

Документ подписан простой электронной подписью
Информация о владельце:
ФИО: Ястребов Олег Александрович
Должность: Ректор
Дата подписания: 25.05.2026 09:26:16
Уникальный программный ключ:
ca953a0120d891083f939673078ef1a989dae18a

**Федеральное государственное автономное образовательное учреждение высшего образования
«Российский университет дружбы народов имени Патриса Лумумбы»**

Институт русского языка

(наименование основного учебного подразделения (ОУП) – разработчика ОП ВО)

РАБОЧАЯ ПРОГРАММА ДИСЦИПЛИНЫ

ВВЕДЕНИЕ В КОРПУСНУЮ ЛИНГВИСТИКУ

(наименование дисциплины/модуля)

Рекомендована МССН для направления подготовки/специальности:

45.03.04 ИНТЕЛЛЕКТУАЛЬНЫЕ СИСТЕМЫ В ГУМАНИТАРНОЙ СФЕРЕ

(код и наименование направления подготовки/специальности)

Освоение дисциплины ведется в рамках реализации основной профессиональной образовательной программы высшего образования (ОП ВО):

КИБЕРТЕХНОЛОГИИ И АНАЛИЗ ДАННЫХ В ГУМАНИТАРНОЙ СФЕРЕ

(наименование (профиль/специализация) ОП ВО)

1. ЦЕЛЬ ОСВОЕНИЯ ДИСЦИПЛИНЫ

Дисциплина «Введение в корпусную лингвистику» входит в программу бакалавриата «Кибертехнологии и анализ данных в гуманитарной сфере» по направлению 45.03.04 «Интеллектуальные системы в гуманитарной сфере» и изучается в 3 семестре 2 курса. Дисциплину реализует Кафедра русского языка 1. Дисциплина состоит из 6 разделов и 13 тем и направлена на изучение основных проблем современной корпусной лингвистики, а также на приобретение практических навыков работы с корпусами и связанными с ними компьютерными инструментами в сфере цифровой филологии и цифровой педагогики. Дисциплина «Введение в корпусную лингвистику» занимает одно из центральных мест в учебном плане бакалавриата «Прикладная цифровая филология» и представляет собой фундамент, на котором строится профессиональное становление специалиста в области языка и литературы. Изучение дисциплины «Введение в корпусную лингвистику» формирует у обучающегося целостное представление о языке и тексте как отражении культуры и истории; посредством разбора языковых явлений, в том числе на материале литературных произведений, обучающиеся учатся систематизировать информацию, выявлять причинно-следственные связи, аргументировать свою точку зрения, критически оценивать источники информации.

Целью освоения дисциплины является знакомство студентов с наиболее известными компьютерными ресурсами – национальными корпусами, программами обработки текстов, с основными принципами их устройства, методами исследований в области гуманитарных наук, основанных на анализе текстов. Освоение дисциплины «Введение в корпусную лингвистику» способствует формированию языковой компетенции, в основе которой лежит глубокое понимание структуры языка, его исторических изменений, стилистических особенностей, и позволяет будущему специалисту-филологу свободно ориентироваться в языковом пространстве и профессионально работать с речью – устной и письменной.

2. ТРЕБОВАНИЯ К РЕЗУЛЬТАТАМ ОСВОЕНИЯ ДИСЦИПЛИНЫ

Освоение дисциплины «Введение в корпусную лингвистику» направлено на формирование у обучающихся следующих компетенций (части компетенций):

Таблица 2.1. Перечень компетенций, формируемых у обучающихся при освоении дисциплины (результаты освоения дисциплины)

