

КОНЦЕПТУАЛЬНОЕ ПРОЕКТИРОВАНИЕ ИНТЕЛЛЕКТУАЛЬНОЙ СИСТЕМЫ ИЗУЧЕНИЯ АНГЛИЙСКОГО ЯЗЫКА В ВУЗЕ¹

Михалев Антон Сергеевич¹,

e-mail: asmikhalev@yandex.ru

Абдуллаев Нурдоолот¹,

e-mail: abdullaevnurdoolot3@gmail.com

Соболева Елена Владимировна¹,

e-mail: e.v.soboleva.mail@gmail.ru

¹Институт космических и информационных технологий
Сибирского федерального университета, г. Красноярск, Россия

В статье рассматривается концептуальное проектирование интеллектуальной системы изучения английского языка, основанной на использовании больших языковых моделей и ориентированной на поддержку преподавателя английского языка в вузе. Актуальность исследования обусловлена ростом роли английского языка в академической и профессиональной коммуникации, а также активным развитием цифровых образовательных технологий. Предлагаемая система рассматривается как ИИ-ассистент преподавателя, способный поддерживать автоматизированную проверку письменных работ, выявление и комментирование языковых ошибок, формирование персонализированной обратной связи, а также подготовку и модерирование учебной дискуссии. В работе представлен обзор современных исследований в области технологии образования. На этой основе предложена концептуальная модель интеллектуальной системы, включающая диагностический, оценочный, диалоговый и аналитический модули. Отдельное внимание уделено описанию вычислительного эксперимента, направленного на экспериментальную проверку эффективности ИИ-ассистента при решении задач преподавателя английского языка в вузе. Показано, что практическая ценность подобных систем определяется не только качеством генерации текста, но и педагогической корректностью обратной связи, степенью управляемости модели и ее интегрируемостью в реальный образовательный процесс.

Ключевые слова: образовательные технологии, интеллектуальная система обучения, английский язык, большие языковые модели, ИИ-ассистент преподавателя, автоматизированная проверка письменных работ, персонализированная обратная связь, учебная дискуссия

CONCEPTUAL DESIGN OF AN INTELLIGENT ENGLISH LANGUAGE LEARNING SYSTEM IN HIGHER EDUCATION

Mikhalev A.S.¹,

e-mail: asmikhalev@yandex.ru

Abdullaev N.¹,

e-mail: abdullaevnurdoolot3@gmail.com

Soboleva E.V.¹,

e-mail: e.v.soboleva.mail@gmail.ru

¹Institute of Space and Information Technology, Siberian Federal University, Krasnoyarsk, Russia

The article examines the conceptual design of an intelligent English language learning system based on the use of large language models and focused on the support of an English teacher at a university. The relevance of the research is due to the growing role of the English language in academic and professional communication, as

¹ Исследование выполнено за счет гранта Российского научного фонда № 26-21-20004 (<https://rscf.ru/project/26-21-20004/>) и за счет целевого финансирования (гранта) Красноярского краевого фонда науки.

well as the active development of digital educational technologies. The proposed system is considered as an AI teaching assistant capable of supporting automated verification of written works, identification and commenting on language errors, the formation of personalized feedback, as well as the preparation and moderation of educational discussions. The article provides an overview of modern research in the field of education technology. On this basis, a conceptual model of an intelligent system is proposed, including diagnostic, evaluative, dialogical and analytical modules. Special attention is paid to the description of a computational experiment aimed at experimentally verifying the effectiveness of an AI assistant in solving the tasks of an English teacher at a university. It is shown that the practical value of such systems is determined not only by the quality of text generation, but also by the pedagogical correctness of feedback, the degree of controllability of the model and its integrability into the real educational process.

Keywords: educational technologies, intelligent learning system, English, large language models, AI teaching assistant, automated verification of written papers, personalized feedback, educational discussion

Введение

Изучение иностранных языков в условиях глобализации и цифровизации сохраняет высокую значимость как фактор образовательного, научного и профессионального развития. Владение английским языком обеспечивает доступ к международным публикациям, профессиональным сообществам и академическим коммуникациям, а также повышает конкурентоспособность специалистов на глобальном рынке труда [1]. Одновременно развивается рынок образовательных технологий, в котором заметную роль играют онлайн-платформы, мобильные приложения и системы адаптивного обучения [2].

Традиционные формы обучения иностранному языку сегодня не исчезают, но все чаще дополняются цифровыми средствами. Однако распространенные языковые платформы, несмотря на удобство, геймификацию и массовую доступность, не всегда обеспечивают глубокую персонализацию, содержательно богатую обратную связь и полноценную имитацию естественной языковой среды. Систематические обзоры по ICALL и AI-enhanced language education показывают, что автоматическая обратная связь и персонализация действительно усиливают обучение, но целостная педагогическая интеграция и адаптация к индивидуальной траектории остаются недостаточно проработанными [3].

Ситуация существенно изменилась с появлением больших языковых моделей (LLM). В отличие от более ранних решений, основанных на сценариях или ограниченных диалоговых механизмах, LLM способны вести развернутый диалог, генерировать задания, адаптировать объяснения под уровень обучающегося, оценивать ответы и создавать персонализированную обратную связь. В обзорах по ChatGPT в ESL/EFL-образовании отмечается, что основные направления применения связаны с письмом, подготовкой учебных материалов, поддержкой самостоятельного обучения и частично разговорной практикой, тогда как устойчивые модели долгосрочной педагогической интеграции еще находятся в стадии формирования [4].

В современной лингводидактике все более актуальным становится вопрос о трансформации роли преподавателя в условиях внедрения технологий искусственного интеллекта. ИИ в этом контексте следует рассматривать не как замену преподавателя, а как интеллектуальный инструмент, способный взять на себя часть рутинных операций: первичную проверку работ, формирование обратной связи, подготовку заданий и сопровождение учебного взаимодействия. В российской литературе также подчеркивается, что подобные технологии требуют педагогического контроля, критического осмысления и сочетания с традиционными методами обучения [5].

