

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

**Институт русского языка**

---

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

## **РАБОЧАЯ ПРОГРАММА ДИСЦИПЛИНЫ**

**Введение в корпусную лингвистику**

---

(наименование дисциплины/модуля)

**Рекомендована МСЧН для направления подготовки/специальности:**

**45.03.01 Филология**

---

(код и наименование направления подготовки/специальности)

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

**Прикладная цифровая филология**

---

(наименование (профиль/специализация) ОП ВО)

**2022 г.**

## 1. ЦЕЛЬ ОСВОЕНИЯ ДИСЦИПЛИНЫ

Целью освоения дисциплины «Введение в корпусную лингвистику» является формирование у обучающихся теоретических знаний в области создания и применения лингвистических корпусов, приобретение практических умений работы с корпусами и связанными с ними компьютерными инструментами в сфере цифровой филологии и цифровой педагогики.

## 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-1.5	Анализирует и контекстно обрабатывает информацию для решения поставленных задач	Знать: алгоритм и принципы трансформации данных в аналитические выводы; Уметь: осуществлять обработку информации в соответствии с поставленными задачами и формировать собственное мнение на основе фактов; Владеть: методами поиска информации в различных источниках и обработки информации; основными положениями, теоретическими принципами и методологическими принципами логики
УК-6.3	Находит и использует источники получения дополнительной информации для повышения уровня общих и профессиональных знаний с учетом условий, средств, личностных возможностей, этапов карьерного роста, временной перспективы развития деятельности и требований рынка труда	Знать: алгоритмы поиска дополнительной информации в целях повышения собственного уровня знаний; Уметь: осуществлять поиск дополнительной информации для повышения уровня общих и профессиональных знаний; Владеть: готовностью находить и использовать источники получения дополнительной информации для повышения уровня общих и профессиональных знаний
УК-12.2	Проводит оценку информации, ее достоверность, строит логические умозаключения на основании поступающих информации и данных	Знать методы и инструменты оценки показателей достоверности источников информации и данных. Уметь анализировать, сравнивать и критически оценивать источники информации и информацию, их достоверность и релевантность; оценивать практическую значимость результатов поиска и выделять в ней необходимые данные для анализа. Владеть цифровыми инструментами и сервисами для проверки достоверности информации; навыками критического фильтрования информации используемых систем
ОПК-1.2	Осуществляет первичный отбор и анализ языкового и (или) литературного материала	Знать: приемы и методы анализа языкового и (или) литературного материала; принципы формирования запроса при поиске информации. Уметь: осуществлять первичный анализ языкового и (или) литературного материала в соответствии с целями и задачами отбора. Владеть: навыками проведения анализа отобранного языкового и (или) литературного материала.
ОПК-1.3	Обладает навыками анализа и интерпретации филологических проблем в диахроническом и синхроническом аспектах	Знать: последовательность развития отдельных языковых явлений и процесс развития языка в целом; проблемы современного русского языка в диахроническом и синхроническом аспектах. Уметь:

