

УДК 372.881.1+004.8

Генеративный искусственный интеллект как основа повышения автоматизации подготовки учебных материалов с заданными параметрами по иностранному языку

Артеменко Ольга Александровна^(*)

artemenko.aa@bmstu.ru

Туананова Надежда Александровна

tunanovan@bmstu.ru

КФ МГТУ им. Н.Э. Баумана, Калуга, Россия

Аннотация. Рассмотрена способность генеративного искусственного интеллекта создавать фрагменты текста или учебные упражнения с заданными грамматическими явлениями и отобранными для запоминания лексическими единицами, и грамматическими конструкциями английского языка. Сравниваются языковой корпус Now и телеграмм чатбот ChatGPT4 Нейросети. Представлено, что на сегодняшний день лингвистические корпуса значительно уступают искусственному интеллекту по результативности выполнения запросов пользователя. Сделаны выводы, что генеративный искусственный интеллект имеет потенциал успешного применения в образовательных учреждениях высшего образования для формирования иноязычной коммуникативной компетенции, но может потребоваться его дальнейшее обучение или тонкая настройка для достижения высоких показателей точности.

Ключевые слова: генеративный искусственный интеллект, большие языковые модели, языковой корпус, искусственный интеллект в образовании, иноязычная коммуникативная компетенция

Generative artificial intelligence as a basis for increasing the automation of the preparation of educational materials with specified parameters in a foreign language

Artemenko Olga Aleksandrovna^(*)

artemenko.aa@bmstu.ru

Tunanova Nadezhda Aleksandrovna

tunanovan@bmstu.ru

BMSTU (KB), Kaluga, Russia

Abstract. The ability of generative artificial intelligence to create text fragments or learning tasks with given grammatical phenomena and selected lexical units and constructions of the English language has been considered. The language corpus Now and the telegram chatbot ChatGPT4 Neural networks have been compared. It has been shown that today linguistic corpora are significantly inferior to artificial intelligence in terms of the effectiveness of executing user queries. It has been concluded that generative artificial intelligence has the potential for successful application in higher education institutions for the formation of foreign language communicative competence, but further training or fine-tuning may be required to achieve high accuracy rates.

Keywords: generative artificial intelligence, large language models, language corpus, artificial intelligence in education, foreign language communicative competence

Введение. На современном этапе развития информационных технологий искусственный интеллект находит все более широкое применение в рамках организации процесса обучения в высшей школе. По мнению российских и зарубежных исследователей в области педагогики (Ворожихин В.В., Константинова Л.В., Петров А.М., Титова Е.С., Штыхно Д.А. и т. д.), а также методики преподавания иностранных языков (Абалян Ж.А., Пивнева С.В., Авраменко А.П., Сысоев П.В., Фадеева В.А., Терновский В.В. и т. д.), несмотря на ряд объективных ограничений генеративный искусственный интеллект постепенно становится одним из инструментов, позволяющих кардинально трансформировать и автоматизировать процесс разработки учебных материалов, а также существенно усовершенствовать методическую организацию аудиторных занятий по дисциплине «Иностранный язык». Развитие программных приложений и телеграмм ботов на базе искусственного интеллекта может лежать в основе повышения учебной автономии студентов.

Материалы и методы. Генеративный искусственный интеллект (ГИИ) определяется как технология, которая использует модели глубокого обучения для создания оригинальных информационных материалов в ответ на запрос человека и рассматривают как новый, значительно более совершенный интеллектуальный цифровой инструмент для работы с информацией [1]. Таким образом, ГИИ позволяет генерировать новые тексты, изображения, видео- и аудиоконтент, программные приложения и т. д. на основе изучения закономерностей, используя уже существующие массивы данных. В основе обучения больших языковых моделей, тип генеративного ИИ, ориентированный на обработку, понимание и создание данных естественных языков, лежат огромные объемы текстовых данных из сети Интернет. Продвинутое генеративные модели, такие как GPT (от OpenAI) или BERT (от Google), используют глубокое обучение для создания более разнообразного и похожего на человеческий текста посредством обработки естественного языка [2]. Традиционно источником текстовых материалов выступают периодические издания или научно-техническая литература, в том числе в электронном формате, доступные в сети Интернет. Для упрощения поиска используются специализированные корпуса текстов, при этом технологии искусственного интеллекта позволили значительно расширить лингводидактический потенциал уже существующих корпусных технологий, наделив их новыми многочисленными способностями [3].