Цель статьи заключается в разработке концептуальной модели интеллектуальной системы изучения английского языка, включающей ИИ-ассистента для поддержки профессиональной деятельности преподавателя английского языка в вузе, а также в экспериментальной оценке эффективности данного модуля при автоматизированной проверке письменных работ, формировании персонализированной обратной связи и сопровождении учебной дискуссии.

1. Обзор научных трудов по теме исследования

Автоматизация языкового обучения развивалась от систем компьютерной поддержки обучения иностранным языкам (CALL) к интеллектуальным системам ICALL. Систематический обзор Weng и Chiu [3], включивший 83 исследования, показывает, что ICALL-среды наиболее успешно реализуют автоматическую обратную связь, интеллектуальное тьюторство и персонализацию. Вместе с тем принципы проблемно-ориентированного обучения и интеграции знаний в таких системах представлены слабее, чем принципы демонстрации учебного материала и практического применения, что указывает на технологическую зрелость отдельных функций при недостаточной дидактической целостности системы.

Следующий этап развития систем связан с более широким внедрением искусственного интеллекта в языковое образование. Обзор Zheng и Yang [6] показывает, что в последние годы наблюдается заметная активизация исследований в области языкового образования с применением искусственного интеллекта. Вместе с тем авторы отмечают, что потенциал индивидуализации обучения реализуется не в полной мере, а уровень человеко-машинного взаимодействия во многих работах остаётся ограниченным. Это особенно важно для проектирования систем, ориентированных не просто на выдачу заданий, а на длительное сопровождение учебной траектории и поддержку преподавателя в принятии педагогических решений.

Значимое место в этой эволюции заняли чат-боты и разговорные агенты. Систематический обзор Du и Daniel [7] по применению чат-ботов на основе искусственного интеллекта для развития навыков устной речи при изучении английского языка как иностранного показал, что такие решения могут способствовать улучшению результатов обучения говорению, повышению уверенности, вовлеченности и мотивации обучающихся, а также в отдельных случаях снижению языковой тревожности. При этом авторы подчеркивают, что область все еще находится на ранней стадии развития и требует дальнейших исследований, особенно в части методически выверенного дизайна и устойчивых педагогических сценариев.

Появление ChatGPT и других LLM сформировало новый этап развития интеллектуальных языковых систем. Систематический обзор Lo и соавторов [4] по 70 эмпирическим исследованиям в ESL/EFL показывает, что ChatGPT используется для поддержки письма, генерации материалов, персонализированного обучения и помощи преподавателю, но одновременно вызывает опасения, связанные с некорректной информацией, академической недобросовестностью и вопросами конфиденциальности. Авторы также отмечают, что такие виды речевой деятельности, как чтение, аудирование и говорение, на данный момент исследованы в меньшей степени, чем письмо.

Сходные выводы представлены в обзоре Li и соавторов [8], анализирующем публикации первого года, посвящённые применению ChatGPT в обучении языкам. Авторы подчёркивают междисциплинарный характер данной области, доминирование эмпирических исследований и выраженный фокус на высшем образовании. В то же время отмечается необходимость дальнейшего изучения качества предоставляемой обратной связи, восприятия технологии участниками образовательного процесса, а также её долгосрочных эффектов.

В российских научных статьях также подчеркивается, что искусственный интеллект следует рассматривать не как замену преподавателя, а как инструмент интенсификации обучения, персонализации и передачи части рутинных задач автоматизированным системам. В работах Е.М. Михайловой [9] отмечается потенциал ИИ для мгновенной обратной связи, экономии времени преподавателя и построения индивидуальных траекторий обучения, однако одновременно подчеркивается необходимость педагогического контроля и критического анализа результатов работы ИИ.

Отдельного внимания заслуживают исследования, ориентированные на поддержку деятельности преподавателя. В публикации о нейросетевом ассистенте ЛИРА [10], представленной МГИМО, описана апробация системы с участием 52 преподавателей английского языка, что подтверждает практический интерес к ИИ-ассистентам в профессиональной деятельности преподавателя. Этот факт полезен для понимания того, что интеллектуальные языковые системы могут проектироваться не только как средство обучения студента, но и как средство усиления педагогической работы.

Таким образом, литературный обзор показывает, что современное поле языкового EdTech смещается в сторону систем, сочетающих генеративный ИИ, диалоговое взаимодействие, автоматический

анализ ответов и поддержку преподавателя. При этом сравнительно слабо проработан вопрос о том, каким образом ИИ-ассистент может быть встроен в повседневную профессиональную деятельность преподавателя английского языка в вузе как инструмент проверки письменных работ, подготовки обратной связи и модерирования учебной коммуникации.

2. Концептуальная модель интеллектуальной системы

Общая архитектура системы

Предлагаемая интеллектуальная система изучения английского языка рассматривается как ИИ-ассистент преподавателя и включает следующие основные модули (рисунок 1):

1. *Диагностический модуль*. Определяет уровень подготовленности обучающихся, типичные ошибки, проблемные языковые темы и особенности конкретной учебной группы.
2. *Оценочный модуль*. Выполняет автоматизированный анализ письменных работ обучающихся, выявляет ошибки, формирует предварительный вердикт и подготавливает обратную связь для преподавателя.
3. *Диалоговый модуль*. Поддерживает сценарии учебного взаимодействия, включая проведение дискуссии на английском языке, генерацию уточняющих вопросов, модерирование ответов и подведение промежуточных итогов.
4. *Аналитический модуль*. Накапливает данные о типичных ошибках, качестве выполнения заданий, динамике прогресса обучающихся и предоставляет преподавателю интерпретируемую статистику.
5. *Интерфейс педагогического контроля*. Обеспечивает возможность проверки, корректировки и утверждения результатов, сгенерированных системой, до их передачи студенту.

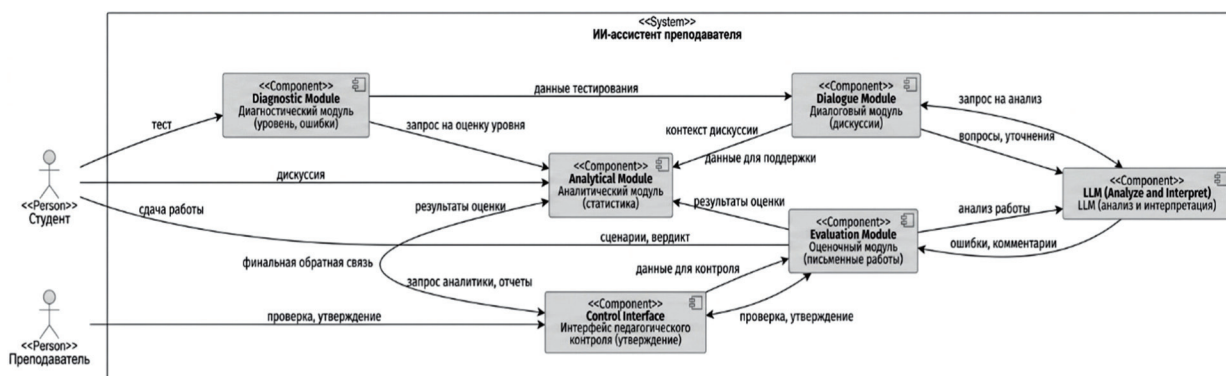


Рисунок 1 – Архитектура интеллектуальной системы¹

Использование LLM в системе

Большие языковые модели в рамках предлагаемой системы выполняют несколько задач.

Во-первых, они поддерживают персонализированное объяснение: одно и то же правило может быть объяснено в разных форматах, с разной степенью подробности и с учетом типичных затруднений конкретной группы студентов.

Во-вторых, LLM выполняют функцию анализатора ответа, выявляя ошибки и формируя содержательную обратную связь. Именно эта функция особенно важна в профессиональной деятельности преподавателя, поскольку позволяет сократить время на первичный анализ письменных работ и сосредоточиться на содержательной педагогической интерпретации.

В-третьих, они обеспечивают поддержку учебной дискуссии, включая генерацию вопросов, уточняющих реплик, промежуточных выводов и структурирование аргументов участников.

¹ Составлено автором.

Сценарии использования

Функциональные возможности предлагаемой интеллектуальной системы проявляются через набор типовых сценариев использования, отражающих ключевые задачи преподавателя английского языка в вузе. Данные сценарии реализуются на основе взаимодействия функциональных модулей системы и большой языковой модели, выступающей в качестве центрального интеллектуального компонента, обеспечивающего анализ ответов обучающихся, выявление и типизацию ошибок, а также формирование педагогически интерпретируемой обратной связи.

Проверка письменной работы

В данном сценарии система используется для автоматизированной первичной проверки письменных работ студентов (рисунок 2). Обучающийся передает текст на английском языке, после чего система выполняет его анализ с использованием большой языковой модели. В ходе анализа выявляются грамматические, лексические и смысловые ошибки, определяется их тип и значимость.

На основе полученных результатов формируется черновой вариант комментариев, который передается преподавателю через интерфейс педагогического контроля. Преподаватель оценивает предложенные системой замечания, при необходимости корректирует их и принимает итоговое решение.