Шифр	Компетенция	Индикаторы достижения компетенции (в рамках данной дисциплины)
УК-12	Способен: искать нужные источники информации и данные, воспринимать, анализировать, запоминать и передавать информацию с использованием цифровых средств, а также с помощью алгоритмов при работе с полученными из различных источников данными с целью эффективного использования полученной информации для решения задач; проводить оценку информации, ее достоверность, строить логические умозаключения на основании поступающих информации и данных	УК-12.1 Осуществляет поиск необходимых для эффективного решения задач источников информации и данных; УК-12.2 Воспринимает, анализирует, запоминает и передаёт информацию с использованием цифровых средств и алгоритмов с целью эффективного использования этой информации для решения задач; УК-12.3 Проводит оценку информации и её достоверности; УК-12.4 Строит логические умозаключения на основании поступающих информации и данных;
УК-4	Способен к коммуникации в межличностном и межкультурном взаимодействии на русском как иностранном и иностранном(ых) языке(ах) на основе владения взаимосвязанными и взаимозависимыми видами репродуктивной и продуктивной иноязычной речевой деятельности, такими как аудирование, говорение, чтение, письмо и перевод в повседневной, социокультурной, учебно-	УК-4.3 Осуществляет поиск необходимой информации для решения стандартных коммуникативных задач на русском и иностранном языках;

Шифр	Компетенция	Индикаторы достижения компетенции (в рамках данной дисциплины)
	профессиональной, официально-деловой и научной сферах общения	
ОПК-1	Способен применять в профессиональной деятельности методы математического анализа, логики и моделирования, теоретического и экспериментального исследования в информатике, лингвистике и гуманитарных науках	ОПК-1.2 Применяет знания теоретического и экспериментального исследования в информатике, лингвистике и гуманитарных науках;

3. МЕСТО ДИСЦИПЛИНЫ В СТРУКТУРЕ ОП ВО

Дисциплина «Введение в корпусную лингвистику» относится к обязательной части блока 1 «Дисциплины (модули)» образовательной программы высшего образования.

В рамках образовательной программы высшего образования обучающиеся также осваивают другие дисциплины и/или практики, способствующие достижению запланированных результатов освоения дисциплины «Введение в корпусную лингвистику».

Таблица 3.1. Перечень компонентов ОП ВО, способствующих достижению запланированных результатов освоения дисциплины

Шифр	Наименование компетенции	Предшествующие дисциплины/модули, практики*	Последующие дисциплины/модули, практики*
УК-12	Способен: искать нужные источники информации и данные, воспринимать, анализировать, запоминать и передавать информацию с использованием цифровых средств, а также с помощью алгоритмов при работе с полученными из различных источников данными с целью эффективного использования полученной информации для решения задач; проводить оценку информации, ее достоверность, строить логические умозаключения на основании поступающих информации и данных	История России; Иностранный язык**; Введение в языкознание; Теория перевода; Методы исследований в лингвистике и переводе**; Методы исследований в истории и литературе**; Ознакомительная практика;	Прикладные аспекты применения кибертехнологий в лингвистике и переводе**; Прикладные аспекты применения кибертехнологий в истории и литературе**; Информационное обеспечение кибертехнологий в гуманитарной сфере; Иностранный язык**; Введение в семиотику; Русский язык как иностранный в профессиональных целях**; Иностранный язык (основной) в профессиональной деятельности**; Разработка проектной документации: практический курс; Технологическая практика (учебная); Научно-исследовательская работа; Преддипломная практика; Проектно-технологическая практика;
УК-4	Способен к коммуникации в межличностном и межкультурном взаимодействии на русском как иностранном и иностранном(ых) языке(ах) на основе владения взаимосвязанными и взаимозависимыми видами репродуктивной и продуктивной иноязычной речевой деятельности,	Иностранный язык**; Русский язык как иностранный**; Деловая коммуникация (для иностранных студентов); Деловая коммуникация (для российских студентов); Введение в языкознание; Второй иностранный язык (практический курс); Педагогика и психология; Практики публичных	Иностранный язык**; Русский язык как иностранный**; Основы риторики и основы публичных выступлений; Деловая коммуникация (для иностранных студентов); Деловая коммуникация (для российских студентов); Иностранный язык (основной) в профессиональной

