

Персонотекст и нейротекст: критерии разграничения естественно-языковых и сгенерированных текстов на основе сопоставительного профилирования

Я. А. Дударева¹, А. А. Воронько²

Кемеровский государственный университет

Ул. Красная, 6, 650000, Кемерово, Россия. E-mail: ¹dudareva-yana@yandex.ru,

²nastyavoronko07@gmail.com

Статья посвящена поиску критериев, позволяющих отличить текст, сгенерированный нейросетью, от текста, продуцированного человеком. На основе комплексной методики автороведческого исследования с добавлением параметров (количество слов и коэффициент лексического разнообразия; коммуникативная тональность; ритм; графические признаки) осуществлялось профилирование реальной языковой личности по персонотексту и нейроязыковой «личности» по нейротексту от LLM «ГигаЧат». Проведенный анализ показал, что для персонотекста характерны смысловые отступления, свободное использование слов разных функциональных стилей, политематическое разнообразие лексики, орфографические и пунктуационные ошибки и опечатки, использование эмодзи. Диалогичность, изменение ритма текста, эмоциональный синтаксис в совокупности со всеми языковыми признаками позволили определить по персонотексту пол, возраст, образование, родной язык, ведущий тип восприятия и темперамент автора. Для нейротекста характерны следующие признаки: смысловое единство, трехчастная композиционная структура (введение, основная часть, заключение), использование книжной нейтральной лексики, шаблонность лексико-грамматических структур, повторяемость синтаксических конструкций, предпочтение полных осложненных предложений, нарочитая орфографическая и пунктуационная правильность. Монотонность, информативная тональность, слабая проявленность эмоций в совокупности с результатами анализа по всем языковым уровням позволили сделать вывод, что LLM «ГигаЧат» проявляет себя как нейроязыковая «личность» с деперсонализированными гендерными и возрастными характеристиками, русскоязычная, имеющая дигитальный тип восприятия (с возможностью настройки на другой тип в зависимости от поставленной в промпте задачи) и обладающая стилем общения флегматика.

Ключевые слова: большие языковые модели (LLM), сгенерированный текст, нейротекст, языковая картина мира, языковая личность, нейроязыковая «личность».

Personatext and Neurotext: Criteria for Distinguishing between Natural Language and AI-Generated Texts Based on Comparative Profiling

Ya. A. Dudareva¹, A. A. Voronko²

Kemerovo State University

6 Krasnaya St., 650000, Kemerovo, Russia. E-mail: ¹dudareva-yana@yandex.ru,

²nastyavoronko07@gmail.com

This article explores criteria for distinguishing text generated by a neural network from text produced by a human. Using a comprehensive authorship attribution methodology with additional parameters (word count and lexical diversity ratio; communicative tone; rhythm; graphic features), we profiled a real linguistic personality based on a persona text and a neurolinguistic "personality" based on a neurotext from the LLM "GigaChat." The analysis revealed that the personatext is characterized by semantic digressions, the free use of words of different functional styles, a polythematic diversity of vocabulary, spelling, punctuation errors, typographical errors, and the use of emoticons. Dialogue, changes in text rhythm, and emotional syntax, combined with all the linguistic features, made it possible to determine the author's gender, age, education, native language, dominant perception style and temperament from the personatext. The neurotext is

characterized by the following features: semantic unity, a three-part compositional structure (introduction, main body, conclusion), the use of bookish, neutral vocabulary, formulaic lexical and grammatical structures, repetitive syntactic constructions, preference for full, complex sentences, and deliberately correct spelling and punctuation. The monotony, informative tone, and lack of emotion, combined with the results of analysis across all language levels, led to the conclusion that the LLM "GigaChat" manifests itself as AI "linguistic personality" with depersonalized gender and age characteristics, Russian-speaking, possessing a digital perception style (with the ability to adjust to a different style depending on the task posed in the prompt), and possessing a phlegmatic communication style.

Key words: large language models (LLM), AI-generated text, neurotext, linguistic picture of the world, linguistic personality, neurolinguistic "personality".

Введение. Персонотекст и нейротекст в парадигме LLM-лингвистики (лингвистики больших языковых моделей)

В основе предпринимаемого исследования лежит проблема стирания границ между категориями «естественное» и «искусственное». Между естественно-языковым текстом, созданным человеком (то есть персонотекстом), и текстом, сгенерированным нейросетью (то есть LLM-текстом, нейротекстом), весьма трудно провести разграничение, поскольку нейротексты представляют собой производный продукт персонотекстов.

Сгенерированный текст является новым объектом изучения, получающим осмысление в рамках формирующегося направления, находящегося в сфере пересечения лингвистики и технологий искусственного интеллекта, которое можно обозначить как LLM-лингвистика, то есть лингвистика больших языковых моделей, или, используя термин Л. О. Микалеффа, «лингвистика нейросетей» [Микалефф 2025: 407]. Также сгенерированные тексты получают активное осмысление лингвистов в рамках предложенной Н. Д. Голевым парадигмы «новой вариантологии», то есть «вариативности текста через призму искусственного интеллекта» [Голев, Мельникова, Напреенко 2025: 201].

Предпринимаемое исследование по разграничению естественно-языкового и сгенерированного текстов вписывается в круг названных актуальных и быстроразвивающихся лингвистических направлений. Мы рассматриваем текст, сгенерированный «большими языковыми моделями, (англ. Large language models, LLM) – компьютерными программами, способными «понимать» и генерировать тексты на естественных человеческих языках» [Лягошина 2024: 113], как находящийся в отношениях множественной производности по отношению к производящим текстам – огромному массиву разнообразных естественно-языковых текстов. Персонотексты как исходные, первичные, естественные тексты и нейротексты, являющиеся вторичными, искусственными текстами, связаны отношениями нейродеривации, в которой принимает участие человек, формулирующий промпт (как текст-стимул), и нейросеть, выступающая инструментом генерации текста-реакции.

Участие человека в генерации текста нейросетью и использование результатов этой генерации обуславливают отнесенность персонотекста и нейротекста к оппозициям «естественность – искусственность», «оригинальность – копиальность», между полюсами которых формируется категория «гибридность». В качестве гибридных мы рассматриваем персонотексты с элементами сгенерированного текста и нейротексты с изменениями, внесенными человеком. Отсюда возникает новая типология текстов (естественно-языковые, гибридные, сгенерированные) и новая типология языковых личностей по отношению к степени участия нейросети в создании текста: оригинальные языковые личности, создающие персонотексты без использования нейросетей; копиальные языковые личности, полностью воспроизводящие результат генерации нейросети; языковые личности гибридного типа, которые формулируют промпт и используют сгенерированный нейросетью текст, внося в него формальные и/или содержательные изменения.

Принимая во внимание диффузность границ между естественно-языковым и искусственным текстом, мы формулируем гипотезу о том, что персонотекст и нейротекст можно дифференцировать между собой. Отсюда цель исследования – найти критерии, позволяющие установить отличия между текстом реальной языковой личности и текстом, сгенерированным нейросетью.

Настоящее исследование, направленное на поиск критериев разграничения нейротекста и персонотекста, является продолжением работы по сопоставлению сочинений по картине В. А. Серова «Девочка с персиками», сгенерированных разными нейросетями и написанных людьми. Поскольку данный материал готовится к опубликованию в монографии, приведем здесь значимые результаты, на которых осуществляется сопоставительное профилирование текстов.

На основе анализа сочинений было установлено, что для нейротекстов, сгенерированных LLM «ГигаЧат», характерны следующие особенности: композиционная полнота, логичность, цельность и связность, значительное количество жанровых клише, активное использование в тексте ключевых слов промпта, однотипность грамматических структур и монотонность, орфографическая и пунктуационная грамотность, непоследовательно проявилась рубрикация текста и наличие в нем фактологических ошибок. Нейротексты «ГигаЧата» самоидентифицируются, идентифицируются программой «Антиплагиат. Структура», а также определяются при перекрестной идентификации через LLM «DeepSeek» как сгенерированные нейросетью.

Нейротексты, сгенерированные LLM «DeepSeek-V3.2», обладают композиционной полнотой, логичностью, цельностью и связностью, ритмичностью, орфографической и пунктуационной грамотностью, содержат повторяющиеся грамматические структуры с одинаковым пунктуационным оформлением (отмечена повторяемость обобщенно-личных предложений, вводных слов и пунктуационного знака «тире»). Нейротексты LLM «DeepSeek»

определяются программой «Антиплагиат. Структура» как содержащие только некоторый процент ИИ-контента, при самоидентификации и перекрестной идентификации через «ГигаЧат» нейротекст определяется как написанный человеком.

Заметим, что полученные результаты лингвистического анализа сочинений нейросетей коррелируют с выводами, представленными в исследовании А. В. Громовой на материале текстов, сгенерированных нейросетями «GPT 4» и «Hailuo» от лица человека, имеющего неудачный опыт употребления наркотиков: монотематичность, стандартная композиция, соблюдение лексико-грамматических норм, построение синтаксически полных конструкций, прямой порядок слов, использование повтора в качестве средства связи, отсутствие орфографических и пунктуационных ошибок, активное использование «тире» [Громова 2025: 6].

Наблюдаемое сходство между нейротекстами позволяет отметить, что если для школьного сочинения перечисленные

характеристики текста соответствуют ожидаемому жанру, то сходство нейротекста о неудачном употреблении наркотиков с шаблонным отличным школьным сочинением (по признакам композиционной полноты, связности и отсутствия ошибок на всех языковых уровнях) демонстрирует его неестественность, искусственный характер. Отмеченное сходство между нейротекстами, полученными от разных LLM и по разным промптам, позволяет сделать вывод о том, что разные большие языковые модели генерируют тексты как инвариантная нейроязыковая «личность», являющаяся «новым типом автора» [Клушина 2025: 20]. Эксперименты лингвистов показывают, что имеются ситуации, когда различия между текстами, сгенерированными разными LLM, нивелируются и в практическом применении возникают такие примечательные совпадения, как сходство языковой личности бывшего наркомана с языковой личностью школьного отличника.

Тексты, сгенерированные нейросетью, но идентифицируемые алгоритмами как написанные человеком, демонстрируют диффузность границ между персонотекстом и нейротекстом. В таких текстах даже без редактуры со стороны человека наблюдается гибридизация, то есть сочетание искусственности и естественности: шаблонности, правильности, повторяемости слов и грамматических структур с их одинаковым пунктуационным оформлением и живого языкового оформления.

Таким образом, анализ сочинений, сгенерированных LLM «DeepSeek» и «ГигаЧат», позволяет дальше развивать противопоставление «естественности – искусственности» как на текстовом уровне, так и на лингвоперсонологическом. В связи с этим особого внимания заслуживают и существенные различия, имеющиеся между нейротекстами разных LLM. Эксперимент с написанием разными нейросетями сочинений по одной и той же картине показал, что разные нейросети генерируют разные тексты, подобно тому как разные языковые личности пишут разные персонотексты. Полученный вывод не противоречит сформулированному выше о сходстве текстов, сгенерированных разными нейросетями, а дополняет его: имеются и сходства, и различия. Следовательно, в рамках оппозиции «естественное – искусственное» можно говорить о нейроязыковой «личности», дрейфующей между полюсами: от полюса «искусственное», когда генеративная природа нейротекста очевидна, к полюсу «естественное», когда граница между персонотекстом и нейротекстом практически стирается.

Методы и материал исследования

В современной юридической лингвистике активно разрабатываются методики, позволяющие отличать текст человека от сгенерированного нейросетью. Классическое автороведение располагает комплексной методикой проведения экспертиз по текстам, содержащим не менее 100 словоформ, для «получения информации о половой принадлежности, возрасте, профессии, образовании, родном языке, национальности, психическом состоянии» человека [Моисеева, Огорелков 2022].

