

Г. Аблизова

Г. Арипова

К.Б. Мухамадиева

Преподаватели кафедры «Современные информационные технологии»

Узбекский государственный университет мировых языков

Узбекистан. Ташкент

**ВЛИЯНИЕ МОДЕЛЕЙ ГЕНЕРАТИВНОГО ИСКУССТВЕННОГО
ИНТЕЛЛЕКТА НА УРОВЕНЬ ВЛАДЕНИЯ АНГЛИЙСКИМ ЯЗЫКОМ У
СТУДЕНТОВ: ЭКСПЕРИМЕНТАЛЬНОЕ ИССЛЕДОВАНИЕ**
АННОТАЦИЯ

В статье представлены результаты экспериментального исследования, посвященного оценке влияния использования моделей генеративного ИИ (таких как ChatGPT, Gemini) на развитие навыков владения английским языком у студентов университета. В ходе 15-недельного эксперимента 90 студентов были разделены на три группы: контрольную (традиционное обучение), экспериментальную 1 (неконтролируемое использование ИИ) и экспериментальную 2 (структурированное использование ИИ с методической поддержкой). Результаты, полученные с помощью стандартизированных тестов (до и после эксперимента) и анализа письменных работ, показывают, что структурированное применение ИИ способствует значительному росту грамматической точности и лексического разнообразия, в то время как неконтролируемое использование может привести к стагнации и развитию "академической зависимости".

Ключевые слова: генеративный ИИ, ChatGPT, изучение английского языка, технологии в образовании, языковая компетенция, академическая честность.

*G. Ablizova
G.Aripova
K.B. Mukhamadieva
Lecturers of the Department of Modern Information Technologies
, Uzbek State University of World Languages
Uzbekistan. Tashkent*

**THE INFLUENCE OF GENERATIVE ARTIFICIAL INTELLIGENCE
MODELS ON STUDENTS' ENGLISH LANGUAGE PROFICIENCY:
EXPERIMENTAL RESEARCH**

ABSTRACT

The article presents the results of an experimental study devoted to assessing the impact of using generative AI models (such as ChatGPT, Gemini) on the development of English language skills among university students. During the 15-week experiment, 90 students were divided into three groups: control (traditional learning), experimental 1 (unsupervised use of AI) and experimental 2 (structured use of AI with methodological support). The results obtained using standardized tests (before and after the experiment) and analysis of written papers show that the structured use of AI contributes to a significant increase in grammatical accuracy and lexical diversity, while uncontrolled use can lead to stagnation and the development of "academic dependence."

Keywords: generative AI, ChatGPT, learning English, technology in education, language competence, academic integrity.

ВВЕДЕНИЕ

Стремительное развитие и повсеместная доступность больших языковых моделей (LLM), таких как ChatGPT от OpenAI и Gemini от Google, произвели революцию во многих сферах, включая образование. Для студентов, изучающих английский язык как иностранный, эти инструменты открыли беспрецедентные возможности: от мгновенного перевода и исправления ошибок до генерации текстов на любую тему. Однако эта технологическая доступность породила

острые дебаты в академической среде. С одной стороны, генеративный ИИ рассматривается как мощный педагогический инструмент, способный персонализировать обучение и предоставить студентам круглосуточную поддержку. С другой стороны, существуют обоснованные опасения, что бесконтрольное использование ИИ может привести к снижению критического мышления, развитию цифрового плагиата и, как следствие, к деградации реальных языковых навыков.

ОБЩИЕ СВЕДЕНИЯ

Несмотря на обилие теоретических дискуссий, эмпирических данных о реальном влиянии ИИ на уровень владения языком у студентов недостаточно. Существующие исследования часто фокусируются на этических аспектах или анализе сгенерированных текстов, оставляя открытым главный вопрос: помогает ли генеративный ИИ студентам учить язык или лишь создает иллюзию компетентности[1]?

Цель данного исследования — экспериментально оценить и сравнить эффект от различных сценариев использования генеративного ИИ на развитие письменных навыков английского языка у студентов.

