

МИНОБРНАУКИ РОССИИ
Федеральное государственное бюджетное образовательное
учреждение высшего образования
«САРАТОВСКИЙ НАЦИОНАЛЬНЫЙ ИССЛЕДОВАТЕЛЬСКИЙ
ГОСУДАРСТВЕННЫЙ УНИВЕРСИТЕТ ИМЕНИ
Н.Г. ЧЕРНЫШЕВСКОГО»

Кафедра теории, истории языка и прикладной лингвистики

**Лингвистические особенности художественных текстов,
сгенерированных ИИ**

АВТОРЕФЕРАТ ВЫПУСКНОЙ КВАЛИФИКАЦИОННОЙ РАБОТЫ
БАКАЛАВРА

Студентки 4 курса 441 группы
направления 45.03.03 Фундаментальная и прикладная лингвистика
Института филологии и журналистики
Алексеевой Ксении Олеговны

Научный руководитель

Канд. филол. наук, доцент

подпись, дата

Е.В. Старостина

Зав.кафедрой

докт. филол. наук, профессор

подпись, дата

О.Ю.Крючкова

Саратов 2025

Искусственный интеллект (далее ИИ) произвёл революцию во многих сферах человеческой жизни: он пишет оригинальные тексты, программы, переводит с иностранных языков, решает математические задачи различной сложности и даже управляет автомобилями. Также он играет решающую роль в современном мире, обеспечивая автоматизацию, улучшая процесс принятия решений, повышая эффективность и производительность, создавая новые возможности для инноваций и роста в различных отраслях, включая здравоохранение, финансы, производство, электронную коммерцию, образование и многие другие. Научно доказано, что эффективность и скорость работы машины гораздо выше человеческой. С развитием ИИ возникает необходимость взаимодействия между техникой и людьми. Одним из важных его аспектов является вопрос о наличии у ИИ так называемой «языковой личности» — способности машины использовать различные языки и стили коммуникации и адаптироваться к ним [Ариськина 2011:15].

Одним из наиболее захватывающих направлений работы с ИИ является генерация текстов с его помощью, что открывает новые горизонты для исследования лингвистических особенностей различных текстов, в том числе художественных. Современные модели, такие как GPT (Generative Pre-trained Transformer), способны создавать тексты, которые по своей структуре и содержанию могут соперничать с произведениями, написанными человеком. В связи с этим возникает логичный вопрос: может ли ИИ писать оригинальные тексты, придерживаясь жанровых особенностей определенного стиля? Насколько эти тексты будут схожи с теми, что создает человеческий мозг?

Целью данной работы является изучение особенностей художественных текстов, сгенерированных ИИ, и выявление их отличий от текстов, написанных человеком.

Для достижения этой цели необходимо решить несколько **задач**:

1. исследовать теоретические основы проблемы генерации текстов;
2. выявить и проанализировать стилистические особенности и характерные черты текстов, сгенерированных ИИ, и художественных текстов;
3. создать базу текстов для анализа, включающую художественные тексты, созданные человеком и сгенерированные ИИ;
4. описать лексические и синтаксические характеристики отобранных текстов (сгенерированных ИИ и написанных человеком), сравнить данные характеристики;
5. провести эксперимент, направленный на выявление особенностей художественного текста и определение восприимчивости людей к искусственно сгенерированному художественному тексту.
6. провести сравнительный анализ полученных результатов и сделать выводы.

Объектом исследования являются художественные тексты, написанные человеком, и тексты художественного стиля, созданные с помощью современных алгоритмов генерации.

Предмет исследования — лингвистические особенности художественных текстов, сгенерированных ИИ.

Методы исследования включают в себя как качественный, так и количественный анализ текстов, что позволит получить более полное представление о природе сгенерированных текстов, а также эксперимент и анкетирование.

Материалом исследования послужили тексты, сгенерированные искусственным интеллектом (Яндекс GPT 5 Pro, DeepSeek и OpenAI), и тексты, написанные известными российскими авторами, получившими награды различного уровня. Так, в наше исследование вошли отрывки из романов Н. Абгарян «С неба упали три яблока», «Молчание цвета», Е.

Водолазкина «Авиатор», М. Степновой «Женщины Лазаря», «Безбожный переулок».

Актуальность данной темы обусловлена не только стремительным развитием технологий, но и необходимостью понимания того, как ИИ интерпретирует и воспроизводит языковые структуры. Вопросы о том, насколько естественными и выразительными могут быть сгенерированные тексты, а также о том, как они соотносятся с традиционными формами художественной речи, становятся все более важными в контексте лингвистических исследований.

Структура работы: работа состоит из введения, заключения, двух глав, списка литературы и двух приложений. В Приложении 1 представлен текст опроса с вариантами ответов. В Приложении 2 – результаты данного опроса.