При этом постоянно ведется поиск новых методик, позволяющих на основе текстовых характеристик получить точную информацию об авторе. Для целей криминалистического автороведения А. В. Громовой обоснована «целесообразность развития междисциплинарного подхода к экспертному анализу таких текстов, подразумевающему использование методов компьютерной и классической лингвистики» [Громова 2025: 50].

Также активно развивается применение метода профилирования автора письменного текста, «т. е. установления характеристик автора по созданному им речевому произведению» [Литвинова 2012: 90]. С развитием интернет-технологий возрастает значимость точного профилирования личности по текстам электронной коммуникации, в связи с чем ведутся разработки методик с большим количеством параметров, «с помощью которых возможно портретирование «цифровой личности» с учетом разнообразного речевого поведения, обусловленного возрастом, гендерной принадлежностью, особенностями темперамента, уровнем психофизиологического и интеллектуального развития, текущим эмоциональным состоянием, наличием или отсутствием аддикции (наркотической, игровой и пр.), акцентуацией поведения в различных типах ситуативной обстановки, возможными видами патологий, местом формирования речевого навыка и т. д. в условиях опосредованной цифровой реальностью коммуникации» [Потапова, Курьянова 2022: 57].

В настоящем исследовании представлено применение метода сопоставительного профилирования авторов письменных текстов с целью представления критериев, позволяющих разграничить персонотекст и нейротекст. Мы исходим из положения о том, что порождение текста обусловлено языковой картиной мира автора (далее – ЯКМ), в результате чего выстраивается следующая цепочка: текст $\Leftrightarrow$ ЯКМ $\Leftrightarrow$ языковая личность. С нашей точки зрения, именно языковая картина мира в силу ее текстопорождающего потенциала [Голев, Шпильная 2009: 11] является ключевым понятием для дифференциации персонотекста от нейротекста. Если ЯКМ спорного текста соответствует ЯКМ типичной для данной языковой личности, то спорный текст является персонотекстом. Если же такого соответствия не наблюдается, то текст сгенерирован нейросетью. Конечно, нейросеть не обладает своей собственной

ЯКМ, но при генерации текстов в них обнаруживаются, воспроизводятся типовые и статистически распространенные ЯКМ.

На данном этапе исследования использованы следующие материалы:

1) персонотекст реальной языковой личности, написанный как пост при размещении фото в личном блоге, который автор ведет в стиле пародийной журналистики. Личный блог содержит более чем 10 поликодовых персонотекстов, объединенных тематикой межличностных школьных отношений;

2) нейротекст LLM «ГигаЧат», работающей на флагманской модели GigaChat Ultra, как реакция на промпт: «Посмотри на фото и сочини текст-репортаж».

Исследовательские материалы приводятся в статье в исходном виде без редактуры.

Персонотекст

«"Борьба за жизнь"»

Срочные новости, друзья!! 🔥

Стало известно, что случилось с бедным школьником, которого затолкала в сумку его собственная одноклассница!!

🤯 Если вы заинтригованы, читайте дальше.

Вы уже догадались, что происходит на фото?.. Это тот самый парнишка из сумки!! 😬 Кадры сделаны после жестокого

заточения одноклассника! Как он смог выбраться?? 🤔 Вы не поверите, когда узнаете, на что может пойти человек,

когда его жизни угрожает опасность! 😬 По словам потерпевшего, у него была с собой расческа(фото прилагается

🎞), с помощью которой он смог проделать отверстие и выкарабкаться наружу. Сообразительный парень заточил

эту расческу своими собственными зубами и только потом смог прорезать себе выход на волю! 🤩

Мы прекрасно знаем, что слева - это та самая обезумевшая после прочтения "Грозы" девушка. Но кто тогда справа на фото?! Это ее сообщница, дорогие мои 🤔 (хотя, возможно не все так просто... О подробностях вы узнаете в скором

времени 😊).

Видно, что обе настроены недоброжелательно: одна угрожает пострадавшему, который еле смог спастись, другая же надменно улыбается и чувствует свою безнаказанность 😡. Парень напуган! Не может сделать и шага! Они окружили

его!! Он в безвыходном положении! 😬 Видно, что сумка все еще в руках сумасшедшей. Вокруг множество людей, и никто не хочет с ней связываться 😭.

Неужели исход истории мальчишки будет плачевным? 🌹 Никто не сможет ему помочь?..

Люди, очнитесь! 🚨 Взгляните на жестокий мир вокруг! Мы не готовы прийти друг другу на помощь в страшных

жизненных ситуациях! 😬 И кто в этом виноват? 🤔 Родители и их воспитание? 😬 Равнодушие и бесчувствие со

стороны людей? 😬 Неправильно привитые обществом ценности? 🤔 Нет, дорогие мои, на самом деле все намного проще: во всем виноваты кобчики.

Автор и редактор статьи:

Лапаева В. В. 🔥 (🚨 Напоминаем, что мы не несем ответственности за возможный вред вашему здоровью, причиненный во время или после прочтения данной статьи 🚨)»

Нейротекст LLM «ГигаЧат»

Репортаж с места событий: «Школа выживания»

Сегодня мы оказались свидетелями необычной сцены в одном из учебных корпусов престижной гимназии города N. Здесь, среди привычных коридоров и классных комнат, разворачивалась настоящая драма человеческих эмоций и характеров.

■ Место действия: школьный холл

Время близилось к обеденному перерыву, ученики спешили отдохнуть и набраться сил перед следующим уроком. Но внимание всех присутствующих привлекла одна группа подростков, расположившаяся на ярко-жёлтом диване в центре холла.

■ Главные герои нашего сюжета

Анна, девушка слева, выглядела спокойной и уверенной. Она сидела прямо, держала сумку на коленях и внимательно наблюдала за происходящим вокруг. Её взгляд выражал лёгкую усмешку и любопытство, словно она была готова ко всему, что могло произойти дальше.

