

Министерство науки и высшего образования Российской Федерации
ФГБОУ ВО «БАЙКАЛЬСКИЙ ГОСУДАРСТВЕННЫЙ УНИВЕРСИТЕТ»

УТВЕРЖДАЮ

Проректор по учебной работе
д.юр.н., доц. Васильева Н.В.



21.06.2024г.

Рабочая программа дисциплины
Б1.О.8. Корпусная лингвистика

Направление подготовки: 45.04.02 Лингвистика

Направленность (профиль): Теория и методика преподавания русского языка
как иностранного

Квалификация выпускника: магистр

Форма обучения: очная

Курс	2
Семестр	21
Лекции (час)	14
Практические (сем, лаб.) занятия (час)	14
Самостоятельная работа, включая подготовку к экзаменам и зачетам (час)	80
Курсовая работа (час)	
Всего часов	108
Зачет (семестр)	21
Экзамен (семестр)	

Иркутск 2024

Программа составлена в соответствии с ФГОС ВО по направлению 45.04.02
Лингвистика.

Автор А.А. Федосеев

Рабочая программа обсуждена и утверждена на заседании кафедры русского
языка и межкультурной коммуникации

Заведующий кафедрой А.М. Литовкина

1. Цели изучения дисциплины

Цель курса – познакомить студентов с наиболее известными компьютерными ресурсами – национальными корпусами, компьютерными словарями, программами обработки текстов, с основными принципами их устройства, методами исследований в области гуманитарных наук, основанных на анализе текстов

- Задачи курса
- В соответствии с поставленной целью в рамках курса стоят следующие задачи:
- познакомить студентов с основными типами языковых корпусов;
 - на примерах показать, какие новые возможности в исследовании грамматики и лексики языка дает использование корпусных, а также применение современных методов обработки этих данных;
 - дать представления о проблемах создания корпусов, об основных принципах разработки данных ресурсов и об основных требованиях, предъявляемых к ним;
 - познакомить с технологиями и проблемами разметки корпусов;
 - научить работать с современными пакетами обработки собственных корпусов;
 - научить применять специальные методы лингвистических исследований, использующие данные корпусов, в том числе и статистические методы исследования;
 - познакомить с некоторыми специальными методами лингвистических исследований, использующих данные корпусов, в том числе количественными методами исследования, методом контент-анализа;
-
- обучить практическим навыкам по применению корпусных методов в своей исследовательской работе.

2. Перечень планируемых результатов обучения по дисциплине (модулю), соотнесенных с планируемыми результатами освоения образовательной программы

Компетенции обучающегося, формируемые в результате освоения дисциплины

Код компетенции по ФГОС ВО	Компетенция
ОПК-7	Способен работать с основными информационно-поисковыