

ДИДАКТИЧЕСКИЙ ПОТЕНЦИАЛ И ОГРАНИЧЕНИЯ CHATGPT ДЛЯ ВЫПОЛНЕНИЯ ПИСЬМЕННЫХ ЗАДАНИЙ



Аннотация

В статье рассматривается дидактический потенциал технологий генеративно-го ИИ при развитии письменно-речевых умений. Приводятся результаты эксперимента по выполнению письменного задания в сотрудничестве с чат-ботом на базе GPT. На основании анализа количественных данных эксперимента сделаны выводы о том, на какие критерии оценки влияет обработка текста с помощью технологий ИИ; выявлены ограничения применения данного средства обучения при работе с письменными текстами.



Abstract

PROSPECTS AND LIMITATIONS OF CHATGPT FOR PERFORMING WRITING TASKS

The article deals with the didactic options of generative artificial intelligence technologies within the development of writing skills. The results of performing written assignments in cooperation with a GPT-based chatbot are presented. On the basis of analyzing the quantitative data, conclusions are drawn about the assessment criteria that are influenced by text editing with the help of artificial intelligence technologies. The limitations of using this learning tool for editing texts are highlighted.



Ключевые слова

генеративный искусственный интеллект, формирование письменно-речевых умений, трансформация критериев оценивания письменных заданий



Keywords

generative artificial intelligence, productive writing strategies, transformation of the writing tasks criteria

Введение. Развитие технологий ИИ, включая распознавание и синтез речи, обработку естественного языка, генеративные нейронные сети, позволило по-новому взглянуть на проблему создания благоприятных условий для речевой практики. Однако, несмотря на ежедневный технологический прогресс в области создания больших языковых моделей, существуют и ограничения, связанные с применением ИИ в области обучения ИЯ.

Технологии генеративного искусственного интеллекта (ИИ), а именно – большие языковые модели (*Large Language Models, LLM*), работающие на базе алгоритмов обработки естественного языка (*Natural Language Processing, NLP*), могут выступать в качестве партнера по обучению (*learning partner*) для учащихся. По сравнению с более ранними образцами, выпущенный компанией *OpenAI*

30 ноября 2022 года *ChatGPT (Generative Pre-trained Transformer)* более эффективен при генерации текстов благодаря интеграции внутри архитектуры типа «трансформер» различных методов NLP. Принципиальным отличием языковых моделей такого типа от других генеративных технологий можно считать их адаптивность – учет контекста и обратной связи, получаемой от конкретного пользователя. Поскольку *ChatGPT* и его российские аналоги (*GigaChat* от Сбер и *YaGPT* от Яндекс) доступны круглосуточно, они открывают перед учениками широкий дидактический потенциал для самообучения.

Нами были проанализированы исследования, проведенные отечественными методистами и педагогами за первый год существования *ChatGPT* [1–6]. По нашему мнению, понимание технических особенностей и ограничений функционирования диалоговых систем вкупе с рефлексией по поводу существующих дидактических практик позволит определить необходимые шаги для адаптации данных систем к решению насущных методических задач (например, повышение эффективности чат-ботов как симуляторов реальной коммуникации на иностранных языках), а также формулирования и последующего решения новых учебных задач, появляющихся в связи с вхождением наиболее продвинутых диалоговых систем в жизнь участников образовательного процесса.

К дидактическим свойствам *GPT*, как и других технических средств обучения (ТСО), относятся адаптивность, интерактивность, доступность, гибкость и открытость. Надо отметить, что дидактические функции рассматриваемой технологии реализуются в автоматизированном режиме. Основные вызовы для учителей иностранных языков представляют такие возможности искусственного интеллекта, как генерация идей, а также создание и редактирование текста.

Методика эксперимента. Целью эксперимента стало выявление дидактических возможностей и ограничений технологий генеративного ИИ при выполнении письменных заданий. В качестве пула испытуемых была выбрана аудитория старшеклассников и студентов от 16 (9–10 классы) до 20 лет (студенты 1 и 2 курсов) с уровнем владения языком не ниже B1. Эксперимент проводился на базе двух общеобразовательных школ – в г. Брянске (ГБОУ Брянский городской лицей № 1 имени А.С. Пушкина) и г. Тюмени (МАОУ гимназия № 21 с углубленным изучением иностранных языков), а также на двух языковых факультетах МГУ имени М.В. Ломоносова – Высшая школа перевода (ВШП) и Факультет иностранных языков и регионоведения (ФИЯР).