Шифр	Наименование компетенции	Предшествующие дисциплины/модули, практики*	Последующие дисциплины/модули, практики*
	такими как аудирование, говорение, чтение, письмо и перевод в повседневной, социокультурной, учебно-профессиональной, официально-деловой и научной сферах общения	выступлений; Методы исследований в лингвистике и переводе**;	деятельности**; Русский язык как иностранный в профессиональных целях**; Второй иностранный язык (практический курс); Русский язык и культура речи (в интернет пространстве); Разработка проектной документации: практический курс;
ОПК-1	Способен применять в профессиональной деятельности методы математического анализа, логики и моделирования, теоретического и экспериментального исследования в информатике, лингвистике и гуманитарных науках	Ознакомительная практика; Математическое обеспечение кибертехнологий в гуманитарной сфере; Математическое обеспечение кибертехнологий в гуманитарной сфере (избранные разделы); Теория перевода;	Технологическая практика (учебная); Научно-исследовательская работа; Преддипломная практика; Проектно-технологическая практика; Математическое обеспечение кибертехнологий в гуманитарной сфере; Программное обеспечение кибертехнологий в гуманитарной сфере; Введение в семиотику;

* - заполняется в соответствии с матрицей компетенций и СУП ОП ВО

** - элективные дисциплины /практики

4. ОБЪЕМ ДИСЦИПЛИНЫ И ВИДЫ УЧЕБНОЙ РАБОТЫ

Общая трудоемкость дисциплины «Введение в корпусную лингвистику» составляет «2» зачетные единицы

Таблица 4.1. Виды учебной работы по периодам освоения образовательной программы высшего образования для очной формы обучения.

Вид учебной работы	ВСЕГО, ак.ч.		Семестр(-ы)
			3
Контактная работа, ак.ч	34		34
Лекции (ЛК)	17		17
Лабораторные работы (ЛР)	0		0
Практические/семинарские занятия (СЗ)	17		17
Самостоятельная работа обучающихся, ак.ч.	20		20
Контроль (экзамен/зачет с оценкой), ак.ч.	18		18
Общая трудоемкость дисциплины ак.ч.	ак.ч.	72	72
	зач.ед.	2	2

Общая трудоемкость дисциплины «Введение в корпусную лингвистику» составляет «2» зачетные единицы

Таблица 4.2. Виды учебной работы по периодам освоения образовательной программы высшего образования для очно-заочной формы обучения.

Вид учебной работы	ВСЕГО, ак.ч.		Семестр(-ы)
			4
Контактная работа, ак.ч	17		17
Лекции (ЛК)	17		17
Лабораторные работы (ЛР)	0		0
Практические/семинарские занятия (СЗ)	0		0
Самостоятельная работа обучающихся, ак.ч.	37		37
Контроль (экзамен/зачет с оценкой), ак.ч.	18		18
Общая трудоемкость дисциплины ак.ч.	ак.ч.	72	72
	зач.ед.	2	2

