

Методика

DOI: 10.31862/2500-2953-2026-2-49-57

УДК 372.881.1+004.89

А.П. Авраменко

Московский педагогический государственный университет,
119435 г. Москва, Российская Федерация;

Московский государственный университет имени М.В. Ломоносова,
119991 г. Москва, Российская Федерация

Сравнительный анализ дидактического потенциала и ограничений больших языковых моделей для генерации учебных материалов по иностранным языкам

В статье представлен теоретический анализ и эмпирическое исследование роли искусственного интеллекта (ИИ) в профессиональной деятельности преподавателя. Рассматриваются современные подходы к формированию ИИ-компетенции педагога, включая мотивационно-целевой, этический и нормативно-правовой компоненты, а также промпт-инжиниринг. Особое внимание уделено дидактически обоснованным стратегиям использования генеративных больших языковых моделей (LLMs): от создания учебных материалов до анализа обратной связи. В эмпирической части статьи описан эксперимент по сравнению десяти современных LLMs (включая GPT-5, Qwen3-Max, YaGPT 5 и др.) при генерации грамматических и лексических тестов для студентов уровня В2. Результаты показывают значительную вариативность в качестве сгенерированных материалов: наиболее эффективным оказался Qwen3-Max, продемонстрировавший понимание требований и соответствие уровню сложности. Сделан вывод о том, что ИИ выступает

© Авраменко А.П., 2026



Контент доступен по лицензии Creative Commons Attribution 4.0 International License
The content is licensed under a Creative Commons Attribution 4.0 International License

не как замена преподавателю, а как инструмент, расширяющий его профессиональные возможности при условии критической верификации и педагогически осмысленного применения.

Ключевые слова: искусственный интеллект, преподаватель, ИИ-компетенция, генеративные модели, промпт-инжиниринг, автоматизация создания учебных текстов, методика преподавания иностранных языков

ДЛЯ ЦИТИРОВАНИЯ: Авраменко А.П. Сравнительный анализ дидактического потенциала и ограничений больших языковых моделей для генерации учебных материалов по иностранным языкам // Рема. Rhema. 2026. № 2. С. 49–57. DOI: 10.31862/2500-2953-2026-2-49-57

DOI: 10.31862/2500-2953-2026-2-49-57

A. Avramenko

Moscow Pedagogical State University,
Moscow, 119435, Russian Federation

Lomonosov Moscow State University,
Moscow, 119991, Russian Federation

A comparative analysis of the pedagogical efficacy and limitations of large language models for generating ESL educational materials

This article presents a theoretical analysis and empirical study of the role of artificial intelligence (AI) in the professional practice of educators. It examines contemporary approaches to developing teachers' AI competence, encompassing motivational-goal, ethical, legal-regulatory, and prompt engineering components. Special attention is given to evidence-based teaching strategies enabled by generative large language models (LLMs) from creating instructional materials to analyzing student feedback. The empirical section describes an experiment comparing eight state-of-the-art LLMs (including GPT-5, Qwen3-Max, YaGPT 5, etc.) in generating grammar and vocabulary tests for B2-level university students. Results reveal significant variability in output quality: Qwen3-Max demonstrated the strongest alignment with

methodological standards and appropriate task complexity. The study concludes that AI should not be viewed as a replacement for teachers but as a supportive tool that enhances professional capacity – provided its outputs undergo critical verification and pedagogically informed adaptation.

Key words: artificial intelligence, teacher, AI competence, generative models, prompt engineering, automation, methodology of foreign language teaching

FOR CITATION: Avramenko A. A comparative analysis of the pedagogical efficacy and limitations of large language models for generating ESL educational materials. *Rhema*. 2026. No. 2. Pp. 49–57. (In Rus.). DOI: 10.31862/2500-2953-2026-2-49-57

Введение

В последние годы ускорился процесс цифровой трансформации системы образования в связи с развитием искусственного интеллекта (ИИ), а именно генеративных больших языковых моделей (Large Language Models, LLMs). Данные технические средства обучения (ТСО) все активнее интегрируются в педагогическую практику, предлагая преподавателям новые возможности для автоматизации рутинных задач, персонализации учебных материалов и форматов заданий. Однако надежность применения ИИ в учебном процессе остается предметом научного анализа и эмпирической проверки.