Необходимо отметить, что корпусные технологии дают возможность производить поиск только отдельных слов или словосочетаний, тогда как ГИИ способны не только делать выборку с несколькими единицами одновременно, но и с требуемыми грамматическими явлениями в предложениях. К рискам использования ГИИ относится вероятность получения некорректной, неполной или недостаточной информации по запросу. Во-первых, это

связано с тем, что искусственный интеллект — это само-обучающаяся система, и по мере того, как в систему попадает недостоверная информация, чат-бот все чаще может генерировать некорректные ответы [4], во-вторых, ГИИ в некоторых случаях неправильно или не полностью «понимает» запрос пользователя.

В силу того что обучение генеративного ИИ основано не на фиксированных правилах, а на фактическом использовании языка, большие языковые модели не имеют реального понимания того, что запросы означают. Генерируемые тексты в некотором смысле представляют собой просто «переупакованные» данные, которые были ранее введены в систему, что достигается посредством сложного математического процесса анализа и реконструкции. Отсутствие истинного понимания языка и его значения может порой приводить к созданию неточных утверждений и неподходящих текстов [5].

Применение корпусных технологий или генеративного ИИ в процессе изучения иностранных языком исследователи разделяют на две категории. К первой относится применение материалов корпусов для создания различных учебных материалов и пособий, а ко второй категории относят непосредственное использование корпусных технологий для исследования языковых явлений, когда обучающиеся сами выступают автономными исследователями языка [6]. Несмотря на перспективность автономного использования корпусных технологий или ГИИ студентами, отмечается низкая готовность обучающихся в неязыковых вузах к самостоятельному анализу языковых данных. Таким образом, в данном исследовании была выбрана первая модель.

Генеративный искусственный интеллект использовался для создания текстовых фрагментов и упражнений на основе отобранного списка лексических единиц и грамматических конструкций для запоминания. В процессе изучения иностранных языков исследователями выделяются три этапа, которые позволяют студентам провести разграничение между разными значениями определенной лексической единицы: 1) этап введения, семантизация нового слова и первичного его воспроизведения; 2) этап ситуативной тренировки и создания прочных речевых связей конкретного слова в заданных пределах в однотипных речевых ситуациях; 3) варьирующий ситуативный этап — этап создания динамичных речевых связей [7]. На всех этих этапах требуются аутентичные фрагменты текста, содержащие активизируемые лексические единицы и грамматические конструкции, их поиск может быть осуществлен средствами корпусных технологий или текстовые фрагменты генерируются большими языковыми моделями как автономно, так и с использованием аутентичных иноязычных текстов из открытых интернет-источников.

Полученные текстовые материалы должны содержать лексические единицы (ранее отобранные для запоминания в соответствии с учебной программой по дисциплине «Иностранный язык (английский)»), а также грамматические явления, характерные для научного стиля речи (страдательный залог, формальное подлежащее и дополнение, выраженные местоимением it

и т. д.). Первичное предъявление слов и словосочетаний для запоминания осуществлялось в процессе выполнения упражнений по серии учебных пособий Cambridge Academic English (издательство Cambridge University Press). В эксперименте принимают участие студенты 3-го курса бакалавриата и специалитета, а также 1-го курса магистратуры. В соответствии с методом «интервального повторения», полученные текстовые фрагменты объединились в упражнения, которые выполнялись по следующей схеме: 1-е повторение — сразу после изучения; 2-е повторение — через 20–30 минут после 1-го повторения; 3-е повторение — через 1 день после 2-го повторения; 4-е повторение — через 2–3 недели после 3-го; 5-е повторение — через 2–3 месяца после 4-го повторения [8].

В качестве источника дополнительного контекста для обеспечения многократно предъявления лексических единиц и грамматических конструкций в рамках начального этапа исследования сравниваются чат-бот ChatGPT4, Нейросети и корпус NOW (News on the Web) (<https://www.english-corpora.org/pow/>), который содержит 20,3 млрд слов данных из веб-газет и журналов с 2010 года по настоящее время. Корпус растет примерно на 1,5 млрд слов каждый год. Для работы с чат-ботом использовались две технологии: 1 — поиск предложений с заданными параметрами в англоязычных аутентичных статьях из журнала Economist (в строку с запросом вставлялся отрывок из статьи; 2 — запросы без указания текстов для поиска.

