

На правах рукописи



КЛОЧИХИН Виталий Владимирович

**МЕТОДИКА ОБУЧЕНИЯ СТУДЕНТОВ КОЛЛОКАЦИЯМ НА ОСНОВЕ
КОРПУСНЫХ ТЕХНОЛОГИЙ ИСКУССТВЕННОГО ИНТЕЛЛЕКТА
(английский язык, языковой факультет)**

Научная специальность 5.8.2. Теория и методика обучения и воспитания
(гуманитарные науки, среднее профессиональное и высшее образование)

АВТОРЕФЕРАТ

**диссертации на соискание ученой степени
кандидата педагогических наук**

Тамбов-2024

Работа выполнена в Федеральном государственном бюджетном образовательном учреждении высшего образования «Тамбовский государственный университет имени Г.Р. Державина»

Научный руководитель: **Сысоев Павел Викторович**,
доктор педагогических наук, профессор,
профессор кафедры лингвистики и лингводидактики
ФГБОУ ВО «Тамбовский государственный
университет имени Г.Р. Державина»

Официальные оппоненты: **Бартош Дана Казимировна**,
доктор педагогических наук, доцент,
профессор кафедры русского языка как иностранного
ФГБОУ ВО «Московский государственный
лингвистический университет»;

Авраменко Анна Петровна,
кандидат педагогических наук, доцент кафедры
лингвистики и информационных технологий
ФГБОУ ВО «Московский государственный
университет имени М.В. Ломоносова»

Ведущая организация: ФГБОУ ВО «Нижегородский государственный
лингвистический университет имени
Н.А. Добролюбова»

Защита состоится «20» июня 2024 г. в 10.00 часов на заседании совета по защите диссертаций на соискание ученой степени кандидата наук, на соискание ученой степени доктора наук 24.2.409.03 на базе ФГБОУ ВО «Тамбовский государственный университет имени Г.Р. Державина» по адресу: 392000, г. Тамбов, ул. Рылеева, д. 52а, зал заседаний диссертационных советов.

С диссертацией можно ознакомиться в научной библиотеке ФГБОУ ВО «Тамбовский государственный университет имени Г.Р. Державина» по адресу: 392008, г. Тамбов, ул. Советская, д. 6 и на официальном сайте ФГБОУ ВО «Тамбовский государственный университет имени Г.Р. Державина» <http://www.tsutmb.ru>.

Автореферат разослан «__» _____ 2024 г.

Ученый секретарь
диссертационного совета
кандидат педагогических наук, доцент



Евстигнеева И.А.

ОБЩАЯ ХАРАКТЕРИСТИКА РАБОТЫ

Актуальность исследования. Современные тенденции общества определяют социальный заказ на подготовку высококвалифицированных педагогических кадров, владеющих навыками, направленными на решение задач в профессиональной сфере. В этой связи обучение в вузе на языковых направлениях подготовки проходит в рамках компетентностного подхода, позволяющего формировать спектр компетенций, необходимых для эффективной профессиональной деятельности обучающихся (Хуторской А.В., 2003; Шадриков В.Д., 2006; Зимняя И.А., 2003; Сафонова В.В., 2004). Одной из основных целей обучения иностранному языку студентов языкового вуза выступает формирование иноязычной коммуникативной компетенции, включающей ряд составных компонентов: языкового, речевого, социокультурного, учебно-познавательного и компенсаторного. Иноязычная коммуникативная компетенция является многоуровневой категорией. Это означает, что при переходе от уровня В1/В2 к уровню С1/С2 студенты продолжают формировать все составляющие компоненты, расширяя свои знания и практические умения в области использования иностранного языка как средства межкультурного и профессионального общения. Примером такого постоянного развития и повышения сложности материала может служить формирование лексической компетенции обучающихся. Если на начальном этапе содержание обучения лексике включает знание, понимание и использование в иноязычной речевой деятельности лимитированного количества лексические единиц, то на более продвинутых уровнях обучения объем лексических единиц значительно увеличивается и их качественный состав путем изучения коллокаций усложняется. Владение коллокациями, обозначающее способность правильно образовывать лексические единства с учетом правил и особенностей лексической сочетаемости, ввиду неизбежной межязыковой интерференции традиционно вызывает трудности у изучающих иностранный язык и требует от преподавателей постоянного поиска новых методов, форм, технологий обучения.

Пандемия коронавирусной инфекции COVID-2019 и экстренное введение дистанционной формы обучения практически во всех странах мира послужили определенным катализатором для переосмысления лингводидактического потенциала современных информационных и коммуникационных технологий и более широкого их внедрения в образование. Это обстоятельство не могло не отразиться на изменениях в политике в сфере образования. В октябре 2021 г. ЮНЕСКО принимает Глобальную декларацию о включении цифровых технологий в образование. Среди ключевых положений декларации отдельно выделен принцип, связанный с изменениями в организационной модели образования. На современном этапе информатизации общества особую актуальность приобретает переход от исключительно очной к гибридной форме обучения, сочетающей очную и дистанционную онлайн- и офлайн-формы (Авраменко А.П., 2020; Титова С.В., Барина К.В., 2020; Сысоев П.В., Твердохлебова И.П., 2020; Поляков О.Г., 2020; Евстигнеев М.Н., 2020). Корпусные технологии искусственного интеллекта – одно из современных цифровых средств, которое может использоваться в учебном процессе для формирования лексических навыков речи обучающихся. Анализ лингводидактических свойств корпусных технологий искусственного интеллекта позволяет утверждать, что данные технологии могут использоваться обучающимися и преподавателем в учебных целях как в аудиторное время, так и во внеаудиторное

время дистанционно (онлайн и офлайн) для выполнения обучающимися поисковых и поисково-исследовательских заданий, а также в научной работе и проектной деятельности обучающихся. Вместе с тем, разработка теоретических основ и практической методики обучения студентов языкового вуза коллокациям на основе корпусных технологий искусственного интеллекта не представляла предмет отдельного изучения, что и определило актуальность данного исследования.