Актуальность исследования обусловлена тем, что LLM открывают новые возможности для персонализированного обучения, генерации учебных материалов и моделирования реальных коммуникационных ситуаций, что особенно ценно в условиях отсутствия языковой среды [Caines, 2023]. Инструменты на основе LLM позволяют интенсифицировать и ускорить процесс освоения всех четырех видов речевой деятельности, обеспечивая необходимую практику и мгновенную обратную связь [Cherednichenko, 2024]. Однако использование ИИ в образовании сопряжено с рядом ограничений. Критически важна оценка качества ИИ-сгенерированных материалов, особенно по критериям подлинности, точности, легитимности и соответствия, которые могут быть нарушены в процессе генерации [Huang, 2025]. Необходим гибридный подход, где ИИ выступает вспомогательным инструментом, а не заменой преподавателя, обеспечивая автоматизацию рутинных задач при сохранении человеческого взаимодействия [Sharma, 2025].

Особую значимость приобретает вопрос о том, насколько точно и методологически корректно ИИ-инструменты могут выполнять функции, традиционно возлагаемые на преподавателя, в частности, разработку

учебных материалов. В условиях растущего числа доступных LLMs (таких как GPT, Gemini, Grok, Qwen, YaGPT и др.) возникает необходимость в систематическом сравнительном анализе их дидактического потенциала и ограничений. Целью настоящего исследования является эмпирическая оценка эффективности десяти современных больших языковых моделей при создании заданий на закрепление грамматического и лексического материала для студентов первого курса бакалавриата с уровнем владения иностранным языком B2.

Теоретический обзор

Современный этап развития образования характеризуется формированием новой триады взаимодействия «педагог – обучающийся – искусственный интеллект» [Сысоев, 2025]. В этой модели ИИ берет на себя некоторые функции: проверку заданий, генерацию учебных материалов, адаптацию содержания обучения под индивидуальные потребности студентов, освобождая время преподавателя для реализации более сложных и творческих аспектов обучения [Курганова, Лапчик, 2024; Каптерев, 2025]. Особое значение приобретают задачи, которые могут быть значительно упрощены с помощью ИИ:

- генерация множества разнообразных примеров;
- предоставление множества объяснений одного и того же понятия;
- анализ обратной связи от студентов и др. [Mollick, Mollick, 2023].

ИИ выступает не как замена педагогу, но как новый вид ТСО, расширяющий границы цифровизации и автоматизации элементов учебного процесса. Все материалы, сгенерированные ИИ, требуют тщательной верификации: проверки на фактологическую точность (ИИ склонен к «галлюцинациям»), соответствие уровню аудитории и педагогическую целесообразность. Однако, как отмечают С.В. Титова и И.В. Харламенко, несмотря на растущий интерес к ИИ в образовании, остается недостаточно исследований, посвященных эмпирической оценке качества сгенерированных ИИ-материалов [Титова, Харламенко, 2025]. Анализ систематического обзора показывает значительный дисбаланс в научной литературе: 65% исследований посвящены применению ИИ в обучении студентов, тогда как лишь 35% фокусируются на профессиональном развитии самих преподавателей [Tan, Cheng, Ling, 2025].

В этом контексте особенно важна концепция антропоцентрического подхода. Опросы преподавателей и студентов в Гонконге показали, что подавляющее большинство респондентов считает ИИ неспособным заменить учителя, поскольку он лишен эмоционального интеллекта, способности к эмпатии и установлению личной связи, которые

являются основой для развития социальных и эмоциональных компетенций учащихся [Chan, Tsi, 2024]. Иными словами, происходит переосмысление роли преподавателя. Если раньше его часто воспринимали как источник знаний, то в эпоху ИИ он становится наставником и дизайнером интерактивной обучающей среды (ИОС). На сегодняшний день ИКТ-компетенция педагога дополняется следующими компонентами:

- понимание нормативно-правовой базы внедрения инструментов ИИ;
- промпт-инжиниринг, т.е. умение формулировать запрос;
- критическая оценка и верификация сгенерированных ИИ учебных материалов;
- этическое сопровождение процесса обучения, включая обсуждение вопросов авторства, плагиата и предвзятости алгоритмов;
- педагогическая рефлексия на уровне персонализации обучения [Резаев, 2023; Елсакова, Маркус, 2024; Fernández Jiménez, 2024; Сысоев, 2025].

Таким образом, роль педагога из исполнителя учебного плана переходит в фазу активного исследователя собственного педагогического опыта. Анализируя данные об успеваемости, вовлеченности и обратной связи студентов, ИИ помогает преподавателю выявлять не только проблемы в знаниях учащихся, но и слабые места в собственной методике [Zhang, Zhang, 2024]. Более того, преподаватели все чаще используют ИИ-инструменты для анализа актуальности и релевантности учебных материалов. На основе анализа больших массивов данных (например, запросов в профессиональных сообществах, трендов на рынке труда) нейронные сети помогают педагогу принимать обоснованные решения о том, какие темы следует углубить, а какие сократить или исключить [Karataş et al., 2024]. ИИ становится партнером в динамическом проектировании учебного процесса, что особенно важно в быстро меняющихся дисциплинах. Ключевым барьером для широкого внедрения ИИ в образовательный процесс остается недостаток целенаправленной педагогической подготовки кадров, который можно объяснить фокусом программ повышения квалификации на техническом освоении инструментов (цифровой грамотности), а не на практико-ориентированных сценариях [Kucuk et al., 2022].

Наше исследование акцентирует внимание на втором и третьем компонентах ИИ-компетенции педагога, а именно в эмпирической проверке эффективности десяти различных языковых моделей в разработке грамматических и лексических упражнений для студентов первого курса с уровнем B2, что позволит выявить их педагогические возможности и ограничения в образовательном контексте.

Результаты эксперимента

В эмпирической части исследования был проведен лабораторный эксперимент по сравнению десяти современных больших языковых моделей в решении типовой задачи – создании заданий на закрепление языкового материала для студентов первого курса бакалавриата (уровень B2). Основным ограничением эксперимента выступило использование запросов формата нулевого промта (zero-shot prompt), т.е. без дальнейшего уточнения и продолжения диалога. Данное ограничение объясняется первичным этапом эксперимента, который может быть продолжен сравнением моделей при нескольких итерациях, что требует дополнительной разработки сценариев диалога педагога с виртуальным ассистентом. Эксперимент состоял из двух частей. В первой части моделям был предложен следующий нулевой запрос для генерации грамматического теста: *You are an ESL instructor. Create a 10-question multiple-choice grammar test on English tenses for first-year university students (level B2)*. Во второй части использовался аналогичный запрос для создания лексического теста по теме «studying abroad»: *You are an ESL instructor. Create a 10-question multiple-choice vocabulary quiz on the topic of studying abroad, for first-year university students (level B2)*. Метриками, или критериями оценки результатов работы моделей, выступили:

- уровень используемой лексики (проверялся согласно списку Oxford 5000 с помощью инструмента TextChecker);
- педагогический компонент (вариативность дистракторов, комментарии к правильному ответу и т.д.);
- охват темы.

Исходя из ранее проведенных нами интервью с преподавателями были проанализированы следующие наиболее популярные в учебном контексте международные модели в актуальных версиях на момент проведения эксперимента (октябрь 2025 г.), а также две ведущие по охвату российские разработки: Claude Sonnet 4.5; DeepSeek V.3.1; Gemini 2.5 Flash; GigaChat 2; GPT 5; Grok 4; Mistral Large 2; Qwen3-Max; Perplexity 2.55.0; YaGPT 5.

Анализ показал, что все модели успешно сгенерировали грамматические задания, не усложняя уровень используемой лексики (A1 от 80% у Grok до 98% у YaGPT). Тем не менее, существенно различалась глубина охвата темы:

- GigaChat 2 и Qwen3-Max включили в тест конструкцию *used to* (наиболее узкий фокус продемонстрировала модель YaGPT 5, ограничившись базовыми временами без объяснения правильных ответов);

- GPT 5 и Qwen3-Max пошли дальше, предложив задания, в которых требовалось выбирать между двумя временами в одном предложении;
- Claude, DeepSeek, Grok, Qwen сразу предоставили ключи с пояснениями, что экономит время преподавателя; в то время как GPT 5, напротив, сначала выдал только вопросы для задания.