Результаты. В КФ МГТУ им. Н.Э. Баумана с сентября 2024–2025 учебного года проводится эксперимент с целью оценки результативности использования корпусов англоязычных текстов и генеративного искусственного интеллекта для создания текстовых фрагментов и/или их поиска в аутентичных текстовых материалах из заданного источника (онлайн версия журнала Economist). Полученные результаты подтверждают сделанные на основе анализа теоретических источников выводы, а именно:

- генеративный искусственный интеллект позволяет подобрать аутентичный контекст на основе заданных параметров достаточно эффективно, хотя и требует постоянного контроля и коррекции как запросов, так и тщательной проверки полученных результатов;

- количество сгенерированных ГИИ учебных упражнений, удовлетворяет принципам реализации метода интервального повторения, что ведет к более качественному усвоению как лексических единиц, так и грамматических конструкций;

- необходима тщательная доработка существующих больших языковых моделей для сокращения временных затрат на преобразование сгенерированных текстовых фрагментов или упражнений в учебные материалы, пригодные для предъявления обучающимся.

Обсуждение полученных результатов. Несмотря на популярность ГИИ обработка промптов существующими моделями зачастую не позволяет получить методические материалы, удовлетворяющие заданным критериям. Частично данная проблема решается многократной переформулировкой запро-

сов, однако ведет к существенному увеличению временных затрат. Таким образом, требуется дальнейшее проведения практического исследования в соответствии с ранее описанной методикой для получения количественных результатов, позволяющих сделать методически осмысленные практические рекомендации по доработке существующих ГИИ.

Заключение. В связи с геополитической изолированностью Российской Федерации именно генеративный искусственный интеллект в целом и большие языковые модели в частности могут существенно повысить эффективность и производительность в области подготовки иноязычных учебно-методических материалов, удовлетворяющих принципу аутентичности. При этом темп развития информационных технологий требует от современных научно-педагогических кадров не только разработки эффективных методик использования ГИИ, но и постоянного мониторинга направлений развития и новых возможностей искусственного интеллекта.

Список источников

- [1] Константинова Л.В., Ворожихин В.В., Петров А.М., Титова Е.С., Штыхно Д.А. Генеративный искусственный интеллект в образовании: дискуссии и прогнозы. *Открытое образование*, 2023, № 2, с. 36–48.
- [2] Авраменко А.П., Фадеева В.А., Терновский В.В. Опыт интеграции технологий искусственного интеллекта в иноязычное высшее образование от цифровизации к автоматизации. *Вестник Московского университета. Сер. 19. Лингвистика и межкультурная коммуникация*, 2024, № 2, с. 55–66.
- [3] Сысоев П.В. Использование технологий искусственного интеллекта в обучении иностранному языку: тематика методических работ за 2023 год и перспективы дальнейших исследований. *Вестник ТГУ*, 2024, № 2, с. 294–308.
- [4] Абалян Ж.А., Пивнева С.В. Потенциал и риски использования искусственного интеллекта в высшей школе при обучении иностранному языку в профессиональной школе. *МНКО*, 2024, № 3(106), с. 5–7.
- [5] Godwin-Jones R. Distributed agency in language learning and teaching through generative AI. *Language Learning & Technology*, 2024, vol. 28 (2), pp. 5–30. <https://doi.org/10125/73570>
- [6] Шульц О.Е., Первова Г.М., Хаусманн-Ушкова Н.В. Языковой корпус на основе генеративного искусственного интеллекта в обучении иностранному языку: преимущества и недостатки. *Вестник ТГУ*, 2024, № 3, с. 608–616.
- [7] Танцура Т.А. Контекстное обучение иностранному языку в профессиональной подготовке студентов. *МНКО*, 2018, № 4 (71), с. 238–240.
- [8] Иголкина М.И., Язынина В.С. Метод интервальных повторений при обучении иностранным языкам. *Гуманитарный вестник*, 2023, № 3 (101), с. 1–8.