Степень разработанности проблемы. В настоящий момент существует обширный корпус научных работ, в которых были рассмотрены частные вопросы обучения студентов коллокациям на основе корпусных технологий искусственного интеллекта. Изучение корпуса работ отечественных и зарубежных исследователей свидетельствует о рассмотрении следующих вопросов:

- разработка моделей и методик обучения студентов коллокациям (Алексеева Л.Б., 2010; Багарян А.А., 2004; Шульгина Е.М., Федорко В.А., 2019; Трифонова И.С., Левенкова А.Ю., 2020; Аксенова И.Н., 2019; McCarthy M., O'Dell F., 2008; Hill J., 1999; Sinclair J., 1991; Firth J., 1957; Lackman K., 2011; Heikkilä T., 2005; Dukali A., 2018; Men H., 2018; Dokchandra D., 2019; El-Dakhs D., 2015; Lewis M., 2000);

- применение технологий искусственного интеллекта в обучении иностранному языку (Сысоев П.В., 2023; Сысоев П.В., Филатов Е.М., Сорокин Д.О., 2023; Сысоев П.В., Филатов Е.М., 2023; Титова С.В., 2023; Кадеева О.Е., Сырицина В.Н., 2020; Харламенко И.В., 2023; Halaweh M., 2023; Hwang G.J., Chang C.Y., 2021; Kim H.S., Cha Y., Kim N.Y., 2021; Schmidt T., Strasser T., 2022; Евстигнеев М.Н., 2023);

- применение корпусных технологий в методике обучения иностранному языку (Сысоев П.В., 2010; Горина О.Г., 2014; Сысоев П.В., Сёмич Ю.И., 2019; Сысоев П.В., Золотов П.Ю., 2020; Павлова О.Ю., 2021; Мухамадырова А.Ф., 2021; Сысоев П.В., Кокорева А.А., 2013; Crosthwaite P., Luciana L., Schweinberger M., 2021; Fang L., Ma Q., Yan J., 2021; O'Keeffe A., Mark G., 2022; Poole R., 2022; Reppen R., 2010; Zaki M., 2022; Zareva A., 2017; Boulton A., 2017; Daskalovska N., 2015).

Вместе с тем, несмотря на наличие целого ряда работ, посвященных отдельным аспектам обучения студентов коллокациям на основе корпусных технологий искусственного интеллекта, следует констатировать, что существует ряд вопросов, которые требуют более детального рассмотрения:

- не определены лингводидактические свойства корпусных технологий искусственного интеллекта, на основе которых можно разрабатывать методики обучения студентов коллокациям;

- не выявлены психолого-педагогические условия обучения студентов коллокациям на основе корпусных технологий искусственного интеллекта;

- не разработана методическая модель обучения студентов коллокациям на основе корпусных технологий искусственного интеллекта;

- не разработаны этапы обучения студентов коллокациям на основе корпусных технологий искусственного интеллекта;

- не разработана система упражнений для обучения студентов коллокациям на основе корпусных технологий искусственного интеллекта.

Вышеперечисленные проблемы указывают на наличие **противоречий** между: потребностью обучающихся языковых направлений в изучении коллокаций и неразработанностью методик их обучения с использованием современных средств; лингводидактическим потенциалом современных корпусных технологий

искусственного интеллекта и их применением в процессе обучения студентов коллокациям.

Обозначенные противоречия определяют **проблему** диссертационного исследования: в чем заключается методика обучения студентов коллокациям на основе корпусных технологий искусственного интеллекта. Постановка проблемы определила тему диссертационного исследования – «**Методика обучения студентов коллокациям на основе корпусных технологий искусственного интеллекта (английский язык, языковой факультет)**».

Объектом диссертационного исследования выступает процесс обучения коллокациям студентов языковых направлений подготовки.

Предмет исследования – методика обучения студентов коллокациям на основе корпусных технологий искусственного интеллекта.

Цель диссертационного исследования – разработать и теоретически обосновать методику обучения студентов коллокациям на основе корпусных технологий искусственного интеллекта с последующей проверкой ее эффективности в ходе педагогического эксперимента.

Для достижения поставленной цели необходимо решить ряд **задач**:

1) определить лингводидактические свойства корпусных технологий искусственного интеллекта, позволяющие обучать студентов коллокациям;

2) выявить и обосновать комплекс психолого-педагогических условий обучения студентов коллокациям на основе корпусных технологий искусственного интеллекта;

3) разработать методическую модель и этапы обучения студентов коллокациям на основе корпусных технологий искусственного интеллекта;

4) разработать систему упражнений на основе корпусных технологий искусственного интеллекта для освоения студентами английских коллокаций.

Гипотеза исследования: обучение студентов коллокациям на основе корпусных технологий искусственного интеллекта будет результативным, если при разработке методики обучения учитываются:

– лингводидактические свойства корпусных технологий искусственного интеллекта: 1) возможность детализированного поиска (многоуровневость, выбор функциональных типов текстов, выбор языка текста); 2) комплексные результаты поиска (контекстность, систематизация, иллюстративность, статистический анализ); 3) актуальность текстовых данных; 4) онлайн-доступ;

– психолого-педагогические условия, которые влияют на эффективность методики обучения студентов коллокациям на основе корпусных технологий искусственного интеллекта: 1) мотивация обучающихся изучать коллокации на основе корпусных технологий искусственного интеллекта; 2) сформированность ИКТ-компетенции у преподавателя иностранного языка; 3) владение обучающимися иностранным языком на уровне В1; 4) следование выделенным этапам проектной деятельности;

– алгоритм обучения, состоящий из восьми этапов: 1) изучение нового лексического материала на занятии по иностранному языку; 2) организационный; 3) обсуждение вопросов обеспечения информационной безопасности в киберпространстве; 4) поисково-исследовательский; 5) обсуждение в мини-группах выявленных коллокаций; 6) выполнение коммуникативных заданий; 7) оценка и самооценка выполнения коммуникативных заданий; 8) рефлексия;

– система упражнений для обучения студентов коллокациям на основе корпусных технологий искусственного интеллекта, состоящая из упражнений на выявление наиболее распространенных коллокаций, определение значения коллокации, определение верных частей коллокации, определение контекста коллокации, замену части коллокации.

Для решения поставленных в диссертационном исследовании задач использовались следующие **методы исследования**:

– теоретические: изучение и анализ научно-методической литературы по теме диссертационного исследования, систематизация и экстраполяция данных;

– эмпирические: планирование и реализация педагогического эксперимента;

– математические: обработка и анализ количественных данных.

