inextricable connection with nature, a religious community with Being, and predestination.

Camel Mythology in Nomadic Culture

According to O.K. Abisheva, in the modern literature of Kazakhstan there is an image that goes back to myth as an archaic-traditional form of thinking [7]. This is the image of the camel, represented by two variants of manifestation: Bactrian (two-humped camel) and Aruana (one-humped camel). The camel is one of the animals that formed the basis of the nomadic type of management; the whole way of life of the peoples of Central Asia relates to it. The camel was domesticated in pastoral as early as culture five thousand years BC. Being the main support of the nomad's economy, the camel is distinguished by its excellent adaptability to environmental conditions: it endures heat, goes without water for a long time, is unpretentious in food, resistant to cold, moves across the steppes and dunes. Its main need is space, expanse. The symbolism of the Bactrian and Aruana goes back to ancient archetypal images of nomadic culture. In the strict sense, this is a mythologeme—an element of a literary text, a collective pattern contained in the subconscious of the people, relayed both in a myth and in a literary work. Thus, this is an archaic motif, based on which a mythological or literary plot can be reconstructed. The image of the camel is one of the mythologems of Turkic culture.

We consider the image of a camel in its connection with the higher spheres of being. Archaic consciousness perceives all phenomena and elements as personifications of various animals and deities; in the hierarchy of these figurative and symbolic representations, the camel occupies a special place as an intermediary between man and Heaven. The symbolism of the image of a camel goes back to the idea of an indivisible single Cosmos; it is created from nothing and is itself the fundamental principle. In the worldview of the Turks, he is a symbol of the primordial world [8].

In the mythology of the Turkic peoples, the camel has a special place. So, in The Book of My Grandfather Korkut [7], common for the Turkic tribes, there is a legend about a winged camel, on which the baksy travels around the four cardinal points, escaping from death. According to the legend, it was predicted to a great old man that he would die if he remembered death. In all corners of the world he sought salvation, however, wherever he went, everywhere people dug graves. Then he killed his companion—a magical camel —and made a musical instrument—kobyz —from the skin around her neck. According to the Kazakh ethnographer A.K. Akishev, the myths about the heavenly camel (sometimes it is a fire-breathing camel) embody the image of the solar thunder god. In Mongolian fairy tales, the Tengrian Thunderers traveled on camels; in the mythology of the Iranians, the

image of a camel was associated with the image of a snake-dragon; he was identified with the cosmos [8. P. 74].

Among the Kazakhs, the camel was the most expensive sacrifice brought to higher powers. It was also considered a talisman: its bones were used in healing magic, meat was valued as apotropaic. The camel embodied the idea of fertility; in this sense, his image was associated with male power. It is a symbol of endurance, vitality, power, perseverance. In the art of the nomads of Central Asia, the image of a camel is extremely widespread. The first rock carvings, according to L.I. Rempel, appear in the 2nd millennium BC [9]. These are scenes of running and fighting camels found in Kazakhstan, Uzbekistan, and Turkmenistan. A censer was discovered on the territory of Semirechye, the symbolism of which A.K. Akishev calls cosmogonic. In the center of the bowl are images of camels with their heads to each other; their four humps symbolize the four cardinal directions. Their pairing could also mean gender differences; in this case, they personify the fruitful forces of the earth, the dual unity of cosmic principles, the dialectics of Indo-Iranian mythology itself [8. P. 75].

An image of an animal with the head of a camel was also found in modern-day Uzbekistan. Various items — statuettes, coins, and gems with images of a camel — testify to the existence of a cult tradition of this animal. I.L. Rempel identifies several significant aspects in the interpretation of the camel: the cultic, the dynastic, and the ludic. Its image was correlated with the indomitability and power of the god of war and victory. “In this aspect, the camel people and its winged double are, as it were, a cosmic deity. It also served as a sign of royalty and greatness. All three aspects named above—cultic, dynastic, and ludic—correspond at the same time historical development of images of thinking: mythological, epic, folklore” [9. P. 97–98].