Шифр	Компетенция	Индикаторы достижения компетенции (в рамках данной дисциплины)
		идентифицировать проблемы современного русского языка, анализировать их в диахроническом и историческом аспекте, в том числе как часть системы русского языка. Владеть: навыком анализа и интерпретации филологических проблем в диахроническом и синхроническом аспектах.
ОПК-4.1	Владеет методикой сбора и филологического анализа языковых и литературных фактов и текстов разной степени сложности	Знать: базовые понятия современной филологии в их историческом развитии, теоретическом и методологическом аспектах; базовые методики сбора и анализа языкового материала и литературных фактов, а также интерпретации текстов различных типов; Уметь: осуществлять адекватную репрезентацию результатов анализа собранных языковых и литературных фактов; производить лингвистический и лингвостилистический анализ художественного и иных типов текста, интерпретировать тексты различных типов, основываясь на результатах анализа текста. Владеть: методиками сбора и анализа языковых и литературных фактов и интерпретации текстов различных типов
ПК-4.1	Владеет базовыми навыками поиска и извлечения графической и/или текстовой информации с учетом требований законодательства Российской Федерации в области интеллектуальной собственности и ограничений, касающихся использования информационных материалов в сети Интернет	Знать: нормативные требования к стилевым особенностям и функционально-смысловым типам речи; принципы поиска и использования информации с учетом требований законодательства; Уметь: осуществлять поиск информации и ее извлечение в соответствии с поставленными целями, использование информации с учетом требований законодательства Российской Федерации в области интеллектуальной собственности; Уметь: осуществлять поиск и обработку информации в соответствии с целеполаганием и учетом правовых норм Российской Федерации; Владеть: способностью обрабатывать найденную информацию с учетом профессиональных целей и действующего законодательства Российской Федерации
ПК-5.1	Владеет навыками контроля качества текстов	Знать: основы стилистики, редактирования и корректуры; Уметь: осуществлять проверку текстов различных стилей на соответствие нормам функциональных стилей изучаемого литературного языка; Владеть: способностью осуществлять проверку и контроль качества текстов различной стилистической принадлежности
ПК-5.2	Владеет навыками работы со словарями и справочниками, электронными информационными ресурсами в избранной сфере профессиональной, в том числе педагогической, деятельности	Знать: типы нормативных словарей современного языка и информационных ресурсов для профессиональной деятельности и принципы работы с ними; Уметь: осуществлять выбор нормативного словаря и/или информационного ресурса (в том числе электронного) для решения профессиональной задачи; Владеть: навыками работы со словарями и справочниками, электронными информационными ресурсами в избранной сфере профессиональной деятельности

В результате освоения дисциплины обучающийся должен:

*знать:*

- ключевые понятия, актуальные проблемы, перспективные направления развития корпусной лингвистики;
- основные типы корпусов, наиболее известные и широко применяемые иноязычные и русскоязычные корпусы;

- способы использования корпусов для различных целей;
- технологию создания корпусов, основные форматы и способы организации корпусов;

*уметь:*

- определять тип и способы применения незнакомых корпусов;
- осваивать новые компьютерные инструменты;
- интерпретировать данные, полученные из корпусов;
- сравнивать и оценивать функциональность корпусов и инструментов по результатам работы с ними;

*владеть навыками*

- работы с несколькими существующими иноязычными и русскоязычными корпусами и корпусными менеджерами;
- формулирования сложных лингвистических запросов с использованием специальных обозначений (тэгов) и регулярных выражений, разметки текстов и чтения разметки;
- проектирования собственного корпуса;
- создания небольших специализированных корпусов.

### 3. МЕСТО ДИСЦИПЛИНЫ В СТРУКТУРЕ ОП ВО

Дисциплина «Введение в корпусную лингвистику» относится к обязательной части (базовая компонента) Б1.О.01.15.

В рамках ОП ВО 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*

<b>Шифр</b>	<b>Наименование компетенции</b>	<b>Предшествующие дисциплины/модули, практики*</b>	<b>Последующие дисциплины/модули, практики*</b>
УК-6.3	Находит и использует источники получения дополнительной информации для повышения уровня общих и профессиональных знаний с учетом условий, средств, личностных возможностей, этапов карьерного роста, временной перспективы развития деятельности и требований рынка труда	Типологическая лингвистика	
УК-12.2	Проводит оценку информации, ее достоверность, строит логические умозаключения на основании поступающих информации и данных		
ОПК-1.1	Знает краткую историю основных этапов филологии, её современное состояние и перспективы развития		
ОПК-1.2	Осуществляет первичный отбор и анализ языкового и (или) литературного материала		
ОПК-1.3	Обладает навыками анализа и интерпретации филологических проблем в диахроническом и синхроническом аспектах		

