

ДИДАКТИЧЕСКИЙ ПОТЕНЦИАЛ ЛИНГВИСТИЧЕСКИХ КОРПУСОВ НА БАЗЕ ТЕХНОЛОГИЙ ИСКУССТВЕННОГО ИНТЕЛЛЕКТА ДЛЯ АДАПТАЦИИ УЧЕБНЫХ МАТЕРИАЛОВ.

Авраменко Анна Петровна

Тишина Маргарита Андреевна

МГУ имени М.В. Ломоносова

119991, Российская Федерация, г. Москва, Ленинские горы, д. 1.

Аннотация. **Актуальность** нашего исследования заключается в растущем интересе к прикладному применению технологий искусственного интеллекта в различных сферах деятельности человека. Нашей **целью** является рассмотрение нейронных сетей как примера технологий искусственного интеллекта для адаптации текстов в преподавании иностранных языков. Для достижения поставленной цели мы обращаемся к **методам** педагогического моделирования и эксперимента, а также статистической обработки данных. **Научная новизна** исследования состоит предложенной методике использования адаптированных нейронными сетями текстов в рамках серии заданий формата «мобильный квест». **Теоретическая значимость** исследования включает в себя уточнение дидактического потенциала выбранного формата заданий. **Практическая значимость**, в свою очередь, обусловлена приведенным примером интеграции задания «мобильный квест» в конкретной учебной ситуации (языковой вуз), а также обзором инструментов анализа текстов на базе технологий искусственного интеллекта. **Результаты** опытно-экспериментальной работы по апробации предложенных материалов и заданий стало статистически значимое подтверждение их эффективности для развития предметных и метапредметных навыков. На основании проведенного пилотирования выдвинутой методики мы делаем **вывод** о необходимости и уместности использования актуальных и адаптированных средствами нейронных сетей текстах в преподавании иностранных языков.

Ключевые слова: методика преподавания иностранных языков, интеграция цифровых технологий в учебный процесс, мобильное обучение, искусственный интеллект в образовании.

DIDACTIC POTENTIAL OF LINGUISTIC CORPORA BASED ON ARTIFICIAL INTELLIGENCE TECHNOLOGIES FOR ADAPTING LEARNING MATERIALS.

Anna P. Avramenko

Margarita A. Tishina

Lomonosov Moscow State University

119991, Russian Federation, Moscow, Leninskiye Gory, 1.

Abstract. The relevance of our study lies in the growing interest in the applied application of artificial intelligence technologies in various fields of human activity. Our goal is to consider neural networks as an example of artificial intelligence technologies for adapting texts in teaching foreign languages. To achieve this goal, we turn to the methods of pedagogical modeling and experiment, as well as statistical

data processing. The scientific novelty of the research consists in the proposed method of using texts adapted by neuron networks within the framework of a “mobile quest” task series. The theoretical significance of the study includes clarifying the didactic potential of the selected task type. The practical significance, in its turn, is due to the given example of integrating the “mobile quest” task in a specific learning environment, as well as a review of text analysis tools based on artificial intelligence technologies. The results of the experiment prove the effectiveness for the development of subject and meta-subject skills. We conclude that it is necessary and appropriate to use relevant and adapted by means of neural networks texts in teaching foreign languages.

Key words: methods of teaching foreign languages, educational technologies, mobile learning, artificial intelligence in education.

ВВЕДЕНИЕ

Развитие *корпусной лингвистики* как отдельного направления относится к концу XX века. Именно в этот период лингвисты начали объединять полученные данные в рамках размеченных массивов лингвистических данных в электронных корпусах. Лингвистическая разметка может быть морфологической, акцентной, синтаксической, дискурсивной и др. Корпусы становятся незаменимым ресурсом, позволяющим исследователям работать с большим объемом данных и анализировать эти данные с точки зрения ключевых для той или иной научной работы критериев. Благодаря использованию корпусов можно, например, сделать выводы о лексическом фонде устойчивых словосочетаний и об особенностях их использования. С момента своего появления электронные лингвистические корпуса находили свое применение в двух направлениях: с одной стороны, получение примеров использования определенной языковой единицы и представления о частотности ее употребления; с другой - решение задач компьютерной лингвистики, в том числе машинное обучение.