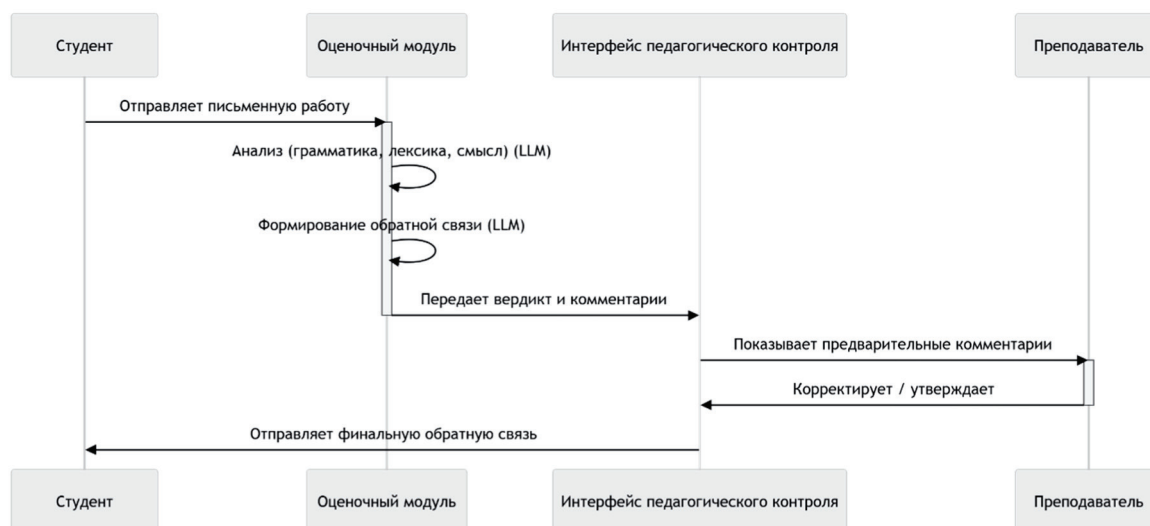


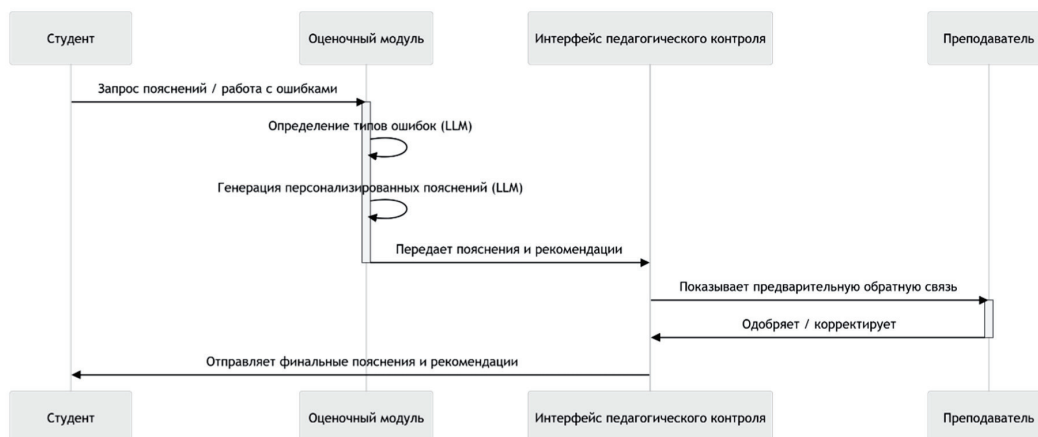
Рисунок 2 – Проверка письменной работы²

Формирование персонализированной обратной связи

На основе выявленных ошибок система с использованием большой языковой модели генерирует пояснения, примеры и рекомендации, адаптированные к уровню студента и характеру допущенных ошибок (рисунок 3).

Особенностью данного сценария является вариативность представления материала: одно и то же правило может быть объяснено с разной степенью детализации, с использованием различных примеров и с учетом типичных затруднений обучающегося. Полученная обратная связь проходит этап педагогического контроля и затем передается студенту.

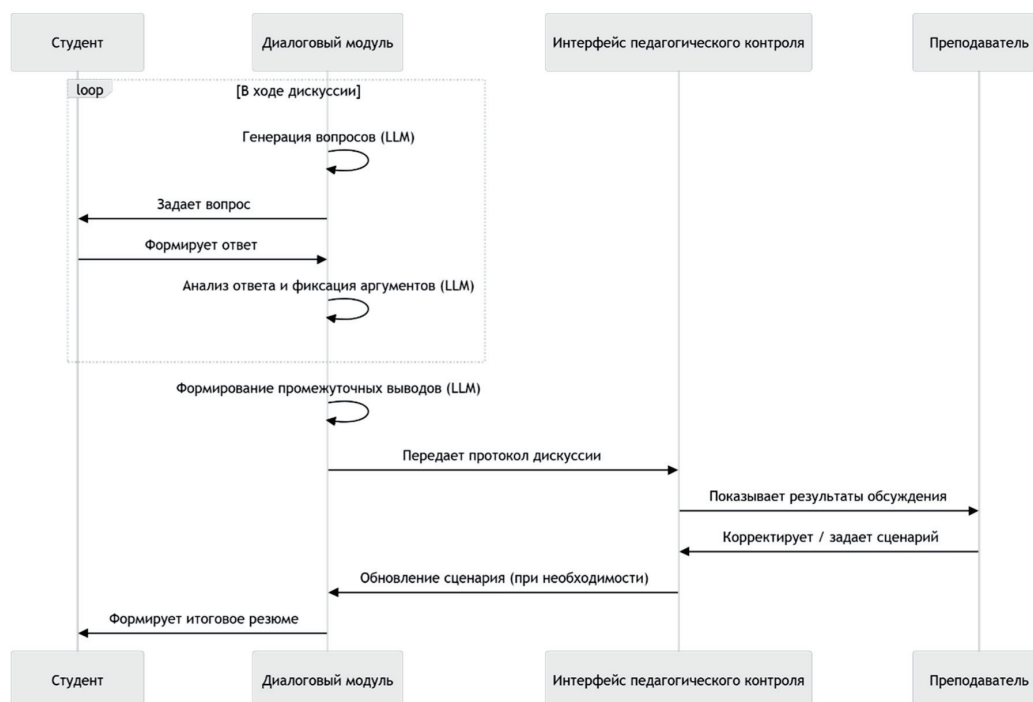
² Составлено автором.

Рисунок 3 – Формирование персонализированной обратной связи³*Модерирование дискуссии*

В рамках данного сценария система выступает в роли интеллектуального модератора учебной дискуссии (рисунок 4). Диалоговый модуль с использованием большой языковой модели генерирует вопросы, поддерживает развитие обсуждения, формирует уточняющие реплики и способствует вовлечению обучающихся в коммуникацию на английском языке.

В процессе взаимодействия система анализирует ответы студентов, выделяет ключевые аргументы и отслеживает логическую структуру дискуссии. При необходимости формируются промежуточные выводы, позволяющие структурировать обсуждение и удерживать его в рамках заданной темы.

Преподаватель сохраняет возможность вмешательства в ход дискуссии, корректируя сценарий взаимодействия и задавая педагогические ориентиры. По завершении обсуждения система формирует итоговое резюме, отражающее основные выводы и позиции участников.

Рисунок 4 – Модерирование дискуссии⁴³ Составлено автором.⁴ Составлено автором.

Анализ типичных ошибок группы

Данный сценарий ориентирован на поддержку преподавателя в принятии педагогических решений на уровне группы (рисунок 5). Аналитический модуль агрегирует данные, поступающие из диагностического, оценочного и диалогового модулей, включая информацию об ошибках, темах обсуждений и результатах выполнения заданий.

С использованием большой языковой модели выполняется выявление закономерностей, определяются наиболее частые ошибки, слабые темы и динамика прогресса обучающихся. Результаты представляются преподавателю в виде интерпретируемой аналитики.

На основе полученной информации преподаватель может корректировать учебный процесс, изменять содержание занятий, акцентировать внимание на проблемных темах и адаптировать образовательную траекторию группы.