Всем участникам эксперимента было рекомендовано использовать чат-бот, который не предлагает выбора версии, – «Марти *ChatGPT*» в мессенджере «Телеграм» – во избежание обращения к разным по своим характеристикам версиям, например *GPT-3.5* и *GPT-4*. В качестве письменного задания было выбрано написание отзыва (*review*), так как данный жанр письменной речи не является наиболее частотным при обучении ИЯ, а значит, все обучающиеся имеют схожий уровень подготовки (нет большого опыта написания подобных текстов). Стандартная формулировка и само задание по написанию *review* были адаптированы с целью снижения порога входа: отзыв требовалось написать на предлагаемую статью, а также обучающимся был дан образец. Далее приводим формулировку задания и начало предложенной статьи:

Your teacher has asked you to write a review for an article you have read recently. Review the article giving your opinion and say whether or not you would recommend it. Write for not longer than 150 words.

Compare online and traditional classroom education: the first thing that comes to mind is the use of computers and physical classrooms. If you are familiar

with technology, you may know that there are various software and online tools available to make online learning easier and to connect with your tutors and friends instantly...

Необходимо подчеркнуть, что использованная нами аутентичная статья была адаптирована и сокращена, чтобы сделать ее наиболее доступной широкому кругу обучающихся с уровнем владения языком от B1 и выше. Для эксперимента были разработаны особые критерии оценивания, так как довольно обобщенные параметры оценки, характерные для *First Certificate Exam (FCE) (content, communication achievement, organisation, language)*, не позволяют в полной мере дифференцировать разноуровневых студентов по основным качественным и количественным характеристикам письменной речи. Работа, выполненная по заданным критериям, могла получить максимум 10 баллов:

- 1) коммуникативная задача (максимум 5 баллов);
- 2) грамматическая корректность (максимум 3 балла);
- 3) лексическая корректность (максимум 2 балла).

При оценивании решения коммуникативной задачи учитывались следующие микроумения: связность / деление на абзацы, наличие вступления, основной описательной части, рекомендаций, выражение личного мнения. Стоит отметить, что сама процедура оценивания проводилась максимально лояльно: каждый обучающийся получал итоговый кумулятивный балл при любой оценке по критерию «коммуникативная задача». Данный подход позволил изучить даже работы тех обучающихся, которые не в полной мере решили коммуникативную задачу.

Важным дополнением к выбранной для эксперимента системе оценки стало использование статистического критерия *Vocabulary range* на основе *Oxford Text Checker*: он позволяет количественно установить принадлежность лексики к уровням *CEFR*. С учетом сложности задания для целей дифференциации учитывалась лексика уровней B1, B2, C1: было подсчитано суммарное процентное соотношение для каждого испытуемого по списку *Oxford 5000*.

Мы изначально предполагали проведение эксперимента с разновозрастными участниками, при этом ключевой задачей являлось наблюдение за использованием *ChatGPT* в сравнимых по языковому уровню группах обучающихся. Данное намерение повлияло на ход проведения эксперимента – была разработана следующая последовательность проведения эксперимента:

- 1) Вводное тестирование на определение уровня владения ИЯ с использованием электронного ресурса *Cambridge*¹ (целевой показатель – не ниже B1).
- 2) Самостоятельное написание обучающимися отзывов (*review*) после ознакомления с образцом.
- 3) Редактирование написанного при помощи чат-бота «Марти *ChatGPT*» с целью улучшения лингвистических характеристик письменной работы (всем испытуемым была предоставлена полная свобода при формулировании запроса, при этом было рекомендовано обращаться на английском языке).

¹ <https://www.cambridgeenglish.org/test-your-english/general-english/>

Результаты эксперимента. Отбор работ для проведения оценивания состоял из двух процедур: 1) на первом этапе были исключены работы обучающихся с уровнем владения ИЯ ниже B1 (один из 54 испытуемых); 2) затем – работы, в которых либо отсутствовал первоначальный текст отзыва до редактирования, либо не была сохранена отредактированная версия с запросом к *ChatGPT* (еще 9). Таким образом, итоговый анализ был проведен на основе 44 работ.