The image of the camel in the cultures of Central Asia was considered sacred. In the works of writers of this region there are plots that retransmit it. The recurrence of this image also indicates that it is a kind of mythopoetic invariant, which contains many meanings in a reduced way.

Satimzhan Sanbaev’s “White Aruana”

The prose of Satimzhan Sanbaev occupies an important place in the contemporary literary process of Kazakhstan. Its distinctive feature is the artist’s excellent knowledge of the national way of aul (village) life. The author’s works (“And the Eternal Battle”, “When They Thirst for a Myth”, “Wells of the Hot Valleys”, “The Old Pipe”, “The Black Whirlwind”, etc.) are characterized by deep philosophicality, contemplation, and an interweaving of the semantic and figurative levels of the artistic word. “White Aruana” is one of the early works of the writer. The story deals with love for the motherland, loneliness, the fragility of human happiness. The work is full of associations and cultural codes. The plot of the story unfolds against the backdrop of the expanse of the steppe; bright author’s sketches allow the reader to plunge into the rhythm of aul life.

Sanbaev's work is materialized in the style of a folk song epic; its fabric is permeated with poetic repetitions, inversions, lyrical monologues, which makes it possible to correlate the text, both compositionally and intonationally, with the authentic tolgau genre—contemplation-reflection [10]. The chosen genre style gives the author the opportunity to recreate the intonation of thoughts about the universe, life, nature. The folklore basis of the text is also the genre of legend (the central motif is the devotion of animals to their native land).

The folklore technique of parallelism is the basis of the compositional drawing of the story: the fates of Myrzagali and Auruana are similar, which emphasizes the idea that a person is close to his roots.

The main characters of the story are old Myrzagali and a beautiful white camel; their destinies are intertwined, in a certain sense contaminated, and through the attitude of one subject of the narrative, we gain access to the psychological experiences of another. It is through the image of Aruana that the reader is immersed in complex human relationships. The artistic fabric of the work remains intact; it does not delaminate, held together by the author's presence in the text.

Myrzagali, who returned from the war infirm, is increasingly thinking about the emptiness and meaninglessness of his life: "He recalled if there was a minute when he tried to leave life ... He tried but could not. And no one, except for the old woman who lived this long and boring life with him, knew about his pain"[3].

His only consolation is a small white camel, personifying for the old man everything pure and bright that was in his life. Aruana does not share Myrzagali's feelings. She yearns for her homeland, for the once abandoned Mangistau. The story, which has a complex subtext, relates to the folklore legend of the Kazakhs about the fidelity of animals to their native land. For Aruana, there is not and cannot be "here-being." The call of another life draws her to her native land. "It was the wind of the country of strange mountains, which she remembered. By the sea, by the Mek mountains, such a wind is born. The mountains are filled with strange, quiet sounds and, as it were, rush upward, straightening up, stretching waking up and—becoming even more impractical. Sensitive saigas tremble, turning their ears and sniffing this resonant dry wind, and run to the north, from the sun, lowering their heads into your short shadows; they hide their beaks in the grass; disturbingly looking at the shimmering sky, the shepherds drive the herds closer to the wells. The dry wind comes to Mangistau for a long time ... on the chest and stomach of the Aruana. She went further into the steppe, dissolving in a sultry haze"[4].

Her homeland beckoned Aruana, and she ran away, but Myrzagali returned her time after time. The plot lines of the old man and the white camel do not just form a parallel; they are superimposed one on top of the other: the history of its owner also passes through the history of Aruana. In Aruana, unhappy and cut off from home, found herself alone among strangers, Myrzagali sees himself. It is quite obvious here that there are some analogies, a significant cross-talk of human and

[3] Sanbayev, S. (2009). Collection of Works in 6 Volumes. Stories and Short Stories. Astana: Agroidat.
[4] Ibid.

animal realities: in perceiving the camel's experience, the reader subtly reframes this perception onto the level of human reactions, emotions, and feelings [11].