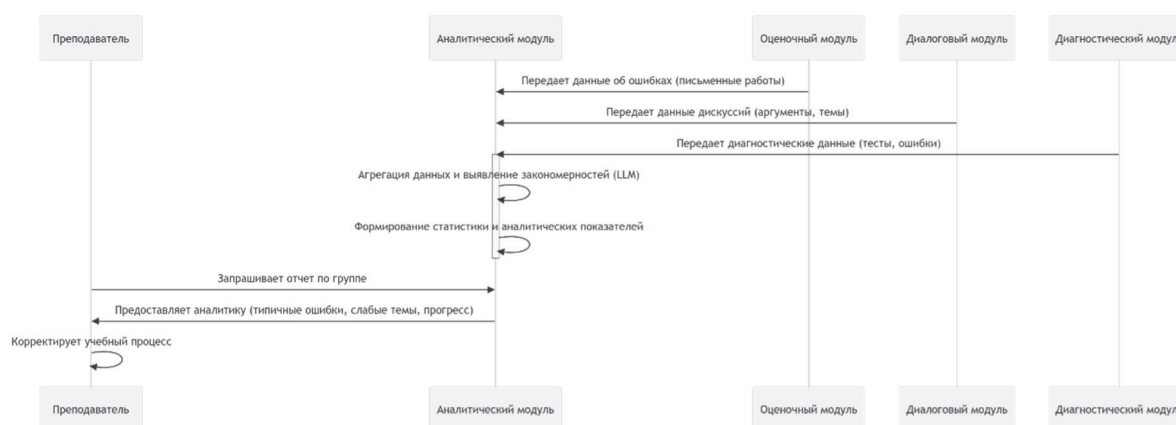


Рисунок 5 – Анализ типичных ошибок группы⁵

3. Экспериментальная проверка реализуемости предложенной системы

Предложенная концептуальная модель интеллектуальной системы изучения английского языка описывает архитектуру и основные сценарии использования, ориентированные на поддержку профессиональной деятельности преподавателя. Однако сама по себе концептуальная проработка не гарантирует практической реализуемости системы, поскольку ее ключевые функции опираются на возможности больших языковых моделей.

В связи с этим возникает необходимость экспериментальной проверки того, в какой степени современные большие языковые модели способны обеспечивать выполнение критически важных задач, лежащих в основе предложенной архитектуры.

С учетом представленных сценариев использования (рисунки 2–5), экспериментальная часть работы направлена на оценку реализуемости ее ключевых функциональных компонентов, непосредственно связанных с применением ИИ-ассистента в профессиональной деятельности преподавателя английского языка в высшей школе.

Для достижения поставленной цели в рамках эксперимента решаются следующие задачи:

- 1) выделить ключевые функции ИИ-ассистента, подлежащие экспериментальной проверке в рамках предложенной интеллектуальной системы изучения английского языка;
- 2) оценить возможности больших языковых моделей при автоматизированном анализе письменных работ студентов на английском языке, включая выявление грамматических, лексических и смысловых ошибок, а также формирование предварительных комментариев;
- 3) сопоставить результаты работы ИИ-ассистента с экспертной оценкой преподавателя по критериям полноты выявления ошибок, точности комментариев, педагогической полезности обратной связи и необходимости ручной корректировки;

⁵ Составлено автором.

4) сформулировать выводы о реализуемости ключевых функций предложенной интеллектуальной системы, а также об условиях и ограничениях ее интеграции в профессиональную деятельность преподавателя английского языка в вузе.

Организация эксперимента

Экспериментальная схема предусматривает оценку ИИ-ассистента по двум основным направлениям: автоматизированная проверка письменных заданий и поддержка учебной дискуссии.

В рамках первого направления в качестве входных данных используются письменные работы студентов, выполненные на английском языке. ИИ-ассистент осуществляет автоматизированный анализ ответов, выявляет языковые ошибки, формирует комментарии и предлагает вариант персонализированной обратной связи. Полученные результаты затем сопоставляются с экспертной оценкой преподавателя, что позволяет оценить качество работы системы по заданным критериям.

В рамках второго направления ИИ-ассистент рассматривается как инструмент сопровождения и модерирования учебной дискуссии по заданной теме. В данном сценарии оценивается способность системы формулировать содержательные вопросы, поддерживать логическое развитие обсуждения, выделять ключевые аргументы участников и подводить итоги дискуссии с учетом педагогической цели занятия.

Материалы и процедура оценки

Для проверки письменных заданий в эксперименте могут использоваться тексты студентов, содержащие грамматические, лексические и смысловые ошибки различного типа. Каждая работа проходит двойную оценку: со стороны преподавателя и со стороны ИИ-ассистента. Затем проводится сопоставление результатов по нескольким параметрам:

- полнота выявления ошибок;
- точность выявления ошибок;
- полезность комментариев;
- педагогическая приемлемость обратной связи;
- необходимость дополнительной ручной корректировки.

Для сценария учебной дискуссии оценивание может проводиться по экспертной шкале, где преподаватель фиксирует качество вопросов, логичность структуры обсуждения, полезность итоговых формулировок и общую пригодность ИИ-ассистента для использования на занятии.

Метрики качества

Для оценки функциональных возможностей интеллектуальной системы могут использоваться различные группы показателей, отражающие как качество выполнения отдельных педагогических задач, так и практическую применимость системы в образовательном процессе. К таким показателям относятся снижение временной нагрузки на преподавателя, корректность выявления ошибок, полнота и педагогическая полезность формируемой обратной связи, качество сопровождения учебной дискуссии, а также удовлетворенность пользователей результатами работы системы.

Однако для экспериментальной оценки одного из ключевых модулей – автоматической проверки письменных ответов – целесообразно использовать специальный набор метрик.

На первом уровне задача формулируется как бинарная классификация: модель должна определить, является ли ответ корректным (Correct) или требует исправлений (Needs Revision). В этом случае используются стандартные метрики:

$$Accuracy(acc) = \frac{(TP + TN)}{(TP + TN + FP + FN)} \cdot$$

$$Precision(P) = \frac{TP}{(TP + FP)} \cdot$$

$$Recall(R) = \frac{TP}{(TP + FN)} \cdot$$

$$F1 = \frac{2 * Precision * Recall}{(Precision + Recall)} \cdot$$

На втором уровне оценивается качество выявления конкретных ошибок. Здесь каждая ошибка из экспертной разметки рассматривается как отдельная единица анализа, а затем вычисляются:

$$Precision_{error}(P_{err}) = \frac{N_{correct_detected}}{N_{reported_by_model}}$$

$$Recall_{error}(R_{err}) = \frac{N_{correct_detected}}{N_{reference_errors}}$$

$F1_{error}$ – гармоническое среднее между $Precision_{error}$ и $Recall_{error}$.