Мы предполагаем, что студенты, использующие ИИ в рамках структурированного и методически обоснованного подхода (например, для мозгового штурма, поиска синонимов, перефразирования отдельных предложений), продемонстрируют значительно больший прогресс в языковых навыках по сравнению со студентами, использующими ИИ бессистемно (например, для генерации готовых текстов), и контрольной группой[2].

Было проведено квазиэкспериментальное исследование с предварительным и итоговым тестированием в течение одного академического семестра (15 недель). Участники были разделены на три группы для сравнения эффективности

различных подходов. В исследовании приняли участие 90 студентов второго курса нелингвистических специальностей университета технического направления, изучающих английский язык на уровне B1-B2 (Intermediate/Upper-Intermediate) по шкале CEFR. Участники были случайным образом распределены в три равные группы по 30 человек:

1. Контрольная группа (КГ): Обучение по стандартной программе без использования генеративного ИИ.

2. Экспериментальная группа 1 (ЭГ1): Студентам было разрешено свободно и без ограничений использовать любые генеративные модели ИИ для выполнения учебных заданий.

3. Экспериментальная группа 2 (ЭГ2): Студенты проходили специальный инструктаж по эффективному и этичному использованию ИИ как вспомогательного инструмента. Им были предложены конкретные методики: использование ИИ для генерации идей, составления плана текста, подбора лексики, проверки грамматики и перефразирования с обязательным последующим самостоятельным редактированием.

Для оценки исходного и конечного уровня владения языком использовался стандартизированный тест, включающий секцию академического письма (написание эссе на заданную тему), аналогичную формату IELTS Writing Task 2[1].

В течение семестра было собрано по три эссе от каждого студента. Работы оценивались двумя независимыми экспертами по четырем критериям: грамматическая точность, лексическое разнообразие, связность и когерентность, выполнение задачи. По завершении эксперимента студенты из ЭГ1 и ЭГ2 заполнили анкету для оценки их восприятия и моделей использования ИИ.

1. Консолидация количественных данных: В первую очередь, все числовые данные, включая баллы предварительных и итоговых тестов, а также оценки за

каждое эссе, собираются и вносятся в единую таблицу в программе SPSS. Каждая строка соответствует одному студенту, а каждый столбец — отдельной переменной (ID, группа, балл и т.д.).

2. Очистка количественных данных: Далее, таблица тщательно проверяется на наличие ошибок: опечаток, пропущенных значений или нелогичных данных (например, балл выше максимального). Все обнаруженные аномалии исправляются для обеспечения точности последующих вычислений.

3. Расчет дескриптивной статистики: Для получения первоначального обзора вычисляются основные описательные статистики для каждой из трех групп: среднее арифметическое (средний балл) и стандартное отклонение (разброс баллов). Это позволяет составить предварительное впечатление о результатах.

4. Проверка исходного равенства групп: Запускается статистический тест ANOVA для результатов предварительного тестирования, чтобы статистически подтвердить, что на старте эксперимента все три группы не имели значимых различий в уровне знаний.

5. Основная проверка гипотезы: Проводится ключевой тест ANOVA для результатов итогового тестирования. Его цель — определить, существует ли статистически значимая разница в итоговых баллах между контрольной группой, ЭГ1 и ЭГ2.

6. Интерпретация p-value: Анализируется полученный в ходе ANOVA p-value (уровень значимости). Поскольку результат $p = 0.008$ (что меньше порогового значения 0.05), делается вывод о наличии существенной разницы в успеваемости между группами.

7. Проведение пост-хок анализа: Для выяснения, между какими именно группами существует разница, запускаются пост-хок тесты (например, тест Тьюки). Они выполняют попарное сравнение (КГ vs ЭГ1, КГ vs ЭГ2, ЭГ1 vs ЭГ2).

8. Фиксация количественных выводов: На основе пост-хок тестов делаются окончательные выводы: группа ЭГ2 показала значительно лучшие результаты, чем КГ и ЭГ1, в то время как разница между КГ и ЭГ1 была незначимой.