По лексическому тесту ключевым критерием стала релевантность лексики теме «studying abroad» и ее соответствие заданному уровню. Лидером по качеству лексического наполнения стала GPT 5, которая использовала 10 из 10 слов/выражений нужного уровня, включая идиоматические единицы. Qwen3-Max и Perplexity 2.55.0 также показали высокий результат (8/10), дополнив задания полезными комментариями и обобщениями. GigaChat 2, напротив, использовал только 4 из 10 релевантных лексических единиц. YaGPT 5 сгенерировал задания в формате карточек на запоминание без контекста, что не соответствует запрошенному формату задания.

Таким образом, согласно нашим метрикам, в обеих задачах Qwen3-Max продемонстрировал наименьшее стремление к упрощению и формализму.

Выводы

Проведенный эксперимент позволяет прийти к заключению о том, что ИИ-инструменты способны генерировать формирующие и диагностические материалы, однако их качество варьируется в зависимости от используемой модели. Ни одна из моделей не гарантирует безусловного соответствия заявленному уровню сложности и методическим стандартам. Наиболее эффективным в решении выбранных задач оказался Qwen3-Max, который продемонстрировал понимание требований уровня B2, использовал релевантную лексику и грамматические конструкции.

Отметим, что роль преподавателя остается центральной, поскольку даже самые продвинутые модели требуют верификации сгенерированного контента на предмет лингвистической точности, методической корректности и соответствия учебным целям. ИИ выступает не как автономный субъект, а как ТСО, расширяющее профессиональные возможности педагога. Промпт-инжиниринг является одной из ключевых компетенцией современного преподавателя.

Перспективным направлением для дальнейших исследований видится разработка стандартизированных протоколов оценки дидактического потенциала LLMs, а также создание методических рекомендаций по адаптации запросов под конкретные учебные дисциплины и уровни подготовки обучающихся и разработка сценариев интеракции педагога с ботом.

Библиографический список / References

Елсакова, Маркусъ, 2024 – Елсакова Р.З., Маркусъ А.М. Повышение квалификации преподавателей вуза в области искусственного интеллекта: современное состояние // Высшее образование в России. 2024. Т. 33. № 11. С. 73–94. [Elsakova R.Z., Markus A.M. Advanced training for university teachers in the field of artificial intelligence: Current state. *Vysshee obrazovanie v Rossii*. 2024. Vol. 33. No.11. Pp. 73–94. (In Rus.)]

Каптерев, 2025 – Каптерев А.И. Роль человеко-машинного взаимодействия в педагогическом дизайне // Вестник МГПУ. Серия: Информатика и информатизация образования. 2025. № 1 (71). С. 21–35. [Kapterev A.I. The role of human-computer interaction in instructional design. *The Academic Journal of Moscow City University. Series: Informatics and Informatization of Education*. 2025. No. 1 (71). Pp. 21–35. (In Rus.)]

Курганова, Лапчик, 2024 – Курганова Н.А., Лапчик Е.С. Приемы разработки учебных заданий педагогами высшей школы с помощью нейросети // Горизонты образования: материалы V Международной научно-практической конференции, Омск, 25–26 апреля 2024 года. Омск, 2024. С. 187–190. [Kurganova N.A., Lapchik E.S. Techniques for developing educational assignments for higher education teachers using neural networks. *Gorizonty obrazovaniya: materialy V Mezhdunarodnoj nauchno-prakticheskoy konferencii, Omsk, 25–26 aprelya 2024 goda*. Omsk, 2024. Pp. 187–190. (In Rus.)]

Резаев, 2023 – Резаев А.В. Искусственный интеллект в университетах: какое будущее нам ожидать? // Высшее образование в России. 2023. Т. 32. № 6. С. 19–37. [Rezaev A.V. Artificial intelligence in universities: What future can we expect? *Vysshee obrazovanie v Rossii*. 2023. Vol. 32. No. 6. Pp. 19–37. (In Rus.)]

Сысоев, 2025 – Сысоев П.В. Компетенция современного педагога в области искусственного интеллекта: структура и содержание // Высшее образование в России. 2025. Т. 34. № 6. С. 58–79. [Sysoev P.V. Competence of a modern teacher in the field of artificial intelligence: Structure and content. *Vysshee obrazovanie v Rossii*. 2025. Vol. 34. No. 6. Pp. 58–79. (In Rus.)]