Myrzagali falls ill and ends up in the hospital; Aruana is subjected to execution: so that she does not run away and gives birth, the veins of her eyes are cut. But the memory of the home does not leave her. When a dry wind comes to the steppe, Aruana gives voice to her camel and runs towards the wind. She is driven by a great instinct; smooth and swift her run. Hearing the clatter of hooves overtaking her, the camel stumbles, makes a leap over a deep ravine and falls soundlessly. Myrzagali cries for a long time, living through his pain. Now he is bringing home his Aruana's camel, a baby who knows only this house, who will be devoted to him for the rest of its life. "Ahead, swaying on long straight legs and stumbling, walked a willow. He wept thinly and pitifully, looking around, not yet realizing that he had lost his mother forever. He walked obediently in front of him, because ahead of him lay his homeland, lay the village where he was born and where yes, he will run away from here, like his mother, who all her life went to Mangistau"[5].

As we noted above, the mythopoetic image of a camel is associated with the idea of originality and indivisibility. The mythological substratum of the image is a kind of prolepsis in the story about the eternal return to the origins. The image of a white camel is associated with one of the basic universal archetypes — the archetype of home.

We consider archetypes as patterns of the collective unconscious retransmitted in culture, carrying universal knowledge and ideas of a person about the world. This information was put into the core of the archetype through the centuries-old repetition of historical retranslation experienced by society from the archaic stage of its existence. Even though the historical context of the existence of the archetype changes, it remains an imprint of the original ideas of man about the world — according to Z. Freud, culture itself is built from the bricks of a long-forgotten worldview [12. P. 163]. Archetypes are constant, but their periphery is relatively mobile and adaptive.

"When Homo erectus, over the course of several millennia turned into Homo sapiens, three of its hypostases were simultaneously formed in its archetype:

1) the human mind appeared;

2) consciousness, thinking, and speech were formed;

3) the significance of his existence appeared," writes A.K. Kairzhanov [13. P. 9].

An archetype is an element of the individual and collective psyche. It is versatile and dynamic. At the level of typical structures, such as motif and image, it is associated with myth — a prototype that repeats itself in culture.

According to A.F. Losev, myth is an intensely experienced, maximally concrete reality. Moreover, the myth is not a product of thought itself, that is, it does not represent an ideal being, or an abstraction. Myth is affective and tangible. It is not scientific, yet mythical truth is inherent. It is not metaphysical, because repelled from sensory experience and deals with "here-being" (this-worldliness,

[5] Sanbayev, S. (2009). Collection of Works in 6 Volumes. Stories and Short Stories. Astana: Agroidat.

according to Losev's formulation). Mythology is "vital and visual." In this respect, it is not allegorical, that is, it is not seen as a set of schemes or parables, but symbolic. According to Losev's final formula, myth is a miracle, or an expanded magical name [14].

Researcher Yu.V. Domansky proposes to understand archetypes as primary plot schemes, images or motifs that arose in the mind of a person at the earliest stage of human development [15. P. 11].

Home is one of the key spatial archetypes. Space is a category that is permanently filled with valuable content. On the one hand, this is everything that is outside the individual and therefore requires development — both material and mental. However, this archetype has another important structure — the existential one, and in this respect, it becomes the receptacle of the human soul itself. "A person's life is in space, which inevitably implies the question of the place of life — a country, a region, a city: what is it and what is the meaning of my life here? A person cannot bear the semantic and value emptiness of the place where he lives, he urgently needs to comprehend it and value it in order"[6]. According to L.N. Gumilyov, "Even primitive man knew the boundaries of his habitat, the so-called feeding and accommodating landscape in which he himself lived, his family and his tribe lived"[7].