Шифр	Наименование компетенции	Предшествующие дисциплины/модули, практики*	Последующие дисциплины/модули, практики*
ПК- 5.2	Владеет навыками работы со словарями и справочниками, электронными информационными ресурсами в избранной сфере профессиональной, в том числе педагогической, деятельности		
ПК-9.1	Решает стандартные задачи профессиональной, в том числе педагогической, деятельности с использованием специализированных цифровых инструментов и технологий с учетом требований информационной безопасности		
УК-6.3	Находит и использует источники получения дополнительной информации для повышения уровня общих и профессиональных знаний с учетом условий, средств, личностных возможностей, этапов карьерного роста, временной перспективы развития деятельности и требований рынка труда		
УК-12.1	Ищет нужные источники информации и данные, воспринимает, анализирует, запоминает и передает 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		
ОПК-7.2	Применяет современные цифровые методы и технологии анализа, обработки и представления информации в профессиональной, в том числе педагогической, сфере деятельности		
ПК-9.1	Решает стандартные задачи в профессиональной, в том числе педагогической, деятельности с использованием специализированных цифровых инструментов и технологий с учетом требований информационной безопасности		Инструменты обработки и визуализации данных
ПК-9.2	Использует цифровые технологии и высокоуровневые языки программирования общего назначения для разработки алгоритмов и программ в области интеллектуального анализа данных		
ПК-8.2	Применяет основные законы естественнонаучных дисциплин в профессиональной в том числе педагогической, деятельности, использует методы математическо-статистического анализа, логики и моделирования, теоретического и экспериментального исследования в гуманитарной сфере, в том числе во взаимосвязи с основными филологическими законами и методами		

<b>Шифр</b>	<b>Наименование компетенции</b>	<b>Предшествующие дисциплины/модули, практики*</b>	<b>Последующие дисциплины/модули, практики*</b>
ПК-10.2	Применяет, комбинирует и адаптирует существующие информационно-коммуникационные технологии для решения задач в профессиональной, в том числе педагогической, деятельности с учетом требований информационной безопасности		
УК-1.3	Осуществляет поиск информации для решения поставленной задачи по различным типам запросов		
УК-6.3	Находит и использует источники получения дополнительной информации для повышения уровня общих и профессиональных знаний с учетом условий, средств, личностных возможностей, этапов карьерного роста, временной перспективы развития деятельности и требований рынка труда		
УК-12.1	Ищет нужные источники информации и данные, воспринимает, анализирует, запоминает и передает 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		
УК-12.2	Проводит оценку информации, ее достоверность, строит логические умозаключения на основании поступающих информации и данных		
ОПК-2.2	Обладает навыками филологического анализа и интерпретации языкового материала, различных типов текстов		
ОПК-3.3	Корректно осуществляет библиографический поиск в традиционных и электронных источниках и составляет корректное описание на основе полученных результатов		Информационные системы анализа данных в гуманитарной сфере
ОПК-4.1	Владеет методикой сбора и филологического анализа языковых и литературных фактов и текстов разной степени сложности		
ОПК-7.1	Знает основные цифровые технологии, принципы их работы и методы для изучения и моделирования объектов профессиональной, в том числе педагогической, деятельности, анализа данных, представления информации и пр.		
ОПК-7.2	Применяет современные цифровые методы и технологии анализа, обработки и представления информации в профессиональной, в том числе педагогической, сфере деятельности		
ПК-9.1	Решает стандартные задачи в профессиональной, в том числе педагогической, деятельности с		

Шифр	Наименование компетенции	Предшествующие дисциплины/модули, практики*	Последующие дисциплины/модули, практики*
	использованием специализированных цифровых инструментов и технологий с учетом требований информационной безопасности		
ПК-9.2	Использует цифровые технологии и высокоуровневые языки программирования общего назначения для разработки алгоритмов и программ в области интеллектуального анализа данных		
ПК-8.2	Применяет основные законы естественнонаучных дисциплин в профессиональной в том числе педагогической, деятельности, использует методы математическо-статистического анализа, логики и моделирования, теоретического и экспериментального исследования в гуманитарной сфере, в том числе во взаимосвязи с основными филологическими законами и методами		
ПК-10.2	Применяет, комбинирует и адаптирует существующие информационно-коммуникационные технологии для решения задач в профессиональной, в том числе педагогической, деятельности с учетом требований информационной безопасности		

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

#### 4. ОБЪЕМ ДИСЦИПЛИНЫ И ВИДЫ УЧЕБНОЙ РАБОТЫ

Общая трудоемкость дисциплины «Введение в корпусную лингвистику» составляет 4 зачетных единицы.

