

Сгенерированный медиатекст в аспекте языковой грамотности

Научный руководитель – Коняева Юлия Михайловна

Мирошников Натан Алексеевич

Студент (магистр)

Санкт-Петербургский государственный университет, Институт "Высшая школа журналистики и массовых коммуникаций Кафедра международной журналистики, Санкт-Петербург, Россия
E-mail: Natheview@vk.com

Исследование направлено на выявление особенностей сгенерированных медиатекстов в соответствии с нормами русского литературного языка. Несмотря на активное внедрение генеративных моделей в профессиональную деятельность и медиапроизводство, уровень грамотности данных текстов рассматривается фрагментарно – как часть общей нормативной теории или в аспекте механизированных сбоев генеративных моделей [Федотова 2025; Осетрова 2025]. Научная проблема, лежащая в основе нашего исследования, заключается в отсутствии целостного представления о границах норм, заложенных в генеративных моделях.

Эмпирическую базу исследования составляют сгенерированные тексты, опубликованные в социальных сетях конкретного медиасубъекта – студии дизайна интерьера Start Design. Хронологические рамки и выборка ограничены временем введения генеративной модели в деятельность компании. В качестве методов используются типологизация и количественный подсчет выявленных нарушений.

В процессе анализа методом сплошной выборки нами был сформирован корпус из 50 медиатекстов компании, которые далее были проанализированы на предмет наличия или отсутствия в них нарушений норм орфографии и пунктуации, а также технических недочетов. Нарушения были зафиксированы в 35 из 50 текстов, что составляет 70% выборки. Во многих случаях один текст содержал ошибки разных типов.

Количественный анализ демонстрирует частотность в сгенерированных текстах ошибок технического характера, ведущих за собой нарушения постановки различных знаков препинания (точка, тире, двоеточие) и нарушения норм правописания строчных и прописных букв.

Наиболее частотными оказались технические погрешности, связанные с наличием **лишних пробелов**: выявлено 60 случаев в 19 текстах. Этот тип ошибки в сгенерированных текстах становится систематическим, особенно в начале и конце абзаца («*Происхождение:* », «*площадью 70 кв.м.,* », «*уединением.* »). При этом отметим, что данная ошибка становится наиболее легко выявляемым маркером неадаптированного текста.

Ошибки в **правописании прописных и строчных букв** зафиксированы в 34 примерах (8 текстов). Типичный случай – употребление заглавной буквы после двоеточия внутри предложения («*Плюсы: Роскошный*», «*Ошибка 1: Экономия*», «*Под кроватью: Да-да!*»). Этот тип ошибки также легко выявляется при чтении, становясь значимым маркером неотредактированного текста.

Ошибки в **постановке двоеточия и тире** выявлены в 18 случаях (5 текстов). Зафиксированные нарушения можно разделить на два типа: во-первых, немотивированное использование двоеточия в заголовках и подзаголовках, где оно не требуется синтаксической конструкцией («*Создаем уют будущего: каждый уголок этого дома – шаг к вашей мечте!*», «*Палитра:*»), во-вторых, избыточное употребление тире на месте других знаков

препинания или при отсутствии необходимых синтаксических условий («Мы рады представить вам проект — интерьер...»). Наличие подобных погрешностей может также рассматриваться как маркер неотредактированного сгенерированного текста, однако для окончательных выводов требуется более детальный анализ контекста, поскольку в ряде случаев интонационное тире или пояснительное двоеточие могут быть нормативными.

Ошибки в **постановке точки** зафиксированы в 5 примерах (5 текстов). Нарушения касаются двух правил: постановка точки после единиц измерения («площадью 160 м².», «площадью 70 кв.м.») и точка в заголовках или подзаголовках («Велосипед: не в прихожей, а в нише.»). Данный тип ошибки в анализируемом материале носит единичный характер и требует дальнейшего наблюдения.

Таким образом, исследование выявило устойчивые типы ошибок в сгенерированных медиатекстах, затрагивающие различные уровни оформления письменной речи. Полученные данные свидетельствуют о важности постредакционной обработки текстов, созданных искусственным интеллектом. Перспективы дальнейшего исследования связаны с привлечением для анализа фактического, логического и композиционного уровней организации текста.

Источники и литература

- 1) Осетрова, Е. В. Характеристики сгенерированного текста: языковой и социально-коммуникативный анализ // Сибирский филологический форум. — 2025. — № 2. — С. 45-55.
- 2) Федотова, А. М. Методика идентификации текстов, сгенерированных большими языковыми моделями // Информатика и автоматизация. — 2025. — № 5. — С. 1444-1470.