Дополнительно вводится коэффициент избыточной коррекции:

$$Overcorrection_{rate}(OC_{rate}) = \frac{N_{false_errors}}{N_{reported_by_model}}$$

Данный коэффициент отражает склонность модели считать ошибками допустимые стилистические варианты и синонимические конструкции. Для образовательной системы это критически важно, поскольку избыточная коррекция снижает доверие к обратной связи и искажает ее педагогическую ценность.

Еще один практический показатель – Format compliance (FC), то есть доля ответов, которые полностью соответствуют заданному JSON-формату. Эта метрика важна для интеграции LLM в реальную программную платформу.

Для оценки сценария дискуссии могут дополнительно использоваться качественные критерии:

- релевантность генерируемых вопросов теме обсуждения;
- логичность переходов между репликами;
- полезность промежуточных выводов;
- корректность выделения ключевых аргументов;
- субъективная оценка преподавателя.

Характеристика эксперимента

Для оценки эффективности ИИ-ассистента в профессиональной деятельности преподавателя английского языка в вузе был сформирован экспериментальный набор данных, включающий письменные работы студентов на английском языке, а также материалы для апробации сценария учебной дискуссии.

Корпус письменных работ включал тексты, содержащие различные типы ошибок, в том числе грамматические, лексические, синтаксические, орфографические и смысловые нарушения. Для каждой работы была подготовлена экспертная разметка, выполненная преподавателем английского языка. Разметка включала итоговый вердикт о качестве работы, указание на ошибки, их типологическую принадлежность, а также комментарии и рекомендации по исправлению.

Следует отметить, что письменные работы студентов представлялись в электронном виде. В связи с этим технические опечатки и орфографические ошибки могли быть устранены обучающимися с использованием встроенных средств проверки орфографии в текстовых редакторах и браузерах. Поэтому в анализируемом корпусе орфографические ошибки не были выделены как значимая самостоятельная категория ошибок; основное внимание в разметке уделялось грамматическим, лексическим, синтаксическим и смысловым нарушениям.

Для оценки возможностей ИИ-ассистента в сценарии поддержки учебной дискуссии использовался набор тем, соответствующих учебной программе дисциплины. Для каждой темы оценивались способность системы генерировать содержательные вопросы, поддерживать логическую структуру обсуждения, фиксировать ключевые аргументы и формулировать итоговые выводы.

Основные характеристики экспериментального набора данных представлены в таблице 1.

Таблица 1 – Характеристика данных для эксперимента⁶

Параметр	Значение
Общее число письменных работ	210
Средний объем одной работы, слов	22
Минимальный объем работы, слов	6
Максимальный объем работы, слов	60

⁶ Составлено автором.

Общее число размеченных ошибок	140
Число грамматических ошибок	68
Число лексических ошибок	18
Число синтаксических ошибок	12
Число орфографических ошибок	0
Число смысловых ошибок	42
Количество тем для дискуссии	1
Среднее число участников одной дискуссии	1

Характеристика исследуемых моделей

В рамках эксперимента оценивались несколько больших языковых моделей, рассматриваемых в качестве основы ИИ-ассистента преподавателя. Отбор моделей осуществлялся с учетом их доступности, распространенности, способности работать в диалоговом режиме и пригодности для решения задач анализа текста, генерации комментариев и сопровождения учебного взаимодействия.

Каждая модель тестировалась в идентичных условиях на одном и том же наборе входных данных. Для повышения сопоставимости результатов использовалась единая инструкция, задающая формат ответа, ограничения на характер исправлений и требования к структуре обратной связи.

Основные сведения об исследуемых моделях приведены в таблице 2.

Таблица 2 – Характеристика исследуемых моделей⁷

Модель	Число параметров	Формат использования	Основное назначение в эксперименте
Openchat-3_5-0106	7B	Локально	Проверка письменных работ
MiniMax-M2.5	229B	HuggingFace API	Проверка письменных работ
Qwen2.5-7B-Instruct	8B	HuggingFace API	Проверка письменных работ
Meta-Llama-3-8B-Instruct	8B	HuggingFace API	Проверка письменных работ, комментарии
Gemma-3-12b-it	12B	HuggingFace API	Проверка письменных работ, генерация обратной связи

Результаты оценки по метрикам

Для оценки качества работы ИИ-ассистента при автоматизированной проверке письменных работ использовались метрики двух уровней: метрики итогового вердикта и метрики обнаружения конкретных ошибок. Дополнительно учитывались коэффициент избыточной коррекции и показатель соблюдения формата ответа. Для сценария учебной дискуссии использовались экспертные оценки по ряду качественных критериев.

Результаты сравнения моделей по основным метрикам представлены в таблице 3.

Таблица 3 – Результаты оценки моделей при проверке письменных работ⁸

Модель	Acc	P	R	$F1$	P_{err}	R_{err}	$F1_{err}$	OC_{rate}	FC
Openchat-3_5-0106	0.84	0.71	0.80	0.73	0.88	0.88	0.88	11.5 %	100 %
MiniMax-M2.5	0.81	0.71	0.63	0.64	0.82	0.92	0.87	18.4 %	100 %
Qwen2.5-7B-Instruct	0.79	0.61	0.40	0.47	0.75	0.98	0.85	24.6 %	100 %
Meta-Llama-3-8B-Instruct	0.79	0.59	0.43	0.46	0.76	0.96	0.85	23.7 %	100 %
Gemma-3-12b-it	0.77	0.85	0.50	0.51	0.68	0.98	0.80	32.6 %	100 %

Результаты оценки сценария учебной дискуссии

Для оценки возможностей ИИ-ассистента в сценарии модерирования учебной дискуссии использовались экспертные критерии, отражающие качество генерируемых вопросов, логичность сопровождения обсуждения, полноту выделения ключевых аргументов и полезность итоговых выводов.