Таблица 4.1. Виды учебной работы по периодам освоения ОП ВО для **ОЧНОЙ** формы обучения

Вид учебной работы	ВСЕГО, ак.ч.	Семестр(-ы)			
		6			
<i>Контактная работа, ак.ч.</i>	48	48			
Лекции (ЛК)	16	16			
Лабораторные работы (ЛР)					
Практические/семинарские занятия (СЗ)	32	32			
<i>Самостоятельная работа обучающихся, ак.ч.</i>	78	78			
<i>Контроль (экзамен/зачет с оценкой), ак.ч.</i>	18	18			
<b>Общая трудоемкость дисциплины</b>	ак.ч.	<b>144</b>	<b>144</b>		
	зач.ед.	<b>4</b>	<b>4</b>		

Таблица 4.2. Виды учебной работы по периодам освоения ОП ВО для ЗАОЧНОЙ формы обучения\*

Вид учебной работы	ВСЕГО, ак. ч.	Семестр(-ы)			
		7			
Контактная работа, ак. ч.	16	16			
Лекции (ЛК)	8	8			
Лабораторные работы (ЛР)					
Практические/семинарские занятия (СЗ)	8	8			
Самостоятельная работа обучающихся, ак. ч.	110	110			
Контроль (экзамен/зачет с оценкой), ак. ч.	18	18			
Общая трудоемкость дисциплины	ак. ч.	144	144		
	зач. ед.	4	4		

\* - заполняется в случае реализации программы в заочной форме

## 5. СОДЕРЖАНИЕ ДИСЦИПЛИНЫ

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

Наименование раздела дисциплины	Содержание раздела (темы)	Вид учебной работы*
Раздел 1. Введение в дисциплину. Практическая значимость использования корпусных данных.	Тема 1.1. Основные понятия корпусной лингвистики.	ЛК, СЗ
	Тема 1.2. Основные характеристики корпусов.	
	Тема 1.3. Преимущества современных корпусных исследований. Практическое применение корпусных данных.	
Раздел 2. Анализ опыта практического использования наиболее актуальных лингвистических корпусов различных языков.	Тема 2.1. Авторитетные зарубежные корпусы: обзор ресурсов, поисковые возможности. Проекты на базе корпусов.	ЛК, СЗ
	Тема 2.2. Корпусы русского языка: обзор ресурсов, поисковые возможности. Проекты на базе национального корпуса русского языка.	
Раздел 3. Разметка корпусов.	Тема 3.1. Лингвистическая разметка. Морфологическая разметка.	ЛК, СЗ
	Тема 3.2. Синтаксическая разметка. Семантическая разметка.	
Раздел 4. Технология создания корпусов.	Тема 4.1. Принципы составления корпусов.	ЛК, СЗ
	Тема 4.2. Технологии создания корпусов.	
	Тема 4.3. Создание корпусов на базе веба.	
Раздел 5. Компьютерные программы, используемые для работы с корпусом.	Тема 5.1. Корпусные менеджеры.	ЛК, СЗ
	Тема 5.2. Сравнение корпусных менеджеров.	
Раздел 6. Практическое применение возможностей	Тема 6.1. Корпусная лексикология и лексикография.	ЛК, СЗ

использования корпуса в различных дисциплинах	Тема 6.2. Корпусная грамматика.
	Тема 6.3. Исследования дискурса, основанные на корпусах.