Мы оценили *reviews*, написанные обучающимися самостоятельно, и версии тех же самых работ после их доработки с помощью чат-бота. Следующим шагом сопоставили средние результаты при создании самостоятельных работ со средними результатами по текстам, созданным совместно с «Марти ChatGPT». Стоит отметить, что сама процедура оценивания проводилась максимально лояльно: каждый обучающийся получал итоговый кумулятивный балл при наличии описательной части и выражении собственного мнения, т.е. по двум критериям сразу. За самостоятельно выполненные задания обучающиеся в среднем получили 7,68 балла, в то время как те же работы после взаимодействия с *ChatGPT* показали более высокое значение – 8,68 балла (представлены ниже на рис. 1).

Эти различия по парному t-критерию Стьюдента можно считать статистическими значимыми (на уровне значимости 0,05). При этом стоит отметить, что рост общего балла обусловлен разнонаправленным влиянием *ChatGPT* на авторские работы испытуемых. На рис. 2 представлены средние значения только по критерию выполнения коммуникативной задачи.

Снижение среднего значения после использования *ChatGPT* не было ярко выраженным. В сотрудничестве с чат-ботом обучающиеся получили за решение коммуникативной задачи в среднем 3,68 балла против 3,75 за текст до вмешательства *ChatGPT*. Следовательно, улучшение качества выполнения задания обусловлено тем, что по двум другим критериям («грамматические ошибки» и «лексические ошибки») испытуемые в сотрудничестве с чат-ботом получили максимум по обоим показателям, т.е. 3 и 2 балла соответственно, в то время как средний показатель без вмешательства *ChatGPT* по грамматическим ошибкам составил 2,3 балла, а по лексическим – 1,61.

Безусловно, отдельного внимания заслуживает сравнение на основе введенного нами дополнительного критерия – уровня сложности лексики по шкале *CEFR*. На представленном ниже рис. 3 – графике линейной регрессии – можно заметить отсутствие прямой зависимости между итоговой оценкой за предложенное задание и уровнем лексики в работах студентов без вмешательства *ChatGPT*, т.е. сформулированных самостоятельно.

В работах обучающихся, получивших 9 итоговых баллов, процент лексики от B1 и выше варьируется от 4 % до 21 %. При этом для всех испытуемых средний уровень наличия лексики высоких уровней (от B1 и выше) составлял 10,18 % от общего

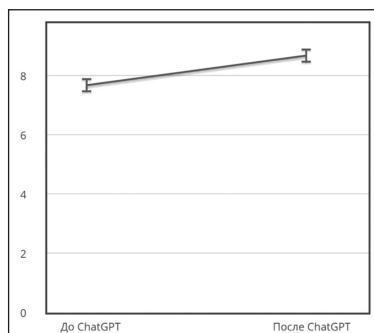


Рис. 1
Сравнение средних показателей

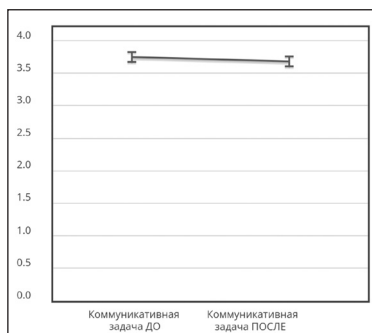


Рис. 2
Средние показатели по коммуникативной задаче

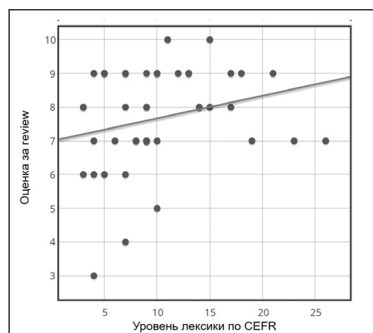


Рис. 2
Регрессионный анализ: оценка / уровень лексики

объема текста. Приведем пример отзыва, в котором объем лексики сложных уровней в первоначальном отзыве составлял 3 % и был увеличен до 22 %.