Основное содержание работы. Первая глава «Отличительные черты текстов, созданных ИИ и человеком» содержит теоретическую информацию по теме исследования. Глава состоит из 3 разделов. Раздел 1.1 «Основные принципы генерации текстов с помощью ИИ» содержит три параграфа. В параграфе 1.1.1 «Понятие «нейронные сети». **Обучение нейронных сетей. Генеративные нейронные сети** рассматривается сущность нейронных сетей как математических моделей, имитирующих работу человеческого мозга, и их ключевая роль в развитии искусственного интеллекта (ИИ). Освещаются исторические предпосылки возникновения ИИ, включая работы Алана Тьюринга и Джона Сёрля.

Особое внимание уделено механизмам обучения нейронных сетей, включая настройку параметров на основе тренировочных данных, минимизацию ошибок и тестирование производительности. Также анализируются генеративные нейронные сети, способные создавать тексты, изображения и музыку, но ограниченные в генерации принципиально нового контента.

Отмечается эволюция нейросетевых технологий: от выдачи множества вариантов ответов к формированию более точных и информативных реакций благодаря структурированным Dataset. Подчеркивается, что современные системы ИИ остаются инструментами обработки данных, не обладая способностью к самостоятельному мышлению.

Первые попытки классификации языков относят к 16 веку. Именно тогда люди стали замечать определенное сходство языков. Современное языкознание знает два подхода классификации: группировку языков по общности строя и типа

В параграфе **1.1.2 «GPT-4 как передовая генеративная нейросеть: возможности и ограничения»** описывается GPT-4, представляющий собой наиболее совершенную на сегодняшний день языковую модель, демонстрирующую качественный скачок в развитии нейросетевых технологий. По сравнению с предыдущей версией GPT-3.5, модель обладает расширенным контекстом (до 25 тыс. слов), повышенной надежностью и мультимодальностью, включая анализ изображений. Как показывают исследования, GPT-4 успешно проходит профессиональные тесты, однако сохраняет склонность к контекстуальным ошибкам.

Ключевой особенностью модели является способность к трансферному обучению — решению задач, не предусмотренных явным обучением. При этом нейросеть демонстрирует существенные ограничения в понимании глубинного смысла и контекста, оставаясь неспособной к подлинно креативному мышлению. Преимущества GPT-4 заключаются в исключительной скорости обработки информации и генерации ответов, что в разы превышает человеческие возможности.

В параграфе **1.1.3 «Семантический и латентно-семантический анализ в обработке естественного языка»** описан латентно-семантический анализ (ЛСА) – важнейший метод в области обработки

естественного языка (NLP), позволяющий выявлять скрытые смысловые связи в текстовых данных. Как отмечают исследователи, данный метод основан на принципах факторного анализа и нашел широкое применение в задачах классификации и кластеризации документов.

Процесс ЛСА включает несколько этапов: сначала большие массивы текстов подвергаются статистической обработке, затем извлекаются контекстно-зависимые значения слов, которые представляются в виде векторных моделей. Эти векторы, являющиеся форматом представления информации в нейросетях, позволяют эффективно анализировать смысловое содержание текстов любого объема.

Такой подход обеспечивает выявление латентных (скрытых) семантических связей, что делает ЛСА мощным инструментом для автоматической обработки и анализа текстовой информации в различных приложениях искусственного интеллекта.

Раздел 2 «Лингвистические особенности текстов разного типа» состоит из двух параграфов. В параграфе **1.2.1 «Лингвистические особенности текстов, сгенерированных ИИ»** описаны черты, по которым можно определить сгенерированный текст.

Современные системы искусственного интеллекта достигли значительных успехов в генерации текстового контента, однако создаваемые ими тексты обладают рядом отличительных характеристик. Прежде всего, такие тексты демонстрируют формальную правильность и соблюдение грамматических норм, но при этом сохраняют определённую шаблонность и предсказуемость структуры.

Важным аспектом является прагматическая составляющая таких текстов. Они часто лишены подлинного авторского стиля, демонстрируют зависимость от исходных данных обучения и испытывают сложности с созданием по-настоящему убедительных аргументов. Эти особенности

необходимо учитывать при использовании ИИ для решения различных текстовых задач.

Далее в параграфе **1.2.2 «Лингвистические особенности художественных текстов»** описаны и проанализированы основные черты художественного текста. Художественный текст представляет собой сложную многоуровневую систему, обладающую уникальными лингвистическими характеристиками, которые отличают его от других функциональных стилей. Прежде всего, он выполняет эстетическую функцию, создавая особую художественную реальность через систему образов и языковых средств. Ключевой особенностью является образность и многозначность, проявляющаяся в активном использовании тропов – метафор, эпитетов, сравнений. Эти средства создают сложные ассоциативные связи и полифоничность смысловых уровней, обогащая текст дополнительными значениями.

Художественный текст создает особую коммуникативную ситуацию, где важную роль играет не только передача информации, но и формирование эстетического переживания. Это реализуется через сложное взаимодействие всех языковых уровней – от лексического до синтаксического, – подчиненное общей художественной задаче.

Раздел 1.3 включает основные выводы по теоретической главе.