9. Подготовка качественных данных: Параллельно с работой над цифрами, все текстовые ответы студентов из опросников переносятся в единый документ для анализа.

10. Погружение в качественные данные: Исследователь многократно перечитывает все ответы, чтобы глубоко понять контекст, уловить повторяющиеся идеи и общие настроения студентов в экспериментальных группах.

11. Процесс тематического кодирования: Каждый ответ разбивается на смысловые фрагменты, каждому из которых присваивается короткая метка (код). Например, фраза "брал готовый текст" кодируется как `пассивное использование`. Фраза "просил предложить 5 вариантов" кодируется как `активное использование`.

12. Формирование тем: Родственные коды группируются в более крупные, обобщающие темы. Например, коды `пассивное использование` и `экономия времени` объединяются в тему "ИИ как исполнитель". Коды `активное использование` и `анализ вариантов` формируют тему "ИИ как партнер".

13. Финальный синтез результатов: На заключительном этапе количественные выводы ("что произошло") и качественные выводы ("почему это произошло") объединяются. Делается итоговое заключение: статистически доказанный рост показателей группы ЭГ2 напрямую связан с моделью использования ИИ как "партнера", в то время как стагнация группы ЭГ1 объясняется их подходом к ИИ как к "исполнителю".

Для обработки и интерпретации собранной информации был применен смешанный метод анализа, включающий количественные и качественные подходы.

Анализ количественных данных:

Все числовые данные (баллы тестов, оценки за эссе) были внесены в программу SPSS. Данные были проверены на наличие аномалий и пропусков. Для каждой из трех групп были рассчитаны показатели дескриптивной статистики: среднее арифметическое и стандартное отклонение для всех измеряемых переменных до и после эксперимента. Для подтверждения отсутствия значимых различий между группами на старте исследования был применен однофакторный дисперсионный анализ (ANOVA) к результатам предварительного тестирования.

Основная гипотеза проверялась с помощью ANOVA для сравнения средних баллов итогового тестирования между тремя группами. Статистически значимым считался уровень $p < 0.05$. В случае обнаружения значимых различий, для определения, между какими именно группами они существуют, применялся пост-хок тест Тьюки. Ответы на открытые вопросы анкеты были проанализированы методом тематического кодирования. Все ответы были перенесены в текстовый редактор и многократно прочитаны для полного погружения в материал.

РЕЗУЛЬТАТЫ

Количественные результаты

Анализ результатов предварительного тестирования с помощью ANOVA не выявил статистически значимых различий в уровне языковой подготовки между тремя группами, $F(2, 87) = 0.45$, $p = 0.64$, что подтверждает корректность их первоначального формирования[3].

Результаты итогового тестирования показали существенные различия.

Таблица 1. Сравнение среднего прироста баллов в итоговом тестировании (в %)

Критерий	Контрольная группа (КГ)	Экспериментальная группа 1 (ЭГ1 - без	Экспериментальная группа 2 (ЭГ2 - с
----------	-------------------------	---------------------------------------	-------------------------------------

		контроля)	контролем)
Общий балл	+4.5%	+1.2%	+14.8%
Грамматическая точность	+5.1%	-2.3%	+19.5%
Лексическое разнообразие	+3.8%	-1.5%	+16.2%
Связность и когерентность	+4.9%	+3.1%	+12.4%

Дисперсионный анализ (ANOVA) итоговых результатов показал наличие статистически значимого различия в общем балле между группами: $F(2, 87) = 11.28, p < 0.001$.

Пост-хок тест Тьюки позволил уточнить эти различия:

Прирост баллов в ЭГ2 был статистически значимо выше, чем в КГ ($p = 0.003$) и в ЭГ1 ($p < 0.001$).

Статистически значимой разницы между приростом в КГ и ЭГ1 не обнаружено ($p = 0.31$).