5. СОДЕРЖАНИЕ ДИСЦИПЛИНЫ

Таблица 5.1. Содержание дисциплины (модуля) по видам учебной работы*

Номер раздела	Наименование раздела дисциплины	Наименование темы		Содержание темы	Вид учебной работы*
Раздел 1	Введение в дисциплину. Практическая значимость использования корпусных данных.	1.1	Основные понятия корпусной лингвистики.	Основные понятия корпусной лингвистики: проблемная область, корпус данных и корпус текстов, единица хранения корпуса данных / текстов.	ЛК, СЗ
		1.2	Основные характеристики корпусов.	Общая характеристика лингвистических корпусов. Машинные фонды, базы данных, корпус. Основные характеристики корпусов. Понятие репрезентативности корпусов. Понятие сбалансированности корпусов. Машинные фонды, базы данных, корпус. Общие черты и отличия. Предпосылки создания корпусов на основе машинных фондов, баз данных. Машинный фонд vs корпус. База данных vs корпус.	ЛК, СЗ
		1.3	Преимущества современных корпусных исследований. Практическое применение корпусных данных.	Качественные и количественные данные в корпусе. Получение корпусных примеров: как и зачем. Программы для работы с корпусом. Лексическое разнообразие текста. Исследование нормы и узуса. Примеры частотных словарей.	ЛК, СЗ
Раздел 2	Анализ опыта практического использования наиболее актуальных лингвистических корпусов различных языков.	2.1	Авторитетные зарубежные корпуса: обзор ресурсов, поисковые возможности. Проекты на базе корпусов.	Фундаментальные корпуса английского языка, французского языка, немецкого языка, китайского языка и др. Британский национальный корпус (БНК), Исторический корпус американского английского, Немецкий справочный корпус (НСК), Чешский национальный корпус и др.: особенности их построения и использования.	ЛК, СЗ
		2.2	Корпусы русского языка: обзор ресурсов, поисковые возможности. Проекты на базе национального корпуса русского языка.	Корпусы текстов русского языка: Особенности их построения и использования. Национальный корпус русского языка (НКРЯ); Русские корпусные проекты: Упсальский корпус, проекты Машинного фонда РЯ, Тюбингенский корпус, Хельсинкский аннотированный корпус (ХАНКО), Корпус газетных текстов и др.	ЛК, СЗ
Раздел 3	Разметка корпусов.	3.1	Лингвистическая разметка. Морфологическая разметка.	Разметка. Виды разметок. Лингвистическая аннотация и метаданные. Понятие разметки. Программы автоматической разметки корпусов. Морфологическая разметка.	ЛК, СЗ
		3.2	Синтаксическая разметка. Семантическая разметка.	Разметка. Виды разметок. Лингвистическая аннотация и метаданные. Понятие синтаксической и семантической разметки. Программы автоматической разметки корпусов. Глубина разметки	ЛК, СЗ
Раздел 4	Технология создания корпусов.	4.1	Принципы составления корпусов.	Этапы формирования корпуса. Методология создания лингвистического корпуса. Автоматическая обработка текста и разметка корпуса. Уровни разметки в лингвистических корпусах. Процессы токенизации, лемматизации, стемминга и парсинга: специфика, цели, результаты.	ЛК, СЗ
Раздел 5	Компьютерные программы, используемые для работы с корпусом.	5.1	Корпусные менеджеры.	Корпусный менеджер как поисковая система. Основные характеристики корпусного менеджера.	ЛК, СЗ
		5.2	Сравнение корпусных менеджеров.	Web-интерфейсы к корпусам. Конкордансер AntConc. Корпусные менеджеры и их сравнительная характеристика.	ЛК, СЗ
Раздел 6	Практическое применение возможностей использования корпуса в различных дисциплинах.	6.1	Корпусная лексикология и лексикография.	Лингвистические исследования, проводимые на базе текстовых корпусов. Применение текстовых корпусов в лексикографии. Лингвистические исследования на базе корпуса: изучение лексики. Частотный словарь. Структурно-вероятностная модель текста. Функции частотного словаря. Единица описания. Понятие токена, леммы, словоформы.	ЛК, СЗ

Номер раздела	Наименование раздела дисциплины	Наименование темы	Содержание темы	Вид учебной работы*
			Принципы составления частотных словарей.	
		6.2	Корпусная грамматика. Лингвистические исследования на базе корпуса: исследование грамматики. Преимущества корпусных грамматик. Виды корпусных грамматик.	ЛК, СЗ
		6.3	Исследования дискурса, основанные на корпусах. Особенности мультимедийных корпусов: описание данных, проблемная область, организация поиска. Проблема автоматизированного выявления коллокаций.	ЛК, СЗ

* - заполняется только по ОЧНОЙ форме обучения: ЛК – лекции; ЛР – лабораторные работы; СЗ – практические/семинарские занятия.

6. МАТЕРИАЛЬНО-ТЕХНИЧЕСКОЕ ОБЕСПЕЧЕНИЕ ДИСЦИПЛИНЫ

Таблица 6.1. Материально-техническое обеспечение дисциплины

Тип аудитории	Оснащение аудитории	Специализированное учебное/лабораторное оборудование, ПО и материалы для освоения дисциплины (при необходимости)
Лекционная	Аудитория для проведения занятий лекционного типа, оснащенная комплектом специализированной мебели; доской (экраном) и техническими средствами мультимедиа презентаций.	
Компьютерный класс	Компьютерный класс для проведения занятий, групповых и индивидуальных консультаций, текущего контроля и промежуточной аттестации, оснащенная персональными компьютерами (в количестве 12 шт.), доской (экраном) и техническими средствами мультимедиа презентаций.	
Семинарская	Аудитория для проведения занятий семинарского типа, групповых и индивидуальных консультаций, текущего контроля и промежуточной аттестации, оснащенная комплектом специализированной мебели и техническими средствами мультимедиа презентаций.	
Для самостоятельной работы	Аудитория для самостоятельной работы обучающихся (может использоваться для проведения семинарских занятий и консультаций), оснащенная комплектом специализированной мебели и компьютерами с доступом в ЭИОС.	