Методологической основой диссертационного исследования выступают положения: **системного подхода** (Блауберг И.В., Садовский В.Н., Юдин Э.Г., 1970, 1978; Садовский В.Н., 1980; Лекторский В.А., Садовский В.Н., 1960; Блауберг И.В., 1980); **компетентностного подхода** (Хуторской А.В., 2003; Болотов В.А., Сериков В.В., 2003; Зимняя И.А., 2006; Шадриков В.Д., 2006); **контекстного подхода** (Вербицкий А.А., 1987, 1991; Johnston E.B., 2002); **корпусного подхода** (Barlow M., 1996; Sinclair J., 1991; Römer U., 2006); **коммуникативно-когнитивного подхода** (Щепилова А.В., 2003; Шамов А.Н., 2008); **личностно-деятельностного подхода** (Выготский Л.С., 1982; Эльконин Д.Б., 1971; Рубинштейн С.Л., 1989; Леонтьев А.Н., 1977; Якиманская И.С., 1979).

В качестве **теоретической базы** исследования выступают работы по **применению современных информационных технологий в обучении иностранному языку** (Титова С.В., 2003, 2023; Евстигнеева И.А., 2013; Сысоев П.В., Твердохлебова И.П., 2020; Сысоев П.В., 2012; Раицкая Л.К., 2007; Сысоев П.В., Филатов Е.М., 2023; Сысоев П.В., Филатов Е.М., Сорокин Д.О., 2023; Харламенко И.В., 2023; Буров В.А., Попова Н.В., 2023), в том числе **корпусных технологий** (Сысоев П.В., Семич Ю.И., 2019; Сысоев П.В., Золотов П.Ю., 2020; Павлова О.Ю., 2021; Мухамадьярова А.Ф., 2021; Crosthwaite P., Luciana L., Schweinberger M., 2021; Fang L., Ma Q., Yan J., 2021; O’Keeffe A., Mark G., 2022; Poole R., 2022; Reppen R., 2010; Zaki M., 2022; Zareva A., 2017; Boulton A., 2017; Daskalovska N., 2015), **методике обучения студентов коллокациям** (Алексеева Л.Б., 2010; Шульгина Е.М., Федорко В.А., 2019; Трифонова И.С., Левенкова А.Ю., 2020; Аксенова И.Н., 2019; McCarthy M., O’Dell F., 2008; Dukali A., 2018; Men H., 2018; Dokchandra D., 2019; El-Dakhs D., 2015).

Экспериментальная база диссертационного исследования. Исследование проводилось на кафедре лингвистики и лингводидактики Тамбовского государственного университета им. Г.Р. Державина. Для апробации экспериментальной методики обучения были выбраны студенты третьего курса направлений подготовки 44.03.01 «Педагогическое образование (Английский язык)» и 45.03.02 «Лингвистика» (профиль «Теория и методика преподавания иностранных языков и культур»).

Обозначение этапов и организация диссертационного исследования. Работа над исследованием проводилась в период с 2022 по 2024 г. и включала в себя три последовательных этапа.

Первый этап (2022 г.) включал в себя изучение и анализ научно-исследовательских работ по теме диссертационного исследования, были поставлены

цель, задачи и проблема исследования, определены гипотеза исследования, объект и предмет научного исследования.

Второй этап (2022–2023 гг.) охватывал разработку теоретических основ методики обучения студентов коллокациям на основе корпусных технологий искусственного интеллекта, в частности: определение лингводидактических свойств корпусных технологий искусственного интеллекта; выявление комплекса психолого-педагогических условий обучения студентов коллокациям на основе корпусных технологий искусственного интеллекта; разработку этапов и методической модели обучения студентов коллокациям на основе корпусных технологий искусственного интеллекта; разработку комплекса упражнений и заданий для обучения студентов коллокациям на основе корпусных технологий искусственного интеллекта.

Третий этап (2023–2024 гг.) включал проведение педагогического эксперимента с целью проверки экспериментальной методики обучения. Была проведена математическая обработка полученных количественных данных. Сформулированы основные выводы по работе и составлено заключение.

Обоснованность и достоверность результатов диссертационного исследования обусловлена опорой на фундаментальные теоретические научно-методические работы по теме исследования, использованием методов исследования, соответствующих поставленным целям и задачам, успешной апробацией инновационной методики обучения студентов коллокациям на основе корпусных технологий искусственного интеллекта.

Научная новизна:

- *разработана* методика обучения студентов коллокациям на основе корпусных технологий искусственного интеллекта;
- *определены* дидактические свойства корпусных технологий искусственного интеллекта, позволяющие обучать студентов коллокациям;
- *обоснован* комплекс психолого-педагогических условий обучения студентов коллокациям на основе корпусных технологий искусственного интеллекта;
- *разработана* методическая модель и этапы обучения студентов коллокациям на основе корпусных технологий искусственного интеллекта;
- *разработана* система упражнений для обучения студентов коллокациям на основе корпусных технологий искусственного интеллекта.

Теоретическая значимость:

- *разработана* теоретическая база методики обучения студентов коллокациям на основе корпусных технологий искусственного интеллекта;
- *доказана* эффективность применения корпусных технологий искусственного интеллекта в обучении студентов коллокациям;
- *раскрыты* проблема диссертационного исследования и лежащие в ее основе противоречия между: потребностью обучающихся языковых направлений в изучении коллокаций и неразработанностью методик их обучения с использованием современных средств; лингводидактическим потенциалом современных корпусных технологий искусственного интеллекта и их применением в процессе обучения студентов коллокациям.

Практическая значимость:

- *дано описание* системы упражнений для обучения студентов коллокациям на основе корпусных технологий искусственного интеллекта;
- *предложены* методические рекомендации по интеграции корпусных технологий искусственного интеллекта;

– результаты исследования могут быть использованы для разработки курсов по методике обучения иностранному языку студентов языковых направлений подготовки.

На защиту выносятся следующие положения:

1. Корпусные технологии искусственного интеллекта – это современное средство обучения иностранному языку, представляющее собой программу, в основе которой используются языковые данные лингвистического корпуса, и с помощью алгоритмов искусственного интеллекта естественный язык, машинное обучение, анализ данных, веб-скрапинг, распознавание речи, интеллектуальная система обучения извлекают, обрабатывают и предоставляют результаты поиска. Корпусные технологии искусственного интеллекта обладают следующими лингводидактическими свойствами: 1) возможность детализированного поиска (многоуровневость, выбор функциональных типов текстов, выбор языка текста); 2) комплексные результаты поиска (контекстность, систематизация, иллюстративность, статистический анализ); 3) актуальность текстовых данных; 4) онлайн-доступ.