Thus, in Sanbaev's story, we can consider the "home" archetype in two aspects. The first is the implementation of the spatial code, place, homeland. It correlates with the motive of return and is actualized in the work in the form of a white aruana. As we mentioned above, the camel and Myrzagali are in the work not only in a relationship of parallelism, but also in a certain contamination. This conclusion allows us to conclude that Myrzagali also needs to return; however, in this case, it is appropriate to talk not about a return to a specific topos (which took place; the old man returned to his native village after the war), but about a return to the origins — that mental and spiritual state that could be characterized as blissful for the hero. Myrzagali will have to return to the house of his own soul, to find peace and tranquility. Restoring the plot of the story, we see that the life of the protagonist is filled with sad events: he returned from the war infirm; wife cheated on him with a neighbor; the daughter living in the city does not communicate with her father. Aruana becomes for the old man the personification of the meaning of life. As an animal strives to return to its native Mangistau, so Myrzagali seek to restore integrity to his life. The camel is a cosmogonic symbol. The meaning of this mythologeme can be explicated through the motifs of cosmogonic originality, death and rebirth. Being an element of nomadic culture, the image of a camel is subject to the mythopoetic ideas of the Kazakh people about the rhythms of life-death-life, subject to the circular rotation of time. Such ideas are conditioned by the habitat of the ethno-collective itself: the world of the steppe, commensurate with the cycles of nature, obeys the laws of circular movement in space: from dzhailau (herds moving to summer pastures) to kuzdeu (autumn pastures), from kuzdeu to kystau

[6] Gumilyov, L.N. (2019). *Ancient Rus' and the Great Steppe*. Moscow: Iris-press. (In Russ.). P. 5.
[7] Ibid. P. 5.

(winter camps). The nomad's world is polycentric: it is centered around a few ontologically significant places. However, the image of Aruana is associated with the theme of the one homeland to which she aspires. The camel is a dichotomous symbol. Realizing in its complex the semantics of the beginning, it also absorbs the meaning of the end. According to the research of A.K. Akishev, the image of the camel also carries the motive of death. The death of Aruana, in this sense, is a foregone conclusion: it should mark the initiatory transition that the protagonist has to make in order to regain his original integrity. This is the node level of the plot, leading the protagonist to catharsis. In this regard, the image of a camel born by Aruana is also significant: in the story it acts as a mythopoetic symbol of a new beginning, a new cosmogony in the soul of Myrzagali. In this case, we consider the "home" archetype in an existential aspect: it actualizes the idea of returning to the original state of the human soul, reborn and transformed.

Conclusions

The camel myth is one of the most significant in Central Asian literature. Camel mythology traces back to the ideas of nomads about cosmogony, the original indivisibility of life and death. It is no coincidence that the image of a camel, according to the views of the ancient nomads, incorporates elements of all known living beings. Folkloric and literary representations of this image are quite frequent. In the work of S. Sanbaev "White Aruana" the myth of the eternal return home is born. Three times the white Aruana runs away to Mangistau calling her; the third time (we note that the idea of finiteness is also embedded in the sacredness of the triad) turns out to be the last: the camel dies, realizing a significant component of her own figurative symbolism: the motive of death.

Being a cosmogonic symbol, the camel also realizes the semantics of death. In the ideas of nomadic peoples, death is a transition to a new cyclic segment of existence. Acting in the text as a kind of parallel to the main character, Myrzagali, aruana marks for him the transition to a new level. This meaning is supported by the image of a camel as the beginning of a new life cycle. The mythologeme of Aruana as a mythopoetic image is associated with the archetype of the home — a sacred spatial code, which is natural to consider in two aspects: toponymic and existential.

The toponymic level of realization of the archetype is associated with the image of the white camel striving to return home. Existential — with the mental and spiritual space of Myrzagali, trying to restore integrity to his soul, and life — meaning. Thus, both protagonists return home. It is symptomatic that the white aruana dies: this mythopoetic image internally completes the motif of death and rebirth. Dying, the camel again becomes part of the universe. The semantics of the image as a cosmogonic symbol is updated. Having survived the death of his beloved, Myrzagali is also reborn: he returns home as a reborn person.

Thus, the camel myth develops not only the idea of devotion to one's native land, but also of devotion to one's own human essence — this idea becomes a concept-forming one for the entire narrative space of S. Sanbaev's story.