* - аудитория для самостоятельной работы обучающихся указывается **ОБЯЗАТЕЛЬНО!**

7. УЧЕБНО-МЕТОДИЧЕСКОЕ И ИНФОРМАЦИОННОЕ ОБЕСПЕЧЕНИЕ ДИСЦИПЛИНЫ

Основная литература:

1. Захаров, В. П. Корпусная лингвистика : учебник / В. П. Захаров, С. Ю. Богданова. — 3-е изд., перераб. — Санкт-Петербург : СПбГУ, 2020. — 234 с. — ISBN 978-5-288-05997-1. — Текст : электронный // Лань : электронно-библиотечная система. — URL: <https://e.lanbook.com/book/139576> (дата обращения: 14.04.2026). — Режим доступа: для авториз. пользователей.

- 2. Шунейко, А. А. Корпусная лингвистика: учебник для вузов / А. А. Шунейко. — Москва: Издательство Юрайт, 2025. — 222 с. — (Высшее образование). — ISBN 978-5- 534-13603-6. — Текст: электронный // Образовательная платформа Юрайт [сайт]. — URL: <https://urait.ru/bcode/567396> (дата обращения: 13.04.2025).

- 3. Грудева, Е. В. Корпусная лингвистика: учебное пособие / Е. В. Грудева. — 3-е изд. — Москва: ФЛИНТА, 2017. — 165 с. — ISBN 978-5-9765-1497-3. — Текст: электронный // Лань: электронно-библиотечная система. — <https://e.lanbook.com/book/504928> (дата обращения: 14.04.2026).

Дополнительная литература:

1. Копотев М., Введение в корпусную лингвистику. Учебное пособие для студентов филологических и лингвистических специальностей университетов. Прага, Animedia Company, 2014.

2. Gries, S. T. The Four Central Corpus Linguistic Methods. In: Gries, S. T. Quantitative Corpus Linguistics with R: A Practical Introduction. Milton Park, Abingdon, Oxon: Routledge, 2017. eBook Collection (EBSCOhost), EBSCOhost (Доступна электронная версия).

Ресурсы информационно-телекоммуникационной сети «Интернет»:

1. ЭБС РУДН и сторонние ЭБС, к которым студенты университета имеют доступ на основании заключенных договоров

- Электронно-библиотечная система РУДН – ЭБС РУДН <http://lib.rudn.ru/MegaPro/Web>

- ЭБС «Университетская библиотека онлайн» <http://www.biblioclub.ru>
- ЭБС Юрайт <http://www.biblio-online.ru>
- ЭБС «Консультант студента» www.studentlibrary.ru
- ЭБС «Троицкий мост»

2. Базы данных и поисковые системы

- электронный фонд правовой и нормативно-технической документации <http://docs.cntd.ru/>
- поисковая система Яндекс <https://www.yandex.ru/>
- Sage <https://journals.sagepub.com/>
- Springer Nature Link <https://link.springer.com/>
- Wiley Journal Database <https://onlinelibrary.wiley.com/>
- поисковая система Google <https://www.google.ru/>
- реферативная база данных SCOPUS <http://www.elsevier.com/locate/scopus/>

Учебно-методические материалы для самостоятельной работы обучающихся при освоении дисциплины/модуля:*

1. Курс лекций по дисциплине «Введение в корпусную лингвистику».

* - все учебно-методические материалы для самостоятельной работы обучающихся размещаются в соответствии с действующим порядком на странице дисциплины **в ТУИС!**

РАЗРАБОТЧИКИ

Доцент кафедры русского языка № 1

Должность

РАЗРАБОТЧИКИ

Старший преподаватель кафедры русского языка
№ 1

Должность

РУКОВОДИТЕЛЬ БУП

Заведующий кафедрой русского языка № 1

Должность

РУКОВОДИТЕЛЬ ОП ВО

Заведующий кафедрой русского языка № 1

Должность

Астахова Я.А.

Фамилия И.О

Лонцкая М.А.

Фамилия И.О

Брагина М.А.

Фамилия И.О

Брагина М.А.

Фамилия И.О