⁷ Составлено автором.

⁸ Составлено автором.

Оценка «высокая» присваивалась в случае, если ответ ИИ-ассистента полностью соответствовал теме и педагогической цели дискуссии, обеспечивал логичное развитие обсуждения, корректно выделял основные аргументы участников и формировал содержательные итоговые выводы, пригодные для использования преподавателем без существенной доработки.

Оценка «средняя» присваивалась в случае частичного соответствия заданному критерию: система в целом сохраняла связь с темой и учебной задачей, однако допускала отдельные обобщенные формулировки, неполное выделение аргументов, недостаточно последовательные переходы между репликами или требовала дополнительной педагогической корректировки.

Оценка «низкая» присваивалась в случае, если результат работы ИИ-ассистента был слабо связан с темой дискуссии, содержал существенные логические разрывы, не фиксировал ключевые аргументы участников либо формировал формальные и педагогически малополезные выводы.

Результаты экспертной оценки представлены в таблице 4.

Таблица 4 – Экспертная оценка ИИ-ассистента в сценарии учебной дискуссии⁹

Модель	Релевантность вопросов	Логичность сопровождения дискуссии	Полнота выделения аргументов	Качество итоговых выводов	Общая полезность
Openchat-3_5-0106	Высокая	Средняя	Средняя	Среднее	Средняя
MiniMax-M2.5	Высокая	Высокая	Высокая	Высокое	Высокая
Qwen2.5-7B-Instruct	Средняя	Средняя	Высокая	Среднее	Средняя
Meta-Llama-3-8B-Instruct	Высокая	Высокая	Высокая	Высокое	Высокая
Gemma-3-12b-it	Средняя	Низкая	Низкая	Низкое	Низкая

Интерпретация результатов

В ходе исследования установлено, что интегральный показатель точности (*Accuracy*) варьируется в диапазоне от 0.77 до 0.84, что подтверждает общую пригодность современных LLM для автоматизации проверки учебных работ.

Наиболее сбалансированные результаты продемонстрировала модель **Openchat-3_5-0106**. Данный результат свидетельствует о высокой степени когнитивной гибкости модели при верификации студенческих ответов. В отличие от нее, модели **Qwen2.5-7B** и **Llama-3-8B** характеризуются выраженной «метрической строгостью»: при высоких показателях точности (*Precision*) их способность признавать корректные, но вариативные ответы (*Recall*) значительно снижена. Это указывает на риск повышения ложноотрицательных оценок в реальном образовательном процессе.

Оценка диагностического модуля системы проводилась через призму детекции языковых и смысловых ошибок.

– Чувствительность к ошибкам ($Recall_{error}$). Экспериментально подтверждено, что модели **Qwen2.5-7B** и **Gemma-3-12b-it** обладают максимальной чувствительностью (0.98), идентифицируя практически весь спектр отклонений от эталона.

– Гармоническая точность ($F1_{error}$). Лидером по совокупному качеству обратной связи является **Openchat-3_5-0106** (0.88), что коррелирует с требованиями педагогической корректности, изложенными в концептуальной модели системы.

Критически важным параметром для ИИ-ассистента является уровень избыточной коррекции ($Overcorrection_{rate}$). По результатам эксперимента модель **Openchat-3.5-0106** продемонстрировала наименьшее значение данного показателя, что свидетельствует о более сдержанном характере формируемой обратной связи и снижает риск демотивации обучающихся вследствие избыточного количества замечаний.

В то же время модель **Gemma-3-12B-it** характеризовалась более высоким уровнем избыточной коррекции. В контексте проектирования интеллектуальной системы изучения английского языка это может рассматриваться как ограничение, поскольку чрезмерное количество замечаний способно повышать когнитивную нагрузку, снижать воспринимаемую полезность обратной связи и вызывать фрустрацию у студентов.

⁹ Составлено автором.

Анализ диалогового модуля показал, что для глубокой проработки аргументации в учебной дискуссии наиболее перспективной является модель **MiniMax-M2.5**. Она обеспечивает максимальную полноту выделения ключевых аргументов и демонстрирует высокое качество итоговых выводов, сопровождая их лингвистическими комментариями (например, исправление грамматических форм в контексте обсуждения).

Подводя итог, можно констатировать, что практическая ценность интеллектуальной системы определяется балансом между строгостью детекции и лояльностью к вариативности естественного языка.

1. Для базовых уровней обучения и массовой автоматизированной проверки рекомендуется использование **Openchat-3_5-0106** как системы с наилучшими показателями педагогической корректности.

2. Для продвинутых уровней, требующих глубокого аналитического сопровождения и детальной диагностики ошибок, целесообразна интеграция **MiniMax-M2.5** или **Llama-3-8B** (при условии дополнительной настройки промптов на снижение стилистического формализма).

Данные результаты подтверждают выдвинутую в статье гипотезу о том, что эффективность ИИ-ассистента детерминирована не только объемом параметров модели, но и её способностью к адекватному интерпретированию образовательного контекста.

Заключение

Проведенный анализ показывает, что современный этап развития интеллектуальных систем языкового обучения характеризуется переходом от частично адаптивных ICALL-сред к генеративным системам, опирающимся на большие языковые модели. Такие модели создают новые возможности для автоматизированной проверки письменных работ, персонализированной обратной связи, генерации учебного контента и сопровождения учебной коммуникации. Вместе с тем сама по себе высокая генеративная способность модели еще не гарантирует педагогической эффективности: необходима интеграция в продуманную учебную архитектуру, учитывающую цели обучения, формат взаимодействия и ведущую роль преподавателя.