Пример первоначального отзыва в авторской редакции: *I think that e-learning is more comfortable and available for many people. It gives you the opportunity to study anywhere in the world. If there are no courses in the place where you live that you would like to take, then you can find them on the Internet. However, for proper e-learning you need to be organized and have a high level of self-control that not many people have. Some people prefer full-time education because of the opportunity to study with their peers.*

Пример того же отзыва после доработки ChatGPT: *I personally believe that e-learning offers numerous advantages, particularly in terms of comfort and accessibility. It allows individuals to pursue their studies from anywhere in the world, eliminating geographical limitations. For those who reside in areas where desired courses are unavailable, the internet provides a vast array of options to satisfy their educational interests. However, it is important to acknowledge that successful e-learning necessitates a high level of self-discipline and organizational skills, which may not come naturally to everyone. Some individuals may prefer a traditional full-time education due to the opportunity it provides for in-person interaction with peers and instructors. The classroom setting offers immediate feedback, group discussions, and collaborative projects that can greatly enrich the learning experience.*

Сотрудничество с ChatGPT, как было сказано ранее, позволило полностью искоренить лексические ошибки, т.е. любые неграмматические ошибки, затрудняющие понимание текста. Тем не менее это не единственный существенный эффект влияния чат-бота на лексический аспект созданных текстов. Так, для всех испытуемых средний уровень наличия лексики высоких уровней существенно вырос после обращения к ChatGPT – 17,5 % в общем объеме текста. Важно отметить, что помощь ChatGPT не оказала положительного влияния на результаты испытуемых с более высоким уровнем владения ИЯ (об этом говорят аналогичные или даже более низкие баллы при сравнении) и не повысила личного уровня владения ИЯ у остальных обучающихся.

Заключение. Проведенный нами количественный анализ результатов оценки отзывов до и после редактирования выявил ряд особенностей чат-бота на базе ChatGPT, который может:

- 1) способствовать получению более корректного с точки зрения лексики и грамматики письменного-речевого продукта;
- 2) не влиять существенно на качество выполнения коммуникативной задачи;
- 3) значительно усложнить лексику полученного текста.

Представленные выводы подтверждают гипотезу о том, что технологии генеративного ИИ могут быть использованы при выполнении письменных заданий. Однако был выявлен тот факт, что эффективным ChatGPT может быть преимущественно для повышения оценки за языковой компонент. На основании результатов эксперимента можно рекомендовать применение технологий ИИ для редактирования и улучшения текста. В то же время нельзя не отметить необходимость поэтапного выполнения и проверки письменных заданий (например, в формате микрообучения [7]) для минимизации угрозы генерации целых письменных-речевых продуктов. Перспективным направлением для дальнейшего исследования представляется качественный анализ запросов (промптов) обучающихся.



Литература

1. Гаркуша Н. С., Городова Ю. С. Педагогические возможности ChatGPT для развития когнитивной активности студентов // Профессиональное образование и рынок труда. – 2023. – № 1 (11). – С. 6–23.
2. Сысоев П. В. Технологии искусственного интеллекта в обучении иностранному языку // Иностр. языки в школе. – 2023. – № 3. – С. 6–16.
3. Сысоев П. В., Филатов Е. М. ChatGPT в исследовательской работе студентов: запрещать или обучать? // Вестник Тамбовского университета. Серия: Гуманитарные науки. – 2023. – Т. 28. – № 2. – С. 276–301.
4. Титова С. В. Технологии разработки и реализации веб-проектов в языковом классе // Иностр. языки в школе. – 2023. – № 3. – С. 17–23.
5. Фадеева В. А., Щедромйрская А. И. Возможности технологий искусственного интеллекта в цифровизации образовательной среды // Иностр. языки в школе. – 2023. – № 3. – С. 81–87.
6. Харламенко И. В. Чат-боты в обучении английскому языку // Иностр. языки в школе. – 2023. – № 3. – С. 55–59.
7. Авраменко А. П. Индивидуализация процесса формирования иноязычной лексической компетенции на основе микрообучения // Вестник Тамбовского университета. Серия: Гуманитарные науки. – 2020. – Т. 25 – № 184. – С. 25–34.



Сведения
об авторах

Авраменко Анна Петровна, канд. пед. наук, доцент кафедры теории преподавания иностранных языков факультета иностранных языков и регионоведения МГУ имени М. В. Ломоносова.

E-mail: avram4ik@gmail.com

Тарасов Алексей Александрович, канд. пед. наук, преподаватель кафедры теории и практики английского языка Высшей школы перевода (факультет) МГУ имени М. В. Ломоносова.

E-mail: lexicon.msk@gmail.com

Никитина Татьяна Сергеевна, учитель английского языка, ГБОУ Брянский городской лицей № 1 имени А. С. Пушкина.

E-mail: tanya3g148@yandex.ru

Чеканов Алексей Андреевич, учитель английского языка, МАОУ гимназия № 21 с углубленным изучением иностранных языков г. Тюмени.

E-mail: chekan6966@yandex.ru