Титова, Харламенко, 2025 – Титова С.В., Харламенко И.В. Подготовка педагогических кадров к использованию искусственного интеллекта в обучении иностранным языкам // Вестник Московского университета. Серия 19: Лингвистика и межкультурная коммуникация. 2025. № 1. С. 45–60. [Titova S.V., Harlamenko I.V. Training teachers to use artificial intelligence in teaching foreign languages. *Vestnik Moskovskogo universiteta. Seriya 19: Lingvistika i mezhkulturnaya kommunikaciya*. 2025. No. 1. Pp. 45–60. (In Rus.)]

Caines et al., 2023 – Caines A., Benedetto L., Taslimipoor S. et al. On the application of Large Language Models for language teaching and assessment technology. *arXiv*. 2023. arXiv: 2307.08393v1.

Chan, Tsi, 2024 – Chan C.K.Y., Tsi L.H.Y. Will generative AI replace teachers in higher education? A study of teacher and student perceptions. *Studies in Educational Evaluation*. 2024. Vol. 83. Article number 101395.

Cherednichenko et al., 2024 – Cherednichenko O., Yanholenko O., Badan A. et al. Large language models for foreign language acquisition. *CLW-2024: Computational Linguistics Workshop at 8th International Conference on Computational Linguistics and Intelligent Systems (CoLInS-2024)*. 2024. Pp. 101–130.

Fernández Jiménez, 2024 – Fernández Jiménez A. Integration of AI helping teachers in traditional teaching roles. *European Public & Social Innovation Review*. 2024. Vol. 9. Pp. 01–17.

Huang et al., 2025 – Huang Q., Lv C., Lu L., Tu S. Evaluating the quality of AI-generated digital educational resources for university teaching and learning. *Systems*. 2025. Vol. 13. Pp. 1--18.

Karataş et al., 2024 – Karataş İ., Yıldırım S., Kovanović V. Reshaping curriculum adaptation in the age of artificial intelligence: Mapping teacher practices and emerging challenges. *British Educational Research Journal*. 2024. Vol. 51. No. 1. Pp. 154–180. DOI: 10.1002/berj.4068

Kucuk et al., 2022 – Kucuk S., Yılmaz R., Sisman M. A systematic review of artificial intelligence studies in teacher education. *Education and Information Technologies*. 2022. Vol. 27. Pp. 8347–8374.

Mollick, Mollick, 2023 – Mollick E., Mollick L. Using AI to implement effective teaching strategies in classrooms: Five strategies, including prompts. Wharton School of the University of Pennsylvania & Wharton Interactive. March 16, 2023. URL: <https://ssrn.com/abstract=4391243>

Tan, Cheng, Ling, 2025 – Tan X., Cheng G., Ling M.H. Artificial intelligence in teaching and teacher professional development: A systematic review. *Computers and Education: Artificial Intelligence*. 2025. Vol. 8.

Sharma et al., 2025 – Sharma S., Mittal P., Kumar M., Bhardwaj V. The role of large language models in personalized learning: A systematic review of educational impact. *Discover Sustainability*. 2025. Vol. 6. Pp. 243.

Zhang, Zhang, 2024 – Zhang Y., Zhang L. AI in teacher education: Unlocking new dimensions in teaching support. *Journal of Computer Assisted Learning*. 2024. Vol. 40 (4). Pp. 1871–1885.

Статья поступила в редакцию 01.11.2025

The article was received 01.11.2025

Сведения об авторе / About the author

Авраменко Анна Петровна – кандидат педагогических наук; докторант кафедры лингводидактики и современных технологий иноязычного образования Института иностранных языков, Московский педагогический государственный университет; доцент кафедры лингвистики и информационных технологий факультета иностранных языков и регионоведения, Московский государственный университет имени М.В. Ломоносова

Anna P. Avramenko – PhD in Pedagogy; Dr. Candidate at the Department of Lingvodidactics and Modern Technologies of Foreign Language Education of the Institute of Foreign Languages, Moscow Pedagogical State University; Associate Professor at the Department of Linguistics and Information Technology of the Faculty of Foreign Languages and Regional Studies, Lomonosov Moscow State University

ORCID: <https://orcid.org/0009-0009-3004-2192>

E-mail: avram4ik@gmail.com