\* - заполняется только по **ОЧНОЙ** форме обучения: ЛК – лекции; ЛР – лабораторные работы; СЗ – 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.	Проектор и/или большой экран/монитор, компьютер, ОС MS Windows 10, MS Office 2016 или MS Office 2019 или MS Office 365, портал Microsoft Azure, доступ к интернету
Лаборатория	Аудитория для проведения лабораторных работ, индивидуальных консультаций, текущего контроля и промежуточной аттестации, оснащенная комплектом специализированной мебели и оборудованием.	Проектор и/или большой экран/монитор, компьютер, ОС MS Windows 10, MS Office 2016 или MS Office 2019 или MS Office 365, портал Microsoft Azure, доступ к интернету
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.	Проектор и/или большой экран/монитор, компьютер, ОС MS Windows 10, MS Office 2016 или MS Office 2019 или MS Office 365, портал Microsoft Azure, доступ к интернету
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.	Проектор и/или большой экран/монитор, компьютер, ОС MS Windows 10, MS Office 2016 или MS Office 2019 или MS Office 365, портал Microsoft Azure, доступ к интернету
Для самостоятельной	Аудитория для самостоятельной работы обучающихся (может использоваться для	Проектор и/или большой экран/монитор,

Тип аудитории	Оснащение аудитории	Специализированное учебное/лабораторное оборудование, ПО и материалы для освоения дисциплины (при необходимости)
работы обучающихся	проведения семинарских занятий и консультаций), оснащенная комплектом специализированной мебели и компьютерами с доступом в ЭИОС.	компьютер, ОС MS Windows 10, MS Office 2016 или MS Office 2019 или MS Office 365, портал Microsoft Azure, доступ к интернету

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

## 7. УЧЕБНО-МЕТОДИЧЕСКОЕ И ИНФОРМАЦИОННОЕ ОБЕСПЕЧЕНИЕ ДИСЦИПЛИНЫ

*Основная литература:*

1. Захаров В. П., Богданова С. Ю. Корпусная лингвистика: учебник. 3-е изд., перераб. — СПб.:Изд-во С.-Петербур. ун-та, 2020. — 234 с. ISBN 978-5-288-05997-1
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).

*Дополнительная литература:*

1. Копотев М.В. Введение в корпусную лингвистику: Учебное пособие для студентов филологических и лингвистических специальностей университетов / М. В. Копотев — «Animedia», 2014
2. Плунгян В. Корпусная лингвистика // ПостНаука [Электронный ресурс]. 2013. URL: <https://postnauka.ru/video/7783>

*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](http://www.studentlibrary.ru)

- ЭБС «Лань» <http://e.lanbook.com/>

- 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/>

- поисковая система Google <https://www.google.ru/>

- реферативная база данных SCOPUS  
<http://www.elsevierscience.ru/products/scopus/>

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

1. Курс лекций по дисциплине «Информационные системы анализа данных в гуманитарной сфере».

2. Презентационные материалы по дисциплине «Информационные системы анализа данных в гуманитарной сфере»

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

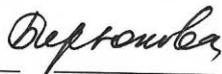
## **8. ОЦЕНОЧНЫЕ МАТЕРИАЛЫ И БАЛЛЬНО-РЕЙТИНГОВАЯ СИСТЕМА ОЦЕНИВАНИЯ УРОВНЯ СФОРМИРОВАННОСТИ КОМПЕТЕНЦИЙ ПО ДИСЦИПЛИНЕ**

Оценочные материалы и балльно-рейтинговая система\* оценивания уровня сформированности компетенций (части компетенций) по итогам освоения дисциплины «Информационные системы анализа данных в гуманитарной сфере» представлены в Приложении к настоящей Рабочей программе дисциплины.

\* - ОМ и БРС формируются на основании требований соответствующего локального нормативного акта РУДН.

**РАЗРАБОТЧИКИ:**

**доцент кафедры русского языка №5  
Института русского языка**



**Бирюкова Ю.Н.**

Должность, БУП

Подпись

Фамилия И.О.

**РУКОВОДИТЕЛЬ ОУП:**

**Директор  
Института русского языка**



**Должикова А.В.**

Наименование БУП

Подпись

Фамилия И.О.

**РУКОВОДИТЕЛЬ ОП ВО:**

**Заведующий кафедрой  
русского языка № 5  
Института русского языка**



**Куриленко В.Б.**

Должность, БУП

Подпись

Фамилия И.О.