



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

EDN: LHQFQK

Научная статья / Research Article

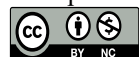
Алгоритмический поворот в интернет-дискурсе

Н.А. Ахренова  

Российский университет дружбы народов, Москва, Российская Федерация
[✉nakhrenova@mail.ru](mailto: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 антропологистики. 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 антропологическая лингвистика

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 антропологистики, которая исследует фундаментальные изменения, вносимые цифровой средой в понимание языка, культуры и их неразрывной связи.

Традиционная антропологистика 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 антропологистики к цифровой знаменует собой смену парадигмы по нескольким ключевым направлениям:

1. Трансформация сообщества. Классическая антропологистика работала с территориально-локализованными сообществами, чьи языковые практики были неразрывно связаны с физическим местом, общими ритуалами и прямым («живым») межличностным взаимодействием («речевой общиной»). Цифровая антропологистика 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 актер, 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) проведен 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 антропологистика становится не просто новым направлением, а необходимостью, поскольку цифровая среда превратилась в основную среду бытования языка для миллионов 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. Двухязычный 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 репозиторий, 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 Commu- nication Corpora	H2H (Human-to- Human)	Тексты, созданные людьми для общения с людьми до 2021 г. (форумы, блоги, лич- ная переписка, ком- ментарии в соцсетях)	512 847	3 427
Корпус промптов / Prompt Corpora	H2A (Human-to-AI)	Диалоги пользовате- лей с ChatGPT, за- просы к генератив- ным нейросетям	463 291	5 891
Корпус ИИ-генераций / AI Generation Corpora	A2H (AI-to-Human)	Тексты, сгенериро- ванные ИИ и марки- рованные как тако- вые (ответы ChatGPT, статьи, созданные нейросе- тями)	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-позиционных 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 антропологистики 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 испытывавшие 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. Используемые 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 раз- нообразие (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 А2Н 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 (Н2А) 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 (Н2А) характеризуются минимальной длиной предложения (русский: 8,9 слов, английский: 10,8 слов). Это соответствует коммуникативной стратегии «экономии усилий» в директивных жанрах. Большая длина английских промптов может объясняться аналитическим строем английского языка, требующим большего количества служебных слов для выражения тех же смыслов.

Тексты А2Н 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 А2Н 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 Н2Н (рус.) / H2N corpora (Russian)	Корпус Н2А (рус.) / H2A corpora (Russian)	Корпус А2Н (рус.) / A2N 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	жест	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 Н2Н (англ.) / H2N corpora (English)	Корпус Н2А (англ.) / H2A corpora (English)	Корпус А2Н (англ.) / A2N 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 Н2Н 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 Н2А 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 А2Н 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 большую 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 Н2Н 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 экстремистской 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 редакции 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, по- правил 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 мета-языковыми 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 антропологистика и трансформация языковой личности. Наконец, наши данные подтверждают тезис о необходимости смены парадигмы, развитый в рамках цифровой антропологистики. 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 эксплицитно выразить 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, редакция) 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 антропологистики.

Проведенное комплексное исследование, опирающееся на репрезентативный корпусный материал и межкультурный 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 антропологистика, 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 переформируют 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. T. 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-Higuaita 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: RJCVCN
11. Ахренова Н.А. Интернет-лингвистика: доминантный подход. М. : Рускайнс, 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-Higuaita, 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: RJCVCМ
 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.