На основе частотности использования определенной языковой единицы можно выделить статистику, то есть наглядно представить и проанализировать полученные данные. Также удобным для исследования является паспортизация или метаразметка текстов, то есть их библиографическое описание. Такое многообразие функций позволяет в конечном счете создать собственный *подкорпус* и применять его в рамках конкретного исследования в качестве базы данных [9, 10]. Однако, согласно результатам проведенного нами среди лингвистов опроса (192 участника, март 2022 года), только 34% опрошенных уже имеют опыт работы с электронными лингвистическими корпусами. В то же время 62% хотели бы подробнее ознакомиться с возможностью использования электронных лингвистических корпусов в исследованиях.

Математические методы используются в сравнительно-историческом языкознании. Анализ больших данных в лингвистике позволяет проследить, как с течением времени менялся язык, то есть провести *диахронические исследования* на основе того или иного корпуса. Примером такого рода исследований может служить такое направление в лингвистике, как грамматика

конструкций, занимающаяся изучением лексико-грамматических конструкций в языке, а именно тех лингвистических процессов, которые влияют на изменение значения и контекстов использования анализируемых конструкций. Нейросети (в частности нейронные сети Кохонена) помогают обнаружить родственные слова и звуковые соответствия, реконструировать формы праязыков и на основе этого изучить типологии, классификации и динамику развития языков.

В лингвистических исследованиях большую роль играют такие функции, как *распознавание имен собственных, текстовая классификация и маркировка* различных частей речи. Данные возможности нейросетей могут не только применяться студентами лингвистических факультетов в своих исследованиях, но и использоваться в развлекательных целях. Так, в приложении «Talk to Books» на базе алгоритмов «Google AI» пользователи могут поговорить с книгой, а именно получить ответ на вопрос в виде цитаты из книги. Нейросеть обрабатывает информацию из книг и находит наиболее близкий по смыслу ответ.

Создание, реферирование и перефразирование нейросетью различных текстов имеет множество применений, таких как автоматическое написание отчетов, генерация текстов в сфере маркетинга и финансов, реферирование текстов электронных медицинских карт, создание текстовых прогнозов погоды на основе данных о погоде и даже создание шуток. Развлекательным применением данной функции является алгоритм «Яндекс. Автопоэт», который составляет собственные стихи на основе заголовков из «Яндекс. Новостей».

Производители товаров особенно заинтересованы в *анализе тональности текста или сентимент-анализе*. Книги жалоб и предложений переместились в интернет, где потребители могут оставить отзыв на товар или услуги в открытом доступе. Из-за огромного количества субъективных высказываний человек не способен сам изучить всю информацию, поэтому здесь помогает нейросеть (зачастую рекуррентная сеть). Более того, она не только помогает с подсчетом количества отзывов, сбора статистики (чем занимаются такие агрегаторы, как «Яндекс. Маркет»), но и может проанализировать отзывы, извлекая из текста мнение автора. При анализе тональности нейросеть классифицирует текст или части текста на две или три категории: положительную/отрицательную и в некоторых случаях нейтральную. Для руководства компаний эта информация может помочь с улучшением товаров и услуг и осуществление индивидуального подбора продукции. Однако, нейросеть может сталкиваться с рядом проблем. Программе все еще сложно определять смешанные и нейтральные чувства, сарказм и юмор, а также классифицировать негативные слова, которые могут иметь положительное значение.

Критика использования больших данных в исследовании связана прежде всего с тем, что количественные данные не всегда отражают качественные данные. Показатель частотности использования одного слова на протяжении истории не показывает, каким словом или выражением пользовались до этого. Следующая трудность, возникающая при работе с большими данными - это трудность разделения лингвистических и нелингвистических факторов. Не

всегда очевидно, связан ли всплеск использования того или иного существительного с появлением новой реалии или это внутриязыковой процесс. Наконец, затруднения вызывает определение валидности полученных данных и качество проведенного анализа.

Итак, нейронные сети могут применяться не только в технических сферах, но и в лингвистических исследованиях. Спектр возможностей, которые предоставляют нейросети огромен: электронные словари, анализ тональности текстов, автоматическое создание, реферирование, классификация и проверка текстов, разработка программ, обеспечивающих общение пользователей с интеллектуальными информационными системами и многое другое. *Предметом* нашего исследования выступает методика внедрения адаптированных нейронными сетями текстов в преподавание иностранных языков. Выбор данного предмета исследования обусловлен тем фактом, что научное знание увеличивается в два раза каждый год, а в повседневной жизни мы постоянно путешествуем в информационном пространстве. В связи с этим заучивание материалов становится бессмысленным, и на смену ему приходит развитие умений критического мышления [2, 4, 5]. Привычка обрабатывать большие объемы информации рождает внутреннюю исследовательскую мотивацию, которая формируется в рамках проблемно-поисковых заданий, например формата «мобильный квест» на базе адаптированных актуальных материалов.

ОСНОВНАЯ ЧАСТЬ

Цель, задачи и методы исследования. *Цель* исследования состоит в разработке теоретически обоснованной и экспериментально апробированной методики интеграции адаптированных с помощью нейронных сетей (как технологий искусственного интеллекта) текстов на изучаемом иностранном языке в основную образовательную программу вуза. Для достижения поставленной цели нами были решены следующие *задачи*: исторический обзор развития корпусной лингвистики как контекста работы современных нейронных сетей; анализ дидактического потенциала задания формата «мобильный квест» для внедрения адаптированных текстов; апробация серии заданий формата «мобильный квест» на базе адаптированных текстов; анализ количественных результатов экспериментальной работы. В ходе исследования мы использовали такие *методы* как анализ методической литературы, педагогическое моделирование, педагогический эксперимент, а также статистическая обработка данных с помощью t-критерия Стьюдента.

Методология и организация исследования. Рассмотрим лингвистические электронные корпуса применительно к *образовательному контексту*. На наш взгляд одним из самых интерактивных инструментов для составления своих подкорпусов, например, из активной лексики модуля является сервис «Autocont», который позволяет преподавателю со студентами вместо сета флешкарт составить свой подкорпус и таким образом расширить семантические поля новых лексических единиц. Инструменты «TextInspector», «Sketch Engine», «Monkey Learn» или даже самый лаконичный и бесплатный

сервис «Text Checker» от «Oxford Learner's Dictionaries» позволяют анализировать текст с точки зрения лексических единиц, выделяя для каждой из них разными цветами: функцию, тему, а также представляющий интерес для нашего исследования уровень владения языком по CEFR [6]. Программы могут быть полезны тем преподавателям, которые актуализируют учебные материалы, проектируют авторские задания и составляют контрольно-измерительные материалы.

На базе всех вышеперечисленных типов открытых ресурсов педагоги могут проектировать *учебные квесты*, призванные индивидуализировать процесс обучения в соответствии с интересами, целями, уровнем и фоновыми знаниями учеников. Если говорить о мобильных (в аудитории) и веб (вне аудитории) квестах в курсе по иностранному языку, то очевидно, главным фактором успеха поискового задания является качественный отбор ресурсов, а также представляется разумной разработка критериев отбора для самих учащихся - это не может не способствовать развитию критического анализа при работе с информацией на иностранном языке [8]. Проектная, творческая и поисковая работа иногда бывает недостаточно структурированной и объективной. Для систематизации и процесса выполнения творческого задания, и его контроля можно использовать ресурс «Zunal», который служит шаблоном для поэтапного описания проектного задания с добавлением мультимедийных и гипертекстовых материалов, а также с таблицами критериев оценки.

Опираясь на классификацию этапов внедрения информационно-коммуникационных технологий в учебный процесс, введенную Н. Хокли и Г. Дадни рассмотрим три этапа внедрения *мобильных заданий*, отражающих постепенную трансформацию их форматов: 1) модификация, 2) частичная трансформация, 3) полная трансформация [7]. На первом этапе сохраняются привычные форматы заданий, однако представление и использование учебных материалов происходит с помощью мобильных устройств. Примером может служить поиск значения слова в мобильном словаре вместо печатного. Второй этап заключается в трансформации традиционных заданий, направленных на развитие речемыслительной деятельности, и постепенном видоизменении форматов заданий с сохранением традиционного результата или продукта. К данному этапу мы относим задание формата «мобильный квест», по результатам которого проводится дискуссионное обсуждение. Наконец, на третьем этапе происходит разработка новых форматов заданий с использованием мобильных технологий, например, задания формата «дополненный квест» с использованием технологий дополненной реальности в аудитории.

«Мобильный квест», как и веб-квест, представляет собой проблемное задание с использованием Интернет-ресурсов, являясь, его разновидностью. В отличие от веб-квеста, который предполагает длительную внеаудиторную работу, «мобильный квест» выполняется в аудитории, что обуславливает его отличительные особенности. «Мобильный квест» имеет те же цели, что и веб-квест, а именно развитие коммуникативной, социокультурной и ИК компетенций, развитие межпредметных умений (исследовательских,

познавательных, работы в команде), навыков и умений обучения через всю жизнь и критического мышления, а также мотивацию обучающихся посредством применения интерактивной модели обучения и интеграции аутентичных и творческих заданий. Различия заключаются в том, что веб-квест направлен на развитие всех четырех видов речевой деятельности, в то время как «мобильный квест» в основном направлен на развитие устных видов речевой деятельности. Если длительность веб-квеста может варьироваться от 4 до 16 академических часов преимущественно самостоятельной работы, то продолжительность мобильного квеста ограничена 0,5-1 академическими часами на аудиторном занятии. Конечным результатом веб-квеста может являться как устное обсуждение, так и создание какого-либо продукта (например, сайта или презентации), а результатом мобильного квеста может быть монолог, диалог, полилог, дискуссия, круглый стол или дебаты. Структура мобильного квеста (задание, ресурсы, дискуссия, оценка) в целом совпадает со структурой веб-квеста (введение, задание, ресурсы, процесс выполнения, оценка, заключение), однако является более упрощенной. Можно отметить, что при проведении мобильного квеста у преподавателя появляется больше возможностей для оперативной обратной связи, а у обучающихся для непосредственной интеракции между собой. Также при таком формате работы проще обеспечить соблюдение принципа общения на изучаемом языке в ходе выполнения задания.

Основной целью мобильного квеста может быть формирование любого из компонентов иноязычной коммуникативной компетенции, в том числе лингвистического. Представляется, что с одной стороны, *«мобильный квест» лингвистического характера* позволяет заострить внимание обучающихся на изучаемых языковых явлениях, которые могут выступать не только как средство решения той или иной задачи, но и как объект изучения. С другой стороны, большинство заданий, предлагаемых для формирования языковых навыков, например, найти примеры употребления слов в словаре или написать текст посредством транскрипции и проверить себя по словарю, являются скорее информационно-поисковыми, а не проблемно-поисковыми, следовательно, при их выполнении не задействуется мышление высшего порядка.

Подводя итог, можно отметить, что существующие работы, посвященные применению мобильных квестов для обучения иностранным языкам, главным образом сосредоточены на вопросах мотивации и выявлении отношения обучающихся к данной форме работы. Поскольку результаты описанных исследований демонстрируют положительный отклик обучающихся и показывают потенциал мобильных квестов для организации активного учебного взаимодействия в аутентичном контексте, представляется, что дальнейшие исследования должны быть направлены на получение объективных данных о результатах обучения различным видам деятельности посредством мобильных квестов. Все *характеристики* мобильных квестов можно разделить на три группы: 1) общие с веб-квестом, 2) отличительные черты мобильного квеста, 3) дополнительные требования к языковому мобильному квесту.

Общие черты с веб-квестом: обязательное наличие аутентичного проблемного задания с неоднозначным решением; использование Интернет-ресурсов («мобильный квест» предполагает использование справочных мобильных приложений, таких как словари, энциклопедии и браузеры); обязательное наличие опор; связь квеста с программой учебной дисциплины и учебным планом. *Отличительные черты* мобильного квеста включают в себя: возможное, но не обязательное использование технологий геолокации, дополненной реальности и специализированных приложений для создания квестов; возможность интеграции реальной окружающей среды в обучение в неизменно виде или с помощью дополненной реальности; возможность организации интенсивного взаимодействия между участниками квеста.

Наконец, обратимся к *требованиям* к языковому мобильному квесту, под которыми понимаются следующие моменты: содержательная насыщенность и разнообразие учебного материала в письменной и устной форме, а также аутентичность и актуальность материалов; выполнение задания требует осмысления языковых явлений, необходимых для решения поставленной задачи; создание условий для общения на изучаемом языке в процессе выполнения задания; наличие опор как для облегчения понимания содержания, так и для заострения внимания обучающихся на языковых формах.

В качестве *примера* интеграции мобильных квестов в иноязычные курсы возьмем дисциплину «Иностранный язык в гуманитарной сфере» на 4 курсе программы бакалавриата по направлению «Лингвистика» на факультете иностранных языков и регионоведения МГУ имени М.В. Ломоносова. В 2022 году в апробации приняли участие две группы студентов: экспериментальная (10 человек) и контрольная (12 человек). В обеих группах студентам предлагается готовить тексты для обсуждения к занятиям по определенным темам. Студентам экспериментальной группы было предложено адаптировать и создавать задания на закрепление лексики с текстами по их интересам (темы обсуждаются на первом занятии в формате мозгового штурма) посредством упомянутого выше ресурса «Text Checker», в то время как в контрольной группе использовались аутентичные тексты без адаптации. При организации работы в экспериментальной группе учитывались трудности, выявленные на предыдущем этапе пилотирования заданий формата «мобильный квест» [1]. В том числе входное тестирование было расширено и учитывало анализ языкового портфолио студентов за три года обучения в вузе.

Результаты исследования и их обсуждение. По итогам освоения 6 тем был проведен промежуточный контроль усвоения активной лексики в обеих группах. Средние результаты промежуточного тестирования по группам составили следующие значения: 76.42% в контрольной группе и 91% в экспериментальной группе соответственно. При обработке приведенных выше количественных данных эксперимента для подтверждения статистической значимости результатов мы применили t-критерий Стьюдента для несвязанных выборок. Значение t-критерия Стьюдента составило 5 при критическом значении 2.086. Приведенные показатели позволяют нам сделать вывод о том, что *различия статистически значимы* ($p=0.000080$).

К *преимуществам* внедрения адаптированных с помощью нейронных сетей материалов в задания формата «мобильный квест» можно отнести реализацию базовых принципов интеграции информационно-коммуникационных технологий следующим образом [3]:

- интерактивность осуществляется на всех трех уровнях: при взаимодействии учащегося с преподавателем в процессе подготовки текста; при взаимодействии учащегося с цифровым сервисом адаптации и автоматического формирования задания; при взаимодействии учащихся между собой, когда подготовивший текст ведущий предлагает остальным послетекстовые задания на формирование лексической компетенции в формате дебатов, круглого стола, ролевой игры и т.д.;
- гибкость наблюдается как в совместном выборе темы на мозговом штурме в ходе первого занятия; так и в языковом уровне по CEFR для адаптации, который был определен в процессе входного тестирования;
- открытость обусловлена широким спектром надежных ресурсов, с которых учащиеся могут выбирать тексты.

Отдельно необходимо отметить высокий уровень развития метапредметных навыков самостоятельной работы при выборе и адаптации текста, а также при разработке задания с ним. К *ограничениям* можно отнести использование в качестве материала одного и того же текста для всех студентов, что открывает перспективу для дальнейшего этапа развития данной формы работы с учетом вариативности подтем внутри темы.

ЗАКЛЮЧЕНИЕ

По результатам проведенного исследования мы можем сделать следующие *выводы*.

1. Цифровизация лингвистических корпусов расширяет массивы больших данных языкового характера. Обработка этих данных возможна благодаря доступным пользователю инструментам, основанным на технологиях искусственного интеллекта (именно на нейронных сетях). Применение больших данных электронных лингвистических корпусов, обработанных нейронными сетями, представляется актуальным вопросом для методистов и разработчиков учебных материалов.

2. Эффективная интеграция адаптированных нейронными сетями текстов в иноязычное образование возможна при условии их систематического использования в определенных форматах заданий. Основываясь на проведенных ранее исследованиях, мы выбрали для апробации формат задания «мобильный квест» на базе адаптированных с помощью искусственного интеллекта текстов.

3. Опытное обучение студентов по системе заданий «мобильный квест» в экспериментальной группе доказало эффективность данной методики по сравнению с традиционной в контрольной группе. Обработка результатов апробации посредством t-критерия Стьюдента показала ее статистическую значимость.

Таким образом, можно говорить о том, что внедрение адаптированных с помощью нейронных сетей текстов в систему заданий формата «мобильный

квест» за счет реализации принципов интерактивности, гибкости и открытости позволяет достичь предметных результатов (в нашем случае развития лексической компетенции) и метапредметных результатов (развития мягких навыков обучения через всю жизнь). *Перспективным направлением для дальнейшего исследования* представляется углубление вариативности текстов для каждого тематического подкорпуса с учетом индивидуальной образовательной траектории студента.

ЛИТЕРАТУРА

1. Авраменко А.П. Лингвокультурная интерференция как фактор формирования индивидуальной траектории обучения иностранному языку школьников // Филологические науки. Вопросы теории и практики. № 4. Тамбов, 2022. С. 1314-1318.
2. Назаренко А.Л. Информационно-коммуникационные технологии в лингводидактике: дистанционное обучение. М., 2013. - 271 с.
3. Полат Е.С., Бухарина М.Ю. Современные педагогические и информационные технологии в системе образования. 3-е изд. М., 2010.
4. Сысоев П.В., Твердохлебова И.П. Дистанционное обучение иностранному языку: новые вызовы в новой реальности // Иностранные языки в школе. № 9. М., 2020. С. 2-4.
5. Тарева Е.Г. Системообразующая функция учебной компетенции в лингводидактической модели иноязычной коммуникативной компетенции // Межкультурное многоязычное образование как фактор социальных трансформаций: становление и развитие научной школы: сборник научных статей. М., 2021. С. 55-63.
6. Companion Volume with New Descriptors (CEFR). Council of Europe, 2018. - 227 p.
7. Dudeney G., Hockey N. Going mobile. UK, 2014. - 120 p.
8. Kaliisa R., Picard M. A. Systematic Review on Mobile Learning in Higher Education: The African Perspective // Turkish Online Journal of Educational Technology, TOJET. Vol. 16. No. 1. 2017. P. 1-18.
9. Koehn P. Neural Language Models // Cambridge University Press. Cambridge, 2020. P. 103-124.
10. Kröger J. L., Lutz O. H., Raschke P. Privacy Implications of Voice and Speech Analysis - Information Disclosure by Inference // IFIP Advances in Information and Communication Technology. Vol. 576. 2020. P. 242-258.

СВЕДЕНИЯ ОБ АВТОРАХ

Авраменко Анна Петровна, кандидат педагогических наук, доцент кафедры теории преподавания иностранных языков факультета иностранных языков и регионоведения МГУ имени М.В. Ломоносова, avram4ik@gmail.com

Тишина Маргарита Андреевна, преподаватель кафедры ... факультета мировой политики МГУ имени М.В. Ломоносова, tishina_margarit@mail.ru