Рядом с Анной сидел Максим. Его поза говорила сама за себя: голова опущена, руки сцеплены в замок, плечи сгорбились. Казалось, весь мир обрушился на парня в этот момент. Максим казался подавленным и уставшим, будто переживал какой-то внутренний конфликт или неудачу.

Справа от Максима находилась Елена, девушка в светлой кофте и черных брюках. Она смотрела на Максима с сочувствием, слегка подняв руку вверх, словно хотела подбодрить друга жестом поддержки. Елена улыбнулась нам, показывая, что ситуация не столь трагична, как кажется на первый взгляд.

■ Что же произошло?

По словам очевидцев, Максим получил плохую оценку на последнем уроке математики. Это стало неожиданностью даже для самого парня, ведь обычно он считался одним из лучших учеников класса. Учительница поставила низкую отметку за контрольную работу, несмотря на старания Максима. Парень, привыкший к успехам, оказался морально не готов к такому повороту событий.

«Мы пытались его успокоить, — рассказала Анна, — но он замкнулся в себе. Мы знаем, что он расстроен, но уверены, что все наладится».

Елена добавила: «Это всего лишь оценка. Главное — поддержка друзей. Я уверена, что Максим справится и исправит ситуацию».

■ Атмосфера школы

Несмотря на драматизм момента, атмосфера в школе оставалась дружелюбной и позитивной. Другие ученики продолжали заниматься своими делами: кто-то читал учебники, кто-то разговаривал с друзьями. Жизнь шла своим чередом, хотя каждый понимал важность взаимопомощи и поддержки друг друга.

■ Итог

Эта сцена стала ярким примером того, насколько важны дружба и понимание в школьной среде. Даже самые сильные переживания можно преодолеть вместе, поддерживая друг друга. Ведь школа — это не только уроки и оценки, но и место, где формируются настоящие человеческие отношения.

Параметры, по которым проводился текстовый анализ:

1. Смысловые признаки.
2. Текстологические признаки.
3. Лексико-стилистические признаки. Количество слов и лексическое разнообразие.
4. Морфолого-синтаксические признаки. Ритм.
5. Коммуникативная тональность.
6. Графические, орфографические и пунктуационные признаки.
7. Психолингвистические признаки.

Параметры для проведения текстового анализа включают традиционные признаки, применяемые в автороведческих исследованиях, дополненные следующими характеристиками:

1. Количество слов и лексическое разнообразие, понимаемое как отношение количества уникальных слов текста к количеству всех слов текста. Параметры определяются с помощью ресурса «Текстометр» (<https://textometr.ru>).

Введение данных параметров в критерии анализа текста основывается на опыте многих лингвистов, отмечающих, что для сгенерированных текстов характерны «формально-языковые и содержательные повторы» [Осетрова, Седова 2025: 48].

2. Ритм текста, представленный как визуализация относительной длины предложений. Параметр позволяет на основе изменения длины предложений охарактеризовать текст как монотонный или динамичный. Он определяется с помощью «Сервиса проверки читабельности, ритма и параметров текста от Шардакова» (<https://text.shardakov.com>).

Введение параметра «ритм текста» обусловлено результатами исследований, в которых установлено, «что с точки зрения собственных лингвистических характеристик «естественные» тексты отличаются от сгенерированных преимущественно большей гибкостью синтаксической структуры, допускающей как пропуск или сокращение структур, так и избыточность, а также – большей лексической вариативностью» [Колмогорова, Марголина 2024: 73].

3. Коммуникативная тональность, понимаемая как «эмоционально-стилевой формат общения, возникающий в процессе взаимодействия коммуникантов и определяющий их меняющиеся установки и выбор всех средств общения» [Карасик 2008: 99].

Введение названного параметра обусловлено представлениями о нейросети как о «новом типе автора», имеющего постоянную коммуникативную тональность в отличие от реальной языковой личности, осуществляющей выбор тональности в зависимости от ситуации общения. Параметр «коммуникативная тональность» определяется в соответствии с классификацией В. И. Карасика.

4. Графические признаки текста, основанные на анализе использования эмодзи в текстах виртуальной коммуникации.

В лингвистике установлено, что эмодзи могут полностью, частично или совсем не соответствовать вербализованному тексту, а их использование «зависит от социальных, психологических черт, особенностей и установок лиц, вступающих в коммуникацию» [Дударева, Шпильная 2023: 99]. Это означает, что наличие/отсутствие эмодзи, их выбор позволяет профилировать автора по тексту.

Анализ персонотекста

1. Смысловые признаки.

Тематическая направленность текста – сфера межличностных школьных отношений.

2. Текстологические признаки.

Текст написан в жанре пародийного новостного репортажа. Нарратив открывается как бытовая сенсация («Срочные новости, друзья!!!»), затем трансформируется в социальную критику («Люди, очнитесь! Взгляните

на жестокий мир вокруг!») и завершается абсурдистским, иррациональным выводом («во всем виноваты кобчики»). Связь между композиционными блоками (новость – мораль – абсурд) осуществляется посредством эмоционально-ассоциативных переходов. Финальный твист о «кобчиках» представляет собой неожиданное изменение в повествовании о школьнике, побочную ассоциацию, не связанную с основной сюжетной линией.

3. Лексико-стилистические признаки. Лексическое разнообразие.

В тексте используются разнообразные стилистические пласты лексики: газетно-публицистическая («потерпевший», «безнаказанность», «жестокое заточение»), разговорная и просторечная лексика («парнишка», «выкарабкаться», «обезумевшая», «затолкала»). Сочетание стилей органично и служит формированию ироничного эффекта. Просторечия используются не как механическая замена нейтральных слов, а как способ создания неформальной коммуникации для «своих».

Коэффициент лексического разнообразия – 0,64. Это высокий показатель для текста из 278 слов, что демонстрирует богатый словарный запас автора, стремление создать качественный творческий продукт.