2. Обучение коллокациям обучающихся языковых направлений подготовки на основе корпусных технологий искусственного интеллекта будет эффективным, если при разработке методики обучения будут учитываться следующие психолого-педагогические условия: 1) мотивация обучающихся изучать коллокации на основе корпусных технологий искусственного интеллекта; 2) сформированность ИКТ-компетенции у преподавателя иностранного языка; 3) владение обучающимися иностранным языком на уровне В1; 4) следование выделенным этапам проектной деятельности.

3. Методическая модель обучения студентов коллокациям на основе корпусных технологий искусственного интеллекта представляет собой систему взаимосвязанных и взаимодополняющих элементов, встроенных в иерархической последовательности. Модель состоит из четырех блоков: целевой, теоретический, организационно-деятельностный и оценочно-результативный. Методологической основой модели выступают системный, компетентностный, контекстный, корпусный, коммуникативно-когнитивный и личностно-деятельностный подходы. Данные подходы реализуются с учетом ряда общедидактических (принцип информатизации процесса обучения, принцип сознательности, принцип активности, принцип доступности и посильности, принцип наглядности, принцип индивидуализации и дифференциации обучения) и методических (принцип коммуникативной направленности, принцип ситуативно-тематической организации учебного материала, принцип методической целесообразности, принцип устного опережения, принцип аппроксимации учебной деятельности) принципов.

4. Практическая методика обучения студентов коллокациям на основе корпусных технологий искусственного интеллекта состоит из восьми этапов: 1) изучение нового лексического материала на занятии по иностранному языку; 2) организационный; 3) обсуждение вопросов обеспечения информационной безопасности в киберпространстве; 4) поисково-исследовательский; 5) обсуждение в мини-группах выявленных коллокаций; 6) выполнение коммуникативных заданий; 7) оценка и самооценка выполнения коммуникативных заданий; 8) рефлексия.

5. Методика обучения коллокациям на основе корпусных технологий искусственного интеллекта включает систему упражнений, состоящую из упражнений на выявление наиболее распространенных коллокаций, определение

значения коллокации, определение верных частей коллокации, определение контекста коллокации, замену части коллокации.

Личный вклад соискателя заключается в разработке методики обучения студентов коллокациям на основе корпусных технологий искусственного интеллекта, включающей: 1) определение лингводидактических свойств корпусных технологий искусственного интеллекта, позволяющих обучать студентов коллокациям; 2) выявление и обоснование комплекса психолого-педагогических условий обучения студентов коллокациям на основе корпусных технологий искусственного интеллекта; 3) разработку методической модели и этапов обучения студентов коллокациям на основе корпусных технологий искусственного интеллекта; 4) разработку системы упражнений для обучения студентов коллокациям на основе корпусных технологий искусственного интеллекта.

Апробация исследования. Материалы исследования обсуждались на заседаниях кафедры лингвистики и лингводидактики ФГБОУ ВО «Тамбовский государственный университет имени Г.Р. Державина» (2022–2024 гг.), а также на всероссийской научной конференции преподавателей и студентов «Державинские чтения» в ФГБОУ ВО «Тамбовский государственный университет имени Г.Р. Державина» (2022, 2023 гг.), международной научной конференции «Язык и культура в ФГАОУ ВО «Национальный исследовательский Томский государственный университет» (2022, 2023 гг.), всероссийской научно-методической конференции «Развивающий и воспитательный потенциал иностранного языка: лингвометодическое наследие Р.П. Мильруда (1948–2020)» (2023 г.), всероссийской научно-практической конференции «Технологии искусственного интеллекта в обучении иностранному языку» в ФГБОУ ВО «Тамбовский государственный университет имени Г.Р. Державина» (2023 г.).

По материалам диссертационного исследования опубликовано двенадцать научных работ, в том числе девять в журналах, рекомендованных ВАК РФ, одна в журнале, индексируемом в МНБ Scopus.

Структура диссертации. Диссертационное исследование состоит из введения, двух глав, выводов после каждой главы, заключения, библиографического списка (175 наименований) и приложения.

ОСНОВНОЕ СОДЕРЖАНИЕ РАБОТЫ

В первой главе диссертационного исследования «**Теоретические основы обучения студентов коллокациям на основе корпусных технологий искусственного интеллекта**» на основе анализа научно-методических работ овладение коллокациями определяется как **знание коллокаций, понимание их значения и употребления, а также способность распознавать и использовать коллокации в иноязычной устной и письменной речи**. В рамках исследования были выделены три элемента обучения коллокациям: 1) распознавание коллокаций – поиск и определение коллокаций в незнакомой устной или письменной речи; 2) понимание коллокаций – определение значения коллокаций; 3) использование коллокаций – способность употреблять коллокации в устной и письменной речи. Овладение коллокациями возможно только при условии овладения каждым из элементов их обучения.

Электронный лингвистический корпус является востребованным средством как для преподавателя, так и для студентов в изучении иностранного языка. Однако следует отметить, что традиционный лингвистический корпус обладает рядом

недостатков, которые ограничивают его дидактический потенциал, к ним относятся: 1) баланс текстовых данных; 2) технические ограничения; 3) недостаток корпусов на материале устной речи; 4) недостаток графических материалов; 5) зависимость от оценки эксперта. Современные корпусные технологии искусственного интеллекта являются выгодным решением в преодолении данных недостатков (табл. 1).

Таблица 1

**Отличия лингвистического корпуса и корпусных технологий
искусственного интеллекта**

| Аспект сравнения | Электронный лингвистический корпус | Корпусные технологии искусственного интеллекта |
|---|--|--|
| Сбор данных | Языковые данные собираются вручную, из заранее подобранных книг, журналов, статей и других письменных источников. Этот процесс подлежит тщательному контролю для обеспечения репрезентативности данных | Алгоритмы искусственного интеллекта самостоятельно ищут в открытых источниках, извлекают и сортируют языковые данные в соответствии с запрашиваемыми критериями. В дополнение, искусственный интеллект может предоставить рекомендации по содержанию корпуса |
| Аннотирование и анализ данных | Лингвисты-исследователи вручную или полуавтоматически осуществляют аннотирование текстовых данных, добавляя теги, метки и т.п. | Алгоритмы искусственного интеллекта способствуют автоматизации процессов аннотирования, маркировки частей речи, именованных объектов, синтаксических структур и др. |
| Взаимодействие с пользователем | Взаимодействие пользователя с традиционными корпусами включает ручной поиск с помощью различных символов и знаков и традиционные инструменты анализа данных | Корпусные технологии на основе искусственного интеллекта предлагают интерактивные интерфейсы пользователя, визуализацию данных и инструменты для изучения и анализа данных |
| Обновление данных и адаптируемость | При необходимости обновления языковых данных и/или добавления новой разметки традиционные корпуса требуют ручного вмешательства лингвиста-исследователя | Алгоритмы искусственного интеллекта используются для постоянного обновления и адаптации корпусных данных на основе развивающихся языковых моделей или отзывов пользователей |