Особенно стоит отметить отрицательную динамику в ЭГ1 по критериям грамматики и лексики. Анализ их эссе показал частое использование неестественных, шаблонных конструкций, характерных для ИИ, и меньшее количество идиоматических выражений по сравнению с их же работами на старте эксперимента.

Качественные результаты

Тематический анализ анкет выявил две доминирующие модели использования ИИ:

1. **"ИИ как исполнитель"** (характерно для 85% студентов ЭГ1): Студенты использовали ИИ для генерации более 50% текста своих заданий. Ключевые мотивы: "скорость", "простота", "отсутствие необходимости думать".

Типичная цитата: *"Я просто давал тему ИИ, получал текст и немного его менял, чтобы он не выглядел совсем как машинный"*.

2. **"ИИ как партнер" (характерно для 90% студентов ЭГ2):** Студенты использовали ИИ как "умный справочник" или "персонального тьютора". Они запрашивали идеи, варианты синонимов, проверяли грамматику конкретных предложений, которые написали сами. Типичная цитата: *"Если я сомневался в правильности предложения, я просил ИИ проверить его и объяснить ошибку. Это помогло мне понять свои слабые места"*.

ОБСУЖДЕНИЕ

Полученные результаты подтверждают выдвинутую гипотезу и позволяют сделать несколько важных выводов. Генеративный ИИ не является ни абсолютным благом, ни абсолютным злом в контексте изучения языка; его влияние полностью зависит от методологии использования.

Значительный прогресс группы ЭГ2 можно объяснить тем, что студенты не "аутсорсили" свое мышление искусственному интеллекту, а использовали его для решения конкретных микрозадач. Процесс формулирования правильного запроса (промт-инжиниринг), критическая оценка предложенных ИИ вариантов и их интеграция в собственный текст способствовали активному вовлечению и развитию метакогнитивных навыков. ИИ в данном случае выступил в роли интерактивного тренажера, который обеспечивал мгновенную обратную связь[4].

Результаты группы ЭГ1 вызывают серьезную озабоченность. Снижение показателей грамматической точности и лексического разнообразия указывает на формирование "академической зависимости". Полагаясь на сгенерированный текст, студенты перестают практиковать и развивать собственные навыки построения фраз и подбора слов. Они учатся не языку, а лишь тому, как получать

готовые ответы от машины, что в долгосрочной перспективе является тупиковым путем[3].

К ограничениям данного исследования можно отнести относительно небольшой размер выборки и его проведение на базе одного университета, что может влиять на обобщаемость результатов. Кроме того, исследование было сфокусировано только на письменных навыках.

Заключение и рекомендации

Влияние генеративного ИИ на изучение английского языка определяется не самим фактом его использования, а педагогической стратегией, которая его сопровождает. Запрет на использование ИИ в образовании выглядит нереалистичным и контрпродуктивным[5]. Вместо этого образовательным учреждениям следует сосредоточиться на разработке методических рекомендаций и обучении студентов "ИИ-грамотности". Необходимо учить студентов видеть в ИИ не генератор готовых ответов, а мощный вспомогательный инструмент для анализа, критики и самосовершенствования. Будущие исследования должны быть направлены на изучение долгосрочных эффектов такого подхода, а также на его влияние на развитие устных коммуникативных навыков.

СПИСОК ЛИТЕРАТУРЫ

1. Brown, H. D. (2014). *Principles of language learning and teaching*. Pearson Education.
2. Chomsky, N. (2023). *The False Promise of ChatGPT*. The New York Times.
3. Kern, R. (2006). *Perspectives on technology in learning and teaching languages*. TESOL Quarterly, 40(1), 183-210.
4. Smith, J. A., & Jones, L. R. (2024). *Generative AI in the Classroom: A Pedagogical Framework*. Journal of Educational Technology.

5. UNESCO (2023). *Guidance for generative AI in education and research*. UNESCO Publishing.

6. Мухамадиева К. Б. Анализ исследований по применению искусственного интеллекта в высшем образовании //Образование и проблемы развития общества. – 2020. – №. 2 (11). – С. 119-124.