4. Морфолого-синтаксические признаки. Ритм.

В тексте часто употребляются вопросительные и побудительные предложения с глаголами в форме повелительного наклонения («очнитесь», «взгляните»), благодаря чему происходит вовлечение читателей в описываемую ситуацию. Наречия выражают пространственные и причинные значения («наружу», «вокруг», «своими», «только потом») и создают осязаемую картину происходящего. Характерно использование усилительных частиц («даже», «же», «всего лишь»), которые придают высказываниям модальность убеждения. Активно используются вводные слова и конструкции: они служат для создания интриги («по словам потерпевшего»), выражения неуверенности («хотя, возможно не все так просто...»), установления контакта с читателем («дорогие мои») и констатации («видно, что»).

В тексте чередуются длинные и короткие предложения, благодаря чему возникает динамичность ритма.

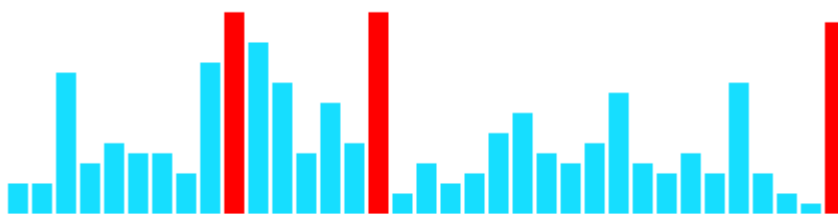


Рис. 1. Ритм персонотекста

5. Коммуникативная тональность.

Анализируемый текст имеет шутливую тональность – «несерьезное сниженное эмоциональное общение, целью которого является сокращение дистанции между коммуникантами» [Карасик 2008: 100] в сочетании с фатической. Показателем названных тональностей является выбранный жанр для создания текста (пародийный новостной репортаж) и реализация названного жанра в языковом оформлении текста. Помимо уже представленных особенностей текста на лексическом и морфологическом уровнях отметим его фоносемантические признаки, благодаря которым также создается шутливая тональность. Анализ ключевых слов («школьник», «сообщница», «расческа», «жестокий», «кобчики»; «одноклассница», «парень», «виноватый», «выкарабкаться», «безнаказанность») показывает частотность фонемы [o], ассоциирующейся с идеей замкнутого пространства (заточение героя, его безвыходное положение), и фонемы [a], ассоциативно связанной с открытым пространством (счастливым избавлением героя от опасности).

6. Графические, орфографические и пунктуационные признаки.

В тексте имеются орфографические ошибки, связанные с написанием Н и НН в причастиях, правописание безударной гласной в корне слова. Допущена пунктуационная ошибка, связанная с пропуском запятой после употребления вводного слова. Имеется слово с пропущенной буквой «прочтения» и слитное написание конструкции «очнитесьВзгляните».

Пунктуационной особенностью текста является эмоционально-экспрессивное использование восклицательных и вопросительных знаков препинания, проявляющееся в их дублировании. Многократные восклицания («друзья!!», «одноклассница!!») выступают средством интенсификации эмоций. Сочетание вопросительного и восклицательного знаков («Но кто тогда справа на фото?!») образует гибридный знак, передающий одновременно удивление и экспрессию. Многоточие («фото?...») используется для создания интригующей паузы и имитации устной речи.

В тексте имеется значительное количество разнообразных эмотиконов, которые усиливают шутливую эмоциональную тональность текста, соответствующую жанру пародийного новостного репортажа.

7. Психолингвистические признаки.

Для текста характерна эмоциональность речи, экзальтированность описания фантазийной ситуации. Проявлена не одна эмоция, а их сложный комплекс. Возбуждение и жалость к «бедному школьнику» сменяются гневом в адрес «обезумевшей» девушки и ее сообщницы, затем пафосным морализаторством, которое обесценивается абсурдным финалом.

Профилирование автора

На основе проведенного лингвистического анализа можно установить, что в тексте отражается языковая картина мира личности, имеющей следующие характеристики: пол – женский; возраст – молодой, с учетом школьной тематики, вероятно, автор – школьница старших классов; родной язык – русский.

Поскольку в тексте имеется множество внешне наблюдаемых деталей, обилие зрительных образов, можно сделать вывод, что по типу восприятия информации автор – визуал.

В тексте, несмотря на монологическую форму представления, заметно проявлена диалогичность (интенция привлечения внимания и вызова у читателей ярких эмоций, ритмизированность текста, экспрессия в использовании вопросительных и восклицательных конструкций), что дает основания сделать вывод об экстраверсии автора текста и, следовательно, о принадлежности к холерическому либо сангвиническому типам темперамента.

Анализ нейротекста

1. Смысловые признаки.

Тематическая направленность текста – сфера межличностных школьных отношений.

2. Текстологические признаки.

В заглавии обозначен жанр текста – «репортаж с места событий». Содержание текста в целом соответствует жанру, поскольку в нем имеется информация о месте («учебный корпус престижной гимназии города N») и времени описываемого события («время близилось к обеденному перерыву»), представлены портретные зарисовки героев (Анны, Елены и Максима), их прямая речь, присутствует автор-наблюдатель. Но поскольку репортаж как журналистский жанр должен выполнять воздействующую функцию, то от текста ожидается большая фактологичность, эмоциональность и динамичность. Завершает текст пафосное рассуждение о важности дружбы и понимания, напоминающее шаблонный вывод в школьном сочинении. Таким образом, наблюдается неполное соответствие заявленному жанру.

Текст четко разделен на содержательные блоки, обозначенные графическими символами (■), каждый из которых выполняет свою функцию в нарративе: обозначение жанра («Репортаж с места событий»), места действия («школьный холл»), представление героев («Главные герои нашего сюжета»), описание события («Что же произошло?»), обобщение («Атмосфера школы», «Итог»). Композиционная структура строго соответствует схеме «введение – основная часть – заключение». Связь между предложениями осуществляется преимущественно за счет формально-логических средств (причинно-следственные связи, временная последовательность) и лексических повторов.