На основе анализа научной литературы было дано определение понятия *корпусные технологии искусственного интеллекта*, под которыми понимаются *программы, в основе которых используются языковые данные лингвистического корпуса, и с помощью алгоритмов искусственного интеллекта естественный язык, машинное обучение, анализ данных, веб-скрапинг, распознавание речи, интеллектуальная система обучения извлекают, обрабатывают и предоставляют результаты поиска*. Алгоритмы искусственного интеллекта, участвующие в создании корпусных технологий искусственного интеллекта, представлены в таблице 2.

Таблица 2

Алгоритмы искусственного интеллекта, участвующие в создании корпусных технологий искусственного интеллекта

| Технология искусственного интеллекта | Краткое описание | Роль в создании корпусных технологий искусственного интеллекта |
|---|---|---|
| Естественный язык (natural language processing) | Технология, направленная на распознавание, интерпретацию, анализ и создание текстов | <ul style="list-style-type: none"> • обработка текста (токенизация, лемматизация и т.д); • частеречная разметка; • категоризация и классификация объектов в тексте; • определение структуры и правил языка (синтаксис); • определение эмоциональной тональности слов; • референциональное тождество; • определение темы, жанра и т.п. текста; • определение значения слов |
| Машинное обучение (machine learning) | Технология, направленная на выявление закономерностей, на основе введенных данных | <ul style="list-style-type: none"> • аннотирование текста; • моделирование языка; • выделение и группировка текстовых данных по различным признакам; • прогнозирование и определение тенденций языка; • персонализация корпусных опций |
| Анализ данных (datascience) | Технология, направленная на обработку и выявление закономерностей в текстовых данных | <ul style="list-style-type: none"> • управление массивом текстовых данных; • контроль интеграции данных из разных ресурсов; • статистический анализ текстовых данных; • прогнозирование и определение тенденций языка; • персонализация корпусных опций |
| Веб-скрапинг и извлечение данных (web scraping and data mining) | Технологии, направленные на поиск и сбор информации из открытых источников | <ul style="list-style-type: none"> • поиск и сбор актуальных данных |
| Распознавание речи (speech recognition) | Технология, направленная на распознавание устной речи | <ul style="list-style-type: none"> • трансформация устной речи в текстовый формат |
| Интеллектуальная система обучения (intelligent computer system) | Технология, направленная на обработку и представление текстовых данных в упорядоченном виде | <ul style="list-style-type: none"> • представление текстовых результатов поиска; • визуализирование результатов поиска |

В рамках решения **первой задачи** исследования были выделены следующие лингводидактические свойства корпусных технологий искусственного интеллекта:

1. **Возможность детализированного поиска** – большой выбор настроек поиска для решения педагогических и исследовательских задач:

а) многоуровневость – представлен широкий спектр языковых данных, ранжированный по уровню языка (от корпусов с письменными работами студентов до корпуса с литературными произведениями английских авторов);

б) выбор функциональных типов текстов – возможность выбора типа, стиля, жанра и т.п. текста (представлен материал научных журналов, газет, рекламных объявлений, литературных произведений и т.д.);

в) выбор языка текста – возможность выбора корпуса разных языков (русский, английский, китайский и др), в том числе выбора корпуса параллельных текстов.

2. **Комплексные результаты поиска** – представляется полная информации о запрашиваемом слове:

а) контекстность – результаты поискового запроса могут быть продемонстрированы в своем естественном контексте;

б) систематизация – имеется широкий набор опций по фильтрации и сортировке результатов поискового запроса;

в) иллюстративность – текстовые результаты поиска могут быть продемонстрированы в графическом виде, с помощью графиков, таблиц, диаграмм;

г) статистический анализ – результаты поиска содержат статистические данные о частоте, месте и времени употребления запрашиваемого слова, а также используются различные математические формулы для выявления уровня сочетаемости слов (LogDice, MI3, T-score).

3. **Актуальность текстовых данных** – алгоритмы искусственного интеллекта обеспечивают постоянную обновляемость текстовых данных, отражая текущие тенденции языка.

4. **Онлайн-доступ** – корпусные технологии искусственного интеллекта доступны в сети Интернет любому пользователю.

В ходе решения **второй задачи** исследования был выявлен и обоснован комплекс психолого-педагогических условий обучения студентов коллокациям на основе корпусных технологий искусственного интеллекта, включающий следующие условия:

1) мотивация обучающихся изучать коллокации на основе корпусных технологий искусственного интеллекта;

2) сформированность ИКТ-компетенции у преподавателя иностранного языка;

3) владение обучающимися иностранным языком на уровне B1;

4) следование выделенным этапам проектной деятельности.

Учет вышеперечисленных условий будет влиять на результативность авторской методики обучения студентов коллокациям на основе корпусных технологий искусственного интеллекта.

Во второй главе диссертационного исследования **«Практические аспекты обучения студентов коллокациям на основе корпусных технологий искусственного интеллекта»** разрабатывается методическая модель и этапы обучения студентов коллокациям на основе корпусных технологий искусственного интеллекта в рамках **третьей задачи** исследования. Методическая модель представляет собой систему взаимосвязанных компонентов-блоков: а) предпосылки; б) целевой блок; в) теоретический блок; г) организационно-деятельностный блок; д) оценочно-результативный блок. Методологической основой модели выступают системный, компетентностный, контекстный, корпусный, коммуникативно-когнитивный и личностно-деятельностный подходы. Данные подходы реализуются с

учетом ряда общедидактических (принцип информатизации процесса обучения, принцип сознательности, принцип активности, принцип доступности и посильности, принцип наглядности, принцип индивидуализации и дифференциации обучения) и методических (принцип коммуникативной направленности, принцип ситуативно-тематической организации учебного материала, принцип методической целесообразности, принцип устного опережения, принцип аппроксимации учебной деятельности) принципов (рис. 1).



Рис. 1. Методическая модель обучения студентов коллокациям на основе корпусных технологий искусственного интеллекта

Методика обучения студентов коллокациям на основе корпусных технологий искусственного интеллекта включает восемь последовательных этапов (табл. 3).