В статье предложена концептуальная модель интеллектуальной системы изучения английского языка, ориентированной на использование в профессиональной деятельности преподавателя вуза. Ее ключевое отличие состоит в том, что большая языковая модель рассматривается не как автономный «заменитель преподавателя», а как интеллектуальное ядро ассистирующей системы, способной поддерживать проверку письменных работ, подготовку персонализированной обратной связи, генерацию учебных материалов и модерирование учебной дискуссии. Такой подход согласуется как с международными обзорами по ICALL и ChatGPT в language education, так и с российскими публикациями, где ИИ рассматривается как эффективный, но контролируемый инструмент усиления учебного процесса.

Проведенный вычислительный эксперимент в предлагаемой постановке позволяет оценить не только формальную правильность работы модели, но и ее практическую пригодность для интеграции в рутинную деятельность преподавателя английского языка в вузе. Перспективы дальнейших исследований связаны с созданием программного прототипа системы, расширением набора педагогических сценариев, включением устной речи и проведением педагогического эксперимента с участием преподавателей и студентов.

Список литературы

1. Curle S., Rose H., Yuksel D. English medium instruction in emerging contexts: An editorial introduction to the special issue // *System*. – 2024. – Vol. 122. – P. 103262.
2. Kaur P., Kumar H., Kaushal S. Technology-assisted language learning adaptive systems: a comprehensive review // *International Journal of Cognitive Computing in Engineering*. – 2023. – Vol. 4. – P. 301–313.
3. Weng X., Chiu T.K.F. Instructional design and learning outcomes of intelligent computer assisted language learning: Systematic review in the field // *Computers and Education: Artificial Intelligence*. – 2023. – Vol. 4. – P. 100117.

4. *Lo C.K. et al.* Exploring the application of ChatGPT in ESL/EFL education and related research issues: A systematic review of empirical studies // *Smart Learning Environments*. – 2024. – Vol. 11, No. 1. – P. 50.
5. *Aco D., Liu M.C., Sari H.* The Shifting Role of the Teacher: A Systematic Review of Teacher Perceptions and Practices with AI in the EFL Classroom // *Journal of Research in Instructional Pedagogy*. – 2025. – Vol. 5, No. 3.
6. *Zheng L., Yang Y.* Research perspectives and trends in Artificial Intelligence-enhanced language education: A review // *Heliyon*. – 2024. – Vol. 10. – No. 19.
7. *Du J., Daniel B.K.* Transforming language education: A systematic review of AI-powered chatbots for English as a foreign language speaking practice // *Computers and Education: Artificial Intelligence*. – 2024. – Vol. 6. – P. 100230.
8. *Li B. et al.* A systematic review of the first year of publications on ChatGPT and language education: Examining research on ChatGPT's use in language learning and teaching // *Computers and Education: Artificial Intelligence*. – 2024. – Vol. 7. – P. 100266.
9. *Михайлова Е.М.* Использование искусственного интеллекта в преподавании иностранного языка (английский) студентам высших учебных заведений // *Мир науки. Педагогика и психология*. – 2025. – Т. 13, № 1.
10. *Конколь М.М., Слесарев М.А., Марьина Е.Д., Черная С.Н.* Опыт использования нейросетевого ассистента (ЛИРА) в обучении иностранным языкам // *Бизнес. Цифровая экономика*. – 2025. – № 4. – С. 41–47.

References

1. *Curle S., Rose H., Yuksel D.* English medium instruction in emerging contexts: An editorial introduction to the special issue // *System*. – 2024. – Vol. 122. – P. 103262.
2. *Kaur P., Kumar H., Kaushal S.* Technology-assisted language learning adaptive systems: a comprehensive review // *International Journal of Cognitive Computing in Engineering*. – 2023. – Vol. 4. – P. 301–313.
3. *Weng X., Chiu T.K.F.* Instructional design and learning outcomes of intelligent computer assisted language learning: Systematic review in the field // *Computers and Education: Artificial Intelligence*. – 2023. – Vol. 4. – P. 100117.
4. *Lo C.K. et al.* Exploring the application of ChatGPT in ESL/EFL education and related research issues: A systematic review of empirical studies // *Smart Learning Environments*. – 2024. – Vol. 11, No. 1. – P. 50.
5. *Aco D., Liu M.C., Sari H.* The Shifting Role of the Teacher: A Systematic Review of Teacher Perceptions and Practices with AI in the EFL Classroom // *Journal of Research in Instructional Pedagogy*. – 2025. – Vol. 5, No. 3.
6. *Zheng L., Yang Y.* Research perspectives and trends in Artificial Intelligence-enhanced language education: A review // *Heliyon*. – 2024. – Vol. 10. – No. 19.
7. *Du J., Daniel B.K.* Transforming language education: A systematic review of AI-powered chatbots for English as a foreign language speaking practice // *Computers and Education: Artificial Intelligence*. – 2024. – Vol. 6. – P. 100230.
8. *Li B. et al.* A systematic review of the first year of publications on ChatGPT and language education: Examining research on ChatGPT's use in language learning and teaching // *Computers and Education: Artificial Intelligence*. – 2024. – Vol. 7. – P. 100266.
9. *Mihajlova E.M.* Ispol'zovanie iskusstvennogo intellekta v prepodavanii inostrannogo yazyka (anglijskij) studentam vysshih uchebnyh zavedenij // *Mir nauki. Pedagogika i psihologiya*. – 2025. – Т. 13, № 1.
10. *Konkol' M.M., Slesarev M.A., Mar'ina E.D., Chernaya S.N.* Opyt ispol'zovaniya nejrosetevogo assistenta (LIRA) v obuchenii inostrannym yazykam // *Biznes. Cifrovaya ekonomika*. – 2025. – № 4. – С. 41–47.

Статья поступила в редакцию: 02.05.2026

Received: 02.05.2026

Статья принята к публикации: 25.06.2026

Accepted: 25.06.2026