3. Лексико-стилистические признаки. Лексическое разнообразие.

Лексика текста характеризуется клишированностью и шаблонностью. Автор использует устойчивые, предсказуемые обороты, характерные для школьных сочинений и газетных репортажей: «настоящая драма человеческих эмоций и характеров», «поза говорила сама за себя», «весь мир обрушился на парня», «морально не готов к такому повороту событий», «жизнь шла своим чередом».

Общее количество слов – 378, коэффициент лексического разнообразия – 0,66, но при этом лексика относится к одной тематической группе «школьная дружба», отсутствуют неожиданные и яркие лексические единицы, наблюдается повторяемость ключевых слов («друг», «дружба», «поддержка»). Отсутствуют просторечные формы, разговорная лексика, иронические обороты. Текст стилистически нейтральный.

4. Морфолого-синтаксические признаки. Ритм.

Текст составлен от лица этикетного «мы» («Сегодня мы оказались свидетелями необычной сцены...»), что характерно для научного или официально-делового стиля. Публицистический стиль предполагает авторскую проявленность, выражение собственного взгляда, оценок и суждений.

Глаголы преимущественно используются в действительном залоге, но их употребление не создает эффекта динамичности: «оказались», «разворачивалась», «спешили», «привлекла», «сидела», «наблюдала», «находилась», «смотрела», «улыбнулась». Глаголы в повелительном наклонении отсутствуют, что снижает апеллятивную функцию текста и возможность вовлечения читателя в коммуникативное пространство. Наречия, имеющие пространственные и причинные значения («среди», «в центре», «слева», «справа», «вверх», «вокруг»), выполняют сугубо информативную функцию без эмоциональной окраски картины происходящего. Усилительные частицы практически отсутствуют. Повторы сравнительных конструкций со словами «казалось», «словно», «будто» создают модальность предположения, неуверенности, осторожности высказывания мнения. Обращает на себя внимание насыщенность текста причастиями и деепричастиями, что увеличивает длину предложений и создает монотонность. Автоматический анализ ритмичности текста показывает отсутствие ярко выраженных контрастов между предложениями. Длина предложений уменьшается только в случаях обозначения рубрик.

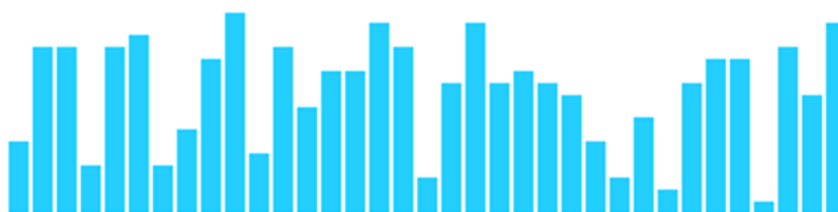


Рис. 2. Ритм нейротекста

5. Коммуникативная тональность

В тексте реализуется информативная тональность – «серьезное нейтрально-эмоциональное общение, целью которого является взаимное либо одностороннее информирование и реакция на него» [Карасик 2008: 100], которая проявляется в фиксации и перечислении действий героев.

6. Графические, орфографические и пунктуационные признаки.

Отсутствуют орфографические и пунктуационные ошибки, опечатки. Отсутствуют эмодзи. Из графических знаков используется вертикальная черта (|) для разделения текста на смысловые блоки.

7. Психолингвистические признаки.

Для анализируемого текста характерна незмоциональность речи в целом, неуверенность в обозначении чувств и переживаний героев, наблюдается ожидаемость эмоций, к которым относится сочувствие к герою, уверенность в позитивном исходе, общий оптимистический настрой. Оценки событий носят предсказуемый, морализаторский, этически выверенный характер: «важны дружба и понимание», «самые сильные переживания можно преодолеть вместе», «школа – это место, где формируются настоящие человеческие отношения».

Профилирование LLM «ГигаЧат» по нейротексту

На основе проведенного лингвистического анализа можно установить, что в тексте отражается языковая картина мира LLM «ГигаЧат» как нейроязыковой «личности», имеющей следующие характеристики:

- гендерная деперсонализация; текст лишен прямой демонстрации гендерной проявленности через обозначение имени автора, грамматические конструкции, в которых категория гендера была бы точно обозначена. Пол автора можно определять только по косвенным указаниям, что является недостаточно убедительным;

- возрастная деперсонализация; текст лишен явной информации об авторе, он не обозначен ни как одноклассник, ни как друг или родственник героев ситуации, ни как журналист, пришедший в школу, чтобы подготовить репортаж.

Родной язык – русский, так как высок процент лексического разнообразия, отсутствует употребление слов в несвойственных значениях, нарушение грамматики и т. п.

Текст наполнен зрительными образами, которые получают аналитическое осмысление. Считаем, что наличие зрительных образов в сгенерированном тексте следует из промпта, направленного на продуцирование текста на основе визуального восприятия. Демонстрации анализа от «ГигаЧата» в промпте не требовалось, но оно нашло проявление, что свидетельствует о признаках дигитала.

Информативная тональность, композиционная структурированность (введение – основная часть – заключение), осложненность лексико-грамматических конструкций вводными словами, причастными и деепричастными оборотами, демонстрирующими детальность и аналитизм наблюдения и создающими монотонность изложения, отсутствие яркой эмоциональности, хотя описывается «настоящая драма человеческих эмоций и характеров», позволяет сделать вывод об интроверсии нейроязыковой «личности» «ГигаЧат» и, следовательно, о вероятной принадлежности коммуникативного стиля к флегматическому типу темперамента.

Заключение

Сопоставление полученных результатов свидетельствует о том, что использование традиционных автороведческих признаков, дополненных новыми данными (с указанием количества слов и коэффициента лексического разнообразия; коммуникативной тональности; ритма; графических признаков), позволяет отличить персонотекст, продуцированный реальной языковой личностью, от нейротекста, сгенерированного нейроязыковой «личностью».