Таблица 3

**Этапы обучения студентов коллокациям на основе корпусных технологий
искусственного интеллекта**

| Этап | Действия преподавателя | Действия студентов | Форма обучения |
|-------------|---|--|---|
| 1 | Изучение нового лексического материала на занятии по иностранному языку | | |
| | Объясняет новый лексический материал студентам | Изучают новый лексический материал | Аудиторная Индивидуальная Групповая |
| 2 | Организационный | | |
| | Объясняет студентам понятие «коллокация» и цель использования корпусных технологий искусственного интеллекта в ее изучении; цель, задачи и критерии оценки выбранного проекта по развитию лексических навыков; определяет этапы выполнения проекта; знакомит с основными функциями программы для выполнения проекта | Узнают о понятии «коллокация» и роли корпусных технологий искусственного интеллекта в ее изучении; цели, задачах и критериях оценки выбранного проекта по развитию лексических навыков; знакомятся с этапами проведения проекта; изучают основные функции программы для выполнения проекта | Аудиторная Групповая |
| 3 | Обсуждение вопросов обеспечения информационной безопасности в киберпространстве | | |
| | Объясняет правила поведения в киберпространстве при работе с корпусными технологиями искусственного интеллекта | Изучают правила поведения в киберпространстве при работе с корпусными технологиями искусственного интеллекта | Аудиторная Групповая |
| 4 | Поисково-исследовательский | | |
| | – | Осуществляют поиск коллокаций в соответствии с целями проекта; проводят сортировку результатов поиска, изучают и анализируют полученные результаты поиска; выполняют тренировочные упражнения | Аудиторная Групповая/ Внеаудиторная Индивидуальная |
| 5 | Обсуждение в мини-группах выявленных коллокаций | | |
| | Мониторинг обсуждения полученных данных об использовании коллокаций в иноязычной речи | Обсуждают полученные данные об использовании коллокаций в иноязычной речи | Аудиторная Групповая |
| 6 | Выполнение коммуникативных заданий | | |
| | Мониторинг использования обучающимися изученных коллокаций при выполнении коммуникативных заданий | Используют изученные коллокации при выполнении коммуникативных заданий | Аудиторная Индивидуальная |
| 7 | Оценка и самооценка выполнения коммуникативных заданий | | |
| | Оценивает выполнение проекта в соответствии с заданными критериями | Проводят самооценку выполнения проекта в соответствии с заданными критериями | Аудиторная Внеаудиторная Индивидуальная |
| 8 | Рефлексия | | |
| | Осуществляет рефлексию реализации проекта | Осуществляют рефлексию участия в проекте | Аудиторная Внеаудиторная Индивидуальная |

Следует отметить, что данные этапы обучения студентов коллокациям на основе корпусных технологий искусственного интеллекта носят общий характер и при возможности или желании преподавателя могут быть видоизменены или пропущены в соответствии с поставленными целями обучения.

Педагогический эксперимент по апробации инновационной методики обучения студентов коллокациям на основе корпусных технологий искусственного интеллекта проходил на базе кафедры лингвистики и лингводидактики Тамбовского государственного университета имени Г.Р. Державина во втором семестре 2022–2023 учебного года. Участниками педагогического эксперимента выступили студенты третьего курса направлений подготовки 44.03.01 «Педагогическое образование (Английский язык)» (контрольная группа – КГ (N = 20) и 45.03.02 «Лингвистика» (профиль «Теория и методика преподавания иностранных языков и культур»)) (экспериментальная группа – ЭГ (N = 20). Экспериментальное обучение проходило в три этапа.

Констатирующий этап. На данном этапе студенты проходили языковой тест, определяющий исходный уровень знаний коллокаций. Тест состоял из заданий следующего типа: множественного выбора, установлений соответствий, раскрытия скобок и написания верной формы, перевода коллокаций, формулировки значения коллокаций, выделения коллокаций в тексте. В тесте было сорок одно задание, максимальный балл – сто.

Формирующий этап. На этом этапе студенты КГ и ЭГ проходили обучение в рамках дисциплины «Иностранный язык (английский)». Неварьируемыми переменными выступали: 1) учебные материалы (студенты обеих групп на занятиях по практическому английскому языку занимались по учебным пособиям: Crace A., Acklam R. *New Total English. Upper Intermediate Students' Book*. N.Y.: Pearson, 2011; O'Dell F., McCarthy M. *English Collocations in Use. Advanced*. Cambridge: Cambridge University Press, 2017); 2) объем аудиторной нагрузки (четыре пары по полтора часа каждая в неделю); 3) преподаватель практического курса по английскому языку. К варьируемой переменной относилась методика обучения коллокациям. Вместе со стандартной учебной нагрузкой, студенты ЭГ принимали участие в проектной деятельности по изучению коллокаций на основе корпусных технологий искусственного интеллекта с помощью программы Sketch Engine. Студенты на исследовательском этапе проводили поисковые запросы новых лексических единиц и в рамках **четвертой задачи** исследования выполняли тренировочные упражнения (рис. 2). Программа обеспечивала статистически подкрепленные и визуализированные результаты поиска, что облегчило восприятие новых коллокаций (рис 3).

WORD SKETCH

partner as noun 7,155,177× ...

| modifiers of "partner" | | | |
|---|---------|-----|-----|
| trading | 62,419 | 8.7 | ... |
| trading partners | | | |
| • concentrated in: politics & government ? | | | |
| • concentrated in: economy & finance & business ? | | | |
| • concentrated in: news ? | | | |
| business | 139,084 | 8.4 | ... |
| business partners | | | |
| • concentrated in: economy & finance & business ? | | | |
| community | 69,021 | 8.0 | ... |
| community partners | | | |
| • concentrated in: education ? | | | |
| strategic | 44,670 | 7.9 | ... |
| strategic partner | | | |
| • concentrated in: economy & finance & business ? | | | |
| industry | 44,029 | 7.8 | ... |
| industry partners | | | |
| • concentrated in: education ? | | | |
| coalition | 19,919 | 7.3 | ... |
| coalition partners | | | |
| • concentrated in: politics & government ? | | | |
| • concentrated in: multi-topic ? | | | |
| • concentrated in: news ? | | | |
| channel | 21,482 | 7.3 | ... |
| channel partners | | | |

| verbs with "partner" as object | | | |
|---|--------|-----|-----|
| manage | 65,436 | 8.4 | ... |
| managing partner | | | |
| • concentrated in: economy & finance & business ? | | | |
| • concentrated in: legal ? | | | |
| trust | 26,471 | 8.4 | ... |
| a trusted partner | | | |
| • concentrated in: economy & finance & business ? | | | |
| spar | 8,688 | 7.6 | ... |
| sparring partner | | | |
| • concentrated in: culture & entertainment ? | | | |
| • concentrated in: sex ? | | | |
| • concentrated in: sports ? | | | |
| show more (1) | | | |
| become | 71,411 | 7.1 | ... |
| became a partner | | | |
| • concentrated in: history ? | | | |
| • concentrated in: reference/encyclopedia ? | | | |
| value | 7,093 | 6.6 | ... |
| a valued partner | | | |
| • concentrated in: economy & finance & business ? | | | |
| choose | 15,951 | 6.3 | ... |
| choose a partner | | | |
| seek | 12,923 | 6.2 | ... |