References

1. Meletinsky, E.M. (2000). *The Poetics of Myth.* Moscow: Izdatelskaya Firma “Vostochnaya Literatura” RAS. (In Russ.). EDN: VNNIUT
2. Jung, K.G. (2020). *Archetypes and the Collective Unconscious*. Moscow: AST. (In Russ.).
3. Kaskabasov, S.A. (1990). *Kazakh Non-Fairy-Tale Prose*. Almaty. (In Russ.).
4. Tursunov, E.D. (2004). *The Origin of the Bearers of Kazakh Folklore*. Almaty: Daik-Press. (In Russ.).
5. Eleukenov, Sh.R. (1987). *From Folklore to the Epic Novel*. Monograph. Almaty: Zhazushy. (In Russ.).
6. Tursunov, E.D. (2001). *Ancient Turkic Folklore: Sources and Formation*. Almaty: Daik-Press. (In Russ.).
7. Abisheva, O.K., & Kogay, E.R. (2016). *Literature of Central Asia. Archetypes, Mythologemes, Concepts*. Almaty: Kazak Universiteti. (In Russ.).
8. Akishev, A.K. (1984). The Image of the Camel in the Legends of Central Asia. In: *Ethnography of the Peoples of Siberia* (pp. 69–76). Novosibirsk. (In Russ.).
9. Rempel, L.I. (1977). Fragment of a Bronze Statue of a Camel from Samarkand and the Winged Camel of Varakhsha (towards the Question of the Nature of Sogdian Art). In: *Central Asia in Antiquity and the Middle Ages. History and Culture* (pp. 70–95). Moscow: Nauka. (In Russ.).
10. Arukenova, O.A. (2020). Dichotomy in Satimzhan Sanbaev's Story “White Aruana”. *Keruen*, 67(2), 1–6. (In Russ.). https://doi.org/10.53871/2078-8134.2020.2-05 EDN: MWBQCE
11. Abdullina, L.I. (2011). The Lyric-Epic Nature of S. Sanbaev’s Story “White Aruana”. *Kazakh American Free University Bulletin*, 2, 195–199. (In Russ.).
12. Freud, Z. (2021). *Introduction to Psychoanalysis*. Moscow. (In Russ.).
13. Kairzhanov, A.K. (2020). *Palaeoturcica. Sign and Semantics. Myth and Culture*. Moscow: Altair. (In Russ.).
14. Losev, A.F. (2001). *The Dialectics of Myth*. Moscow: Mysl'. (In Russ.).
15. Domansky, Yu.V. (2001). *The Meaning-Forming Role of Archetypal Meanings in a Literary Text*. Tver: Tver State University publ. (In Russ.). EDN: TJSQZZ

Список литературы

1. *Мелетинский Е.М.* Поэтика мифа. М. : Восточная литература РАН, 2000. EDN: VNNIUT
2. *Юнг К.Г.* Архетипы и коллективное бессознательное. М. : АСТ, 2020.
3. *Каскабасов С.А.* Казахская несказочная проза. Алма-Ата : Наука КазССР, 1990.
4. *Турсунов Е.Д.* Происхождение носителей казахского фольклора. Алматы : Дайк-Пресс, 2004.
5. *Елеукенов Ш.Р.* От фольклора до романа-эпопеи. Алма-Ата : Жазушы, 1987.
6. *Турсунов Е.Д.* Древнетюркский фольклор: истоки и становление. Алматы : Дайк-Пресс, 2001.
7. *Абишева У.К., Когай Э.Р.* Литература Центральной Азии. Архетипы. Мифологемы. Концепты: монография. Алма-Ата : Казахский университет, 2016.
8. *Акишев А.К.* Образ верблюда в легендах Центральной Азии // Этнография народов Сибири. Новосибирск, 1984. С. 69–76.
9. *Ремпель Л.И.* Фрагмент бронзовой статуи верблюда из Самарканда и крылатый верблюд Варахши (к вопросу о природе согдийского искусства) // Средняя Азия в древности и средневековье. История и культура. М. : Наука, 1977. С. 70–95.
10. *Арукенова О.А.* Дихотомия в повести Сатимжана Санбаева «Белая Аруана» // Keruen. 2020. Т. 2. № 67. С. 1—6. https://doi.org/10.53871/2078-8134.2020.2-05 EDN: MWBQCE
11. *Абдуллина Л.И.* (2011) Лироэпическая природа повести С. Санбаева «Белая Аруана» // Вестник КАСУ. 2011. № 2. С. 195–199.
12. *Фрейд З.* Введение в психоанализ. М. : Азбука, 2021.
13. *Каиржанов А.К.* Palaeoturcica: Знак и семантика. Миф и культура. М. : Алтаир, 2020.
14. *Лосев А.Ф.* Диалектика мифа. М. : Мысль, 2001.