Во **второй главе «Анализ художественных текстов, сгенерированных ИИ и созданных человеком»** описан проведенный нами эксперимент, в рамках которого требовалось оценить, насколько качественно на данном этапе разработки системы искусственного интеллекта способны создать текст в художественном стиле. Для этого был создан специальный опрос, состоящий из 5 наборов по 4 текста, в каждом из которых 3 текста были сгенерированы с помощью инструментов ИИ, а один взят из произведений современных русских

писателей Н. Абгарян, Е. Водолазкина и М. Степновой. Для генерации текстов мы использовали Яндекс GPT 5 Pro, DeepSeek и OpenAI. Также испытуемым были предоставлены варианты ответов «Все тексты сгенерированы ИИ» и «Все тексты написаны человеком». После каждого набора текста было поле с открытыми ответами, куда респонденты должны были вписать основание, почему был выбран конкретный вариант ответа.

Всего в эксперименте приняли участие 75 человек, среди них 64 женщины (85,3%) и 11 мужчин (14,7%). Больше всего реципиентов в возрасте 18-25 лет (57 человек, 76%), 25-35 лет – 4 человека (5,3%), 35-45 лет – 6 человек (8%), в диапазонах 45-55, 55-65, старше 65 по 1 человеку (по 1,3%), а также 5 участников были до 18 лет (6,7%).

Стоит отметить, что все сгенерированные тексты включали в себя однотипные черты художественных текстов: часто использовались тропы (эпитеты, метафоры, олицетворения, и т.д.), однако они были не так образны, как аналогичные тропы в текстах, написанных людьми, все тексты отличаются также высокой эстетической образностью. Все сгенерированные тексты, однако, были обезличены, в них редко можно было найти авторское начало или особую эмоциональность, вероятно, машина не может писать о том, чего не чувствует, как человек.

Можно также отметить такую интересную черту всех текстов, написанных ИИ, как отсутствие привязки ко времени и месту. Во всех текстах ИИ, проанализированных нами, отсутствуют конкретные имена или фамилии героев, упоминания конкретных географических мест или временных периодов. Именно поэтому упоминания имен становятся для испытуемых признаком того, что текст написан человеком.

Кроме того, ИИ никогда не выходит за пределы литературного языка. В созданных им текстах нет ни разговорных элементов, ни

элементов каких-то других стилей, тем более элементов других социальных вариантов (жаргона, просторечия, сленга).

В текстах ИИ нет никаких отклонений от нормы, несообразностей, они всегда очень «гладкие», там нет окказионализмов, нет сочетания разностилевых элементов, языковой игры, юмора, иронии. Все это встречается только в человеческих текстах и по этой причине также становится для испытуемых маркером авторства текста.

В заключении подводятся итоги исследования. Проведенное исследование было направлено на изучение лингвистических особенностей художественных текстов, сгенерированных искусственным интеллектом, и их сравнение с текстами, созданными человеком. В ходе работы были успешно решены поставленные задачи, что позволило достичь основной цели — выявить ключевые различия между этими типами текстов и определить, насколько ИИ способен имитировать художественный стиль.

Теоретический анализ подтвердил, что современные нейросетевые модели, такие как GPT-4, способны генерировать связные и логичные тексты, но их работа основана на вероятностном предсказании, а не на осмысленном творчестве. ИИ демонстрирует высокую скорость обработки данных и способность копировать стилистические особенности, но не обладает креативным мышлением, эмоциональным опытом или глубоким пониманием контекста.

Лингвистические особенности текстов ИИ включают клишированные синтаксические структуры, использование пассивного залога, меньшее лексическое разнообразие и ограниченную вариативность длины предложений. Тексты ИИ часто лишены образности, избегают сложных метафор и идиом, что делает их менее выразительными по сравнению с человеческими. Художественные тексты, созданные человеком, характеризуются эстетической функцией, образностью,

субъективной модальностью и многоплановостью смыслов. Они активно используют тропы, стилистические фигуры и авторский стиль, что позволяет создавать уникальные художественные миры.

Экспериментальная часть показала, что, несмотря на значительные успехи ИИ в генерации текстов, респонденты чаще правильно идентифицировали человеческие тексты по таким критериям, как необычные метафоры, эмоциональность, авторский слог и конкретные детали. Однако некоторые тексты, созданные ИИ (особенно с помощью DeepSeek и OpenAI), были восприняты как человеческие, что свидетельствует о высоком уровне их имитации.

Ограничения ИИ связаны с отсутствием личного опыта, эмоций и способности к глубокой интерпретации контекста. ИИ может успешно имитировать формальные признаки художественного стиля, но не способен создавать принципиально новые идеи или передавать подлинные эмоциональные переживания.

Перспективы дальнейших исследований могут быть связаны с изучением возможностей ИИ в имитации различных жанров и стилей, разработкой методов оценки «естественности» сгенерированных текстов, а также углубленным анализом взаимодействия между человеческим творчеством и искусственным интеллектом в литературном процессе.

Таким образом, несмотря на впечатляющие достижения в области генерации текстов, искусственный интеллект пока не может полностью заменить человеческое творчество, оставаясь мощным, но ограниченным инструментом в руках авторов. Результаты исследования подчеркивают уникальность художественного текста как продукта человеческого мышления и эмоций, что открывает новые горизонты для дискуссий о природе творчества в эпоху цифровых технологий.