Рис. 2. Результаты поискового запроса леммы «partner»

WORD SKETCH

partner as noun 7,155,177× ...

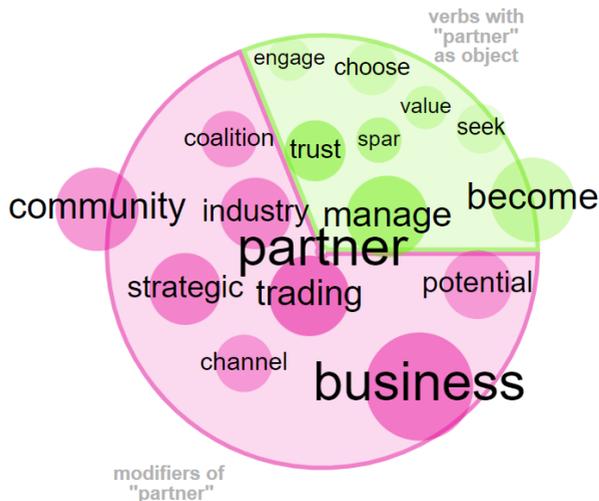


Рис. 3. Результаты поискового запроса леммы «partner»

Контрольный этап. На данном этапе после завершения педагогического эксперимента студенты обеих групп выполнили тот же тест, который они выполняли на констатирующем этапе.

Для подтверждения эффективности авторской методики обучения студентов коллокациям на основе корпусных технологий искусственного интеллекта результаты контрольного и экспериментального срезов в контрольной и экспериментальной группах были обработаны с использованием ПО IBM SPSS Statistics 28. В частности, была применена методика сравнения средних величин t-критерия Стьюдента. Данный метод статистического анализа позволяет выявить наличие или отсутствие статистической значимости в различиях между результатами контрольного и экспериментального срезов в каждой из групп, а также между контрольной и экспериментальной группами до и после эксперимента.

Анализ данных, представленных в таблице 4, демонстрирует отсутствие статистической значимости между результатами выполнения теста на констатирующем этапе. По трем контролируемым элементам обучения коллокациям р-значение t-критерия Стьюдента превышает 0,05.

Таблица 4

Данные сопоставления результатов теста в КГ и ЭГ на констатирующем этапе

| Элемент обучения коллокациям | Группа | N | t-критерий | p-значение |
|------------------------------|--------|---|------------|------------|
|------------------------------|--------|---|------------|------------|

| | | | | |
|---------------|----|----|------|-------|
| Распознавание | КГ | 20 | 0,24 | 0,4* |
| | ЭГ | 20 | | |
| Понимание | КГ | 20 | 0,22 | 0,41* |
| | ЭГ | 20 | | |
| Использование | КГ | 20 | 0,34 | 0,36* |
| | ЭГ | 20 | | |

* $p > 0,05$

Приведенные в таблице 5 данные указывают на статистически значимый прирост по всем контролируемым элементам обучения коллокациям в ходе проведения эксперимента.

Таблица 5

Данные сопоставления результатов выполнения студентами КГ и ЭГ тестов до и после участия в педагогическом эксперименте

| Элемент обучения коллокациям | Группа | N | t-критерий | p-значение |
|------------------------------|--------|----|------------|---------------|
| Распознавание | КГ | 20 | 7,07 | $p \leq 0,05$ |
| | ЭГ | 20 | 9,84 | $p \leq 0,05$ |
| Понимание | КГ | 20 | 9,75 | $p \leq 0,05$ |
| | ЭГ | 20 | 10,94 | $p \leq 0,05$ |
| Использование | КГ | 20 | 10,83 | $p \leq 0,05$ |
| | ЭГ | 20 | 14,16 | $p \leq 0,05$ |

С целью выявления эффективности авторской методики обучения студентов коллокациям на основе корпусных технологий искусственного интеллекта были сопоставлены результаты формирующего среза в двух группах (табл. 6).

Таблица 6

Данные сопоставления результатов формирующего среза КГ и ЭГ

| Элемент обучения коллокациям | Группа | N | t-критерий | p-значение |
|------------------------------|--------|----|------------|------------|
| Распознавание | КГ | 20 | 3,44 | 0,0007* |
| | ЭГ | 20 | | |
| Понимание | КГ | 20 | 2,82 | 0,003* |
| | ЭГ | 20 | | |
| Использование | КГ | 20 | 2,64 | 0,005* |
| | ЭГ | 20 | | |

* $p \leq 0,05$

Материалы таблицы 6 подтверждают эффективность инновационной методики обучения студентов коллокациям на основе корпусных технологий искусственного интеллекта по всем трем контролируемым аспектам: распознавание коллокаций ($t = 3,44$ при $p = 0,0007$), понимание коллокаций ($t = 2,82$ при $p = 0,003$), использование коллокаций ($t = 2,64$ при $p = 0,005$).

В **заключении** диссертационной работы подводятся итоги, подтверждающие выдвинутую гипотезу, формируются следующие **выводы**:

1. Корпусные технологии искусственного интеллекта – это программы, в основе которых используются языковые данные лингвистического корпуса, и с помощью алгоритмов искусственного интеллекта естественный язык, машинное обучение, анализ данных, веб-скрапинг, распознавание речи, интеллектуальная система обучения извлекают, обрабатывают и предоставляют результаты поиска.

Определены лингводидактические свойства корпусных технологий искусственного интеллекта: 1) возможность детализированного поиска (многоуровневость, выбор функциональных типов текстов, выбор языка текста); 2) комплексные результаты поиска (контекстность, систематизация, иллюстративность, статистический анализ); 3) актуальность текстовых данных; 4) онлайн-доступ.