Текстовые и смысловые признаки демонстрируют, что для персонотекста характерны содержательные отступления, побочные ассоциации, типичные для спонтанной естественной речи, индивидуальность композиции текста. Для нейротекста типично смысловое единство и формальное соответствие заявленному жанру, если это указано в промпте. Нейротекст строится в соответствии с классической трехчастной композицией: введение, основная часть, заключение.

Лексико-стилистическими признаками нейротекста являются шаблонность, клишированность, использование предсказуемых оборотов, ориентация на книжную нейтральную лексику. Для персонотекста характерна свободная комбинация слов разных функциональных стилей, авторская проявленность (пол, возраст, образование и т. д.) через лексику.

По результатам проведенного эксперимента нейротекст по количеству слов получился большего объема, чем персонотекст, что коррелирует с параметрами обстоятельного изложения фактов, построением сложных синтаксических конструкций, типичных для классических письменных текстов, но не для текстов виртуальной коммуникации с их стремлением к легкости и живости естественной речи. Параметр «лексическое разнообразие» показал, что для текста, продуцированного реальной языковой личностью, характерно политематическое разнообразие лексических единиц, у нейротекста лексическое разнообразие монотематично и обусловлено промптом.

Важным диагностическим признаком нейротекста в рассмотренном примере является намеренное присвоение героям вымышленных имен: Анна, Максим, Елена. Этот прием демонстрирует алгоритмическую стратегию имитации живой речи через создание иллюзии конкретности и достоверности. Для человека как автора фантазийного репортажа имена героев принципиально не нужны. Отсутствие имен в используемом жанре позволяет осуществить перенос реальной бытовой ситуации, ставшей поводом для фото, на фантазийную занимательную историю. Нейросеть генерирует имена участников описываемой ситуации как маркеры индивидуализации, не принимая во

внимание, что они незнакомы широкому кругу читателей. Это явление можно охарактеризовать как алгоритмическую гиперкомпенсацию: нейросеть создает избыточную конкретику там, где человек как автор использует обобщение или функциональное описание.

Персонотекст демонстрирует способность реальной языковой личности к самоиронии и деконструкции собственного пафоса, что является уникальной чертой человеческого сознания, его рефлексивной природы. Для нейротекста характерны высокопафосные глобальные рассуждения, философские выводы, не подкрепленные реальной конкретикой.

Морфолого-синтаксические признаки персонотекста демонстрируют авторскую интенцию вовлечения в коммуникативное пространство текста, что проявляется в диалогичности и ритмичности текста, использовании эмоционального синтаксиса. Разнообразие модальностей реализует признак живого, рефлексивного автора, который не просто излагает факты, но и комментирует их, выражает свое отношение, сомневается и играет с читателем. Нейротекст отличается прямым порядком слов, полными грамматическими конструкциями, осложненными вводными словами, причастными и деепричастными оборотами, повторяемостью синтаксических конструкций, в результате чего текст становится монотонным.

С точки зрения коммуникативной тональности в нашем примере языковая личность использовала шутливую тональность с элементами фатической. Нейросеть «ГигаЧат» предпочитает информативную тональность, что соответствует ее проявлению как аналитической «личности».

По орфографическим и пунктуационным признакам LLM «ГигаЧат» проявляет нарочитую правильность. Отсутствуют даже опечатки и пропуски, характерные для естественной письменной речи. При генерации текста нейросеть не использовала эмодзи, что нетипично для языковой личности. Пунктуационные знаки нейротекста выполняют смысловую функцию, их экспрессивно-эмоциональная нагрузка минимальна.

Сопоставление по психолингвистическим признакам обнаруживает, что эмоциональные проявления нейроязыковой «личности» сводятся к читательски ожидаемым. При описании ситуации, потенциально вызывающей эмпатию (школьник расстроен из-за плохой оценки), эмоциональная реакция «ГигаЧата» остается ровной, без признаков глубокого сопереживания, что характерно для алгоритма, следующего заданной моральной программе, а не человеческому спонтанному выражению эмоционального интеллекта. В персонотексте авторские эмоции проявляются ярко, многократно. Это заметно и при содержательном анализе текста, и при анализе его графического оформления.

Наблюдаемые текстовые различия позволили охарактеризовать LLM «ГигаЧат» как нейроязыковую «личность» с деперсонализированными гендерными и возрастными характеристиками, русскоязычную, имеющую цифровой тип восприятия (с возможностью настройки на другой тип в зависимости от поставленной в промпте задачи) и обладающую коммуникативным стилем флегматика.

Литература

- Голев Н. Д., Мельникова В. С., Напреенко Г. В. Русский юридический текст в парадигме «новой вариантологии» / Сибирский филологический журнал. – 2025. – № 4. – С. 201–215. <https://doi.org/10.17223/18137083/93/15>.
- Голев Н. Д., Шпильная Н. Н. Языковая картина мира как текstopорождающая категория (на материале сочинений по картине «Девочка с персиками» В. А. Серова) / Вестник Томского государственного университета. – 2009. – № 322. – С. 11–16.
- Громова А. В. Признаки сгенерированных нейросетями текстов в аспекте криминалистического автороведения (на основе экспериментального исследования) / Теория и практика судебной экспертизы. – 2025. – Т. 20. – № 4. – С. 50–58. <https://doi.org/10.30764/1819-2785-2025-4-50-58>.
- Дударева Я. А., Шпильная Н. Н. Диалог как текст, содержащий эмодзи: юрислингвистический аспект рассмотрения / Юрислингвистика. – 2023. – 27. – С. 99–105. [https://doi.org/10.14258/leglin\(2023\)2718](https://doi.org/10.14258/leglin(2023)2718).
- Карасик В. И. Коммуникативная тональность / Актуальные проблемы филологии и педагогической лингвистики. – 2008. – № 10. – С. 99–109.
- Клушина Н. И. Автор, языковая личность и искусственный интеллект / Вестник Московского университета. – Серия 9. Филология. – 2025. – № 2. – С. 20–31. <https://doi.org/10.55959/MSU0130-0075-9-2025-48-02-2>.
- Колмогорова А. В., Марголина А. В. Написанный vs сгенерированный текст: «естественность» как категория текстовая и психолингвистическая / 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>.
- Литвинова Т. А. Установление характеристик (профилирование) автора письменного текста / Филологические науки. Вопросы теории и практики. – 2012. – № 2 (13). – С. 90–94.
- Лягошина Т. В. Большие языковые модели: влияние на публичный дискурс и общество в целом / Вестник Томского государственного университета. Философия. Социология. Политология. – 2024. – № 79. – С. 111–124. <https://doi.org/10.17223/1998863X/79/11>.
- Микалеф Л. О. Лингвистика нейросетей как парадигма современной науки о языке / Мир науки, культуры, образования. – 2025. – № 1. – С. 467–469. <https://doi.org/10.24412/1991-5497-2025-1110-467-469>.
- Моисеева Т. Ф., Огорелков И. В. Судебная автороведческая экспертиза: учебное пособие. М., 2022. URL: <https://op.raj.ru/vpo/1027-moiseeva-t-f-ogorelkov-i-v-sudebnaya-avtorovedcheskaya-ekspertiza-uchebnoe-posobie?ysclid=mnb91139o7427605843>

Осетрова Е. В., Седова А. В. Характеристики сгенерированного текста: языковой и социально-коммуникативный анализ / Сибирский филологический форум. – 2025. – №2 (31). – С. 45–55. <https://doi.org/10.24412/2587-7844-2025-2-45-55>.

Потапова Р. К., Курьянова И. В. Научно-исследовательская основа лингвокриминалистики и особенности портретирования «цифровой личности» / Вестник Московского государственного лингвистического университета. Гуманитарные науки. – 2022. – Вып. 13 (868). – С. 56–61. https://doi.org/10.52070/2542-2197_2022_13_868_56.

References

- Golev, N. D., Melnikova, V. S., Napreenko, G. V. (2025). Russian legal text in the paradigm of “new variantology”. *Siberian Philological Journal*, 4, 201-215. DOI: 10.17223/18137083/93/15 (in Russian).
- Golev, N. D., Shpilnaya, N. N. (2009). The linguistic picture of the world as a text-generating category (based on essays on the painting “Girl with Peaches” by V.A. Serov). *Bulletin of Tomsk State University*, 322, 11-16 (in Russian).
- Gromova, A. V. Features of texts generated by neural networks in the aspect of forensic authorship (based on an experimental study) (2025). *Theory and practice of forensic examination*, 20, 4, 50-58. DOI: 10.30764/1819-2785-2025-4-50-58 (in Russian).
- Dudareva, Ya. A., Shpilnaya, N. N. (2023). Dialogue as a text containing emoticons: a legal linguistic aspect of consideration. *Jurilinguistics*, 27, 99-105. DOI: 10.14258/leglin(2023)2718 (in Russian).
- Karasik, V. I. (2008). Communicative tonality. *Actual problems of philology and pedagogical linguistics*, 10, 99-109 (in Russian).
- Klushina, N. I. (2025). Author, linguistic personality and artificial intelligence. *Bulletin of Moscow University*, 9, Philology, 2, 20-31. DOI: 10.55959/MSU0130-0075-9-2025-48-02-2 (in Russian).
- Kolmogorova, A. V., Margolina, A. V. (2024). Written vs. generated text: “naturalness” as a textual and psycholinguistic category. *Research result. Issues of theoretical and applied linguistics*, 10, 2, 71-99. DOI: 10.18413/2313-8912-2024-10-2-0-4 (in English).
- Litvinova, T. A. (2012). Establishing characteristics (profiling) of the author of a written text. *Philological sciences. Theoretical and practical issues*, 2 (13), 90-94 (in Russian).
- Lyagoshina, T. V. (2024). Large linguistic models: influence on public discourse and society as a whole. *Bulletin of Tomsk State University. Philosophy. Sociology. Political Science*, 79, 111-124. DOI: 10.17223/1998863X/79/11 (in Russian).
- Mikaleff, L. O. (2025). Linguistics of neural networks as a paradigm of modern language science. *World of science, culture, education*, 1, 467-469. DOI: 10.24412/1991-5497-2025-1110-467-469 (in Russian).
- Moiseeva, T. F., Ogorelkov, I. V. (2022). Forensic authorship examination: a tutorial. Available from: <https://op.raj.ru/vpo/1027-moiseeva-t-f-ogorelkov-i-v-sudebnaya-avtorovedcheskaya-ekspertiza-uchebnoe-posobie?ysclid=mnbn91139o7427605843> (in Russian).
- Osetrova, E. V., Sedova, A. V. (2025). Characteristics of the generated text: linguistic and socio-communicative analysis. *Siberian Philological Forum*, 2 (31), 45-55. DOI: 10.24412/2587-7844-2025-2-45-55 (in Russian).
- Potapova, R. K., Kuryanova, I. V. (2022). Research basis of forensic linguistics and features of portraying a “digital personality”. *Bulletin of the Moscow State Linguistic University. Humanities*, 13 (868), 56-61. DOI: 10.52070/2542-2197_2022_13_868_56 (in Russian).

Citation:

Дударева Я. А., Воронько А. А. Персонотекст и нейротекст: критерии разграничения естественно-языковых и сгенерированных текстов на основе сопоставительного профилирования // Юрислингвистика. – 2026 – 41. – С. 50-59.

Dudareva Ya. A., Voronko A. A. (2026) Personatext and Neurotext: Criteria for Distinguishing between Natural Language and AI-Generated Texts Based on Comparative Profiling. *Legal Linguistics*, 41, 50-59.



This work is licensed under a Creative Commons Attribution 4.0. License