15. *Доманский Ю.В.* Смыслообразующая роль архетипических значений в литературном тексте. Тверь : Тверской государственный университет, 2001. EDN: TJSQZZ

Information about the authors:
Sholpan A. Ismailova, PhD student, Senior Lecturer, lecturer in the Department of Language Disciplines, Asfendiyarov Kazakh National Medical University (94 Tole Bi Str., Almaty, the Republic of Kazakhstan, 050000); *Research interests*: theory and methods of teaching Russian language and literature, the use of ICT in language teaching, literary criticism, and Russian literature; *e-mail*: sholpan_is@mail.ru
ORCID: 0000-0002-6476-5702.
Zhadyra A. Bayanbayeva, PhD in Philology, Deputy Dean of the Faculty of Philology for Research and Innovation and International Cooperation, Al-Farabi Kazakh National University (71 al-Farabi Ave., Almaty, the Republic of Kazakhstan, 050040); *Research interests*: Theory and methods of teaching Russian language and literature, literary studies, Russian literature; *e-mail*: bayanbaevazhadra@mail.ru
ORCID: 0000-0002-1710-0791.
Samalbai D. Daribaev, PhD in Philology, Senior Lecturer, Department of Kazakh Literature and Literary Theory, Al-Farabi Kazakh National University (71 al- Farabi Ave., Almaty, the Republic of Kazakhstan, 050040); *Research interests*: Kazakh literature, ethnographic foundations of Kazakh folklore, scientific code (values), and folk pedagogical conclusions; *e-mail*: samal.daribaev@gmail.com
ORCID: 0000-0002-3551-770X.

Сведения об авторах:
Исмаилова Шолпан Аймурзаевна, докторант PhD, лектор кафедры языковых дисциплин, Казахский национальный медицинский университет имени С.Д. Асфендиярова (050000, Республика Казахстан, г. Алматы, ул. Толе Би, д. 94); *сфера научных интересов*: теория и методика преподавания русского языка и литературы, применение ИКТ в обучении языкам, литературоведение, русская литература; *e-mail*: sholpan_is@mail.ru
ORCID: 0000-0002-6476-5702.
Баянбаева Жадыра Амангельдиевна, кандидат филологических наук, заместитель декана филологического факультета по научно-инновационной работе и международному сотрудничеству, Казахский национальный университет им. Аль-Фараби (050040, Республика Казахстан, г. Алматы, пр. аль-Фараби, д. 71); *сфера научных интересов*: теория и методика преподавания русского языка и литературы, литературоведение, русская литература; *e-mail*: bayanbaevazhadra@mail.ru
ORCID: 0000-0002-1710-0791.
Дарибаев Самалбай Дарибаевич, кандидат филологических наук, старший преподаватель кафедры казахской литературы и теории литературы, Казахский национальный университет им. Аль-Фараби (050040, Республика Казахстан, г. Алматы, пр. аль-Фараби, д. 71); *сфера научных интересов*: казахская литература, изучение этнографических основ казахского фольклора, научного кода (ценностей) и народных педагогических выводов; *e-mail*: samal.daribaev@gmail.com
ORCID: 0000-0002-3551-770X.