2. Был выявлен и обоснован комплекс психолого-педагогических условий, учет которых будет влиять на эффективность методики обучения студентов коллокациям на основе корпусных технологий искусственного интеллекта: 1) мотивация обучающихся изучать коллокации на основе корпусных технологий искусственного интеллекта; 2) сформированность ИКТ-компетенции у преподавателя иностранного языка; 3) владение обучающимися иностранным языком на уровне В1; 4) следование выделенным этапам проектной деятельности.

3. Была разработана методическая модель обучения студентов коллокациям на основе корпусных технологий искусственного интеллекта, представляющая собой систему взаимосвязанных компонентов-блоков: а) предпосылки; б) целевой блок; в) теоретический блок; г) организационно-деятельностный блок; д) оценочно-результативный блок. Методологической основой модели выступают системный, компетентностный, контекстный, корпусный, коммуникативно-когнитивный и личностно-деятельностный подходы. Данные подходы реализуются с учетом ряда общедидактических (принцип информатизации процесса обучения, принцип сознательности, принцип активности, принцип доступности и посильности, принцип наглядности, принцип индивидуализации и дифференциации обучения) и методических (принцип коммуникативной направленности, принцип ситуативно-тематической организации учебного материала, принцип методической целесообразности, принцип устного опережения, принцип аппроксимации учебной деятельности) принципов.

4. Разработан алгоритм обучения студентов коллокациям на основе корпусных технологий искусственного интеллекта, включающий восемь этапов: 1) изучение нового лексического материала на занятии по иностранному языку; 2) организационный; 3) обсуждение вопросов обеспечения информационной безопасности в киберпространстве; 4) поисково-исследовательский; 5) обсуждение в мини-группах выявленных коллокаций; 6) выполнение коммуникативных заданий; 7) оценка и самооценка выполнения коммуникативных заданий; 8) рефлексия.

5. Разработана система упражнений для обучения студентов коллокациям на основе корпусных технологий искусственного интеллекта, состоящая из упражнений на выявление наиболее распространенных коллокаций, определение значения коллокации, определение верных частей коллокации, определение контекста коллокации, замену части коллокации.

6. В ходе педагогического эксперимента была доказана гипотеза и эффективность инновационной методики обучения.

Перспективность данной работы состоит в дальнейшем изучении возможности использования корпусных технологий искусственного интеллекта при обучении другим аспектам иностранного языка.

Основные положения диссертации отражены в следующих публикациях автора:

Монографии:

1. Сысоев П.В., Поляков О.Г., Евстигнеев М.Н. и др. Обучение иностранному языку на основе технологий искусственного интеллекта: Монография / под науч. ред.

П.В. Сысоева. Тамбов: Изд. дом «Державинский», – 2023. – 132 с. (авторская доля 0,6 п.л.)

Статьи в научных журналах и материалах конференций:

2. Сысоев П.В., Ключихин В.В. **Формирование коллокационной компетенции студентов на основе корпусных технологий // Перспективы науки и образования. – 2022. – №4 (58). – С. 320-335. – 2 п.л., авторская доля – 1 п.л. (ВАК РФ (К1), Scopus - Q2)**

3. Ключихин В.В. Современный электронный лингвистический корпус: свойства и характеристики // Теория и практика обучения иностранным языкам: традиции и перспективы развития: мат-лы IV Междунар. науч.-практ. конф., г. Москва, 27 марта 2019 г. – М.: Изд-во Мос. гос. обл. ун-та, 2019. – С. 123-128. – 0,37 п.л.

4. Ключихин В.В. **Формирование коллокационной компетенции обучающихся на основе электронного лингвистического корпуса // Вестник Тамбовского университета. Серия: Гуманитарные науки. – 2019. – Т. 24. – № 179. – С. 69-80. – 1,5 п.л.**

5. Ключихин В.В. **Этапы формирования коллокационной компетенции студентов на основе лингвистического корпуса // Вестник Тамбовского университета. Серия: Гуманитарные науки. – 2020. – Т. 25. – № 186. – С. 14-24. – 1,37 п.л.**

6. Ключихин В.В. Коллокационная компетенция: определение понятий и компонентный состав // Державинский форум. – 2021. – Т. 5. – №17. – С. 96-101. – 0,37 п.л.

7. Ключихин В.В. **Психологические аспекты формирования коллокационной компетенции // Вестник Тамбовского университета. Серия: Гуманитарные науки. – 2021. – Т. 26. – №191. – С. 32-37. – 0,75 п.л.**

8. Евстигнеева И.А., Евстигнеев М.Н., Ключихин В.В. **Обеспечение информационной безопасности студентов в процессе использования проектной методики в обучении иностранному языку в университете // Вестник Тамбовского университета. Серия: Гуманитарные науки. – 2022. – Т. 27. – №4. – С. 1009-1019. – 1,37 п.л., авторская доля – 0,45 п.л.**

9. Ключихин В.В. **Методическая модель обучения студентов коллокационной компетенции на основе корпусных технологий // Вопросы методики преподавания в вузе. – 2023. – Т. 12. – №2. – С. 24–36. – 1,62 п.л.**

10. Ключихин В.В. **Психолого-педагогические условия формирования коллокационной компетенции студентов на основе корпусных технологий // Вестник Тамбовского университета. Серия: Гуманитарные науки. – 2023. – Т. 28. – №2. – С. 395-404. – 0,8 п.л**

11. Ключихин В.В., Поляков О.Г. **Технологии искусственного интеллекта: инструменты корпусного анализа в обучении иностранному языку // Иностранные языки в школе. – 2023. – №3. – С. 24–30. – 0,43 п.л. авторская доля – 0,22**

12. Ключихин В.В. **Корпусные технологии искусственного интеллекта в обучении сочетаемости слов и исследовательской работе // Иностранные языки в школе. – 2024. – №3. – С. 39-46. – 0,5 п.л.**

Подписано в печать 18.04.2024 г. Формат 60×84/16. Усл. печ. л. 1,22.
Тираж 100 экз. Заказ № 24106. Бесплатно.
ФГБОУ ВО «Тамбовский государственный университет имени Г.Р. Державина»
392000, г. Тамбов, ул. Интернациональная, 33.
Отпечатано в Издательском доме «Державинский».
392008, г. Тамбов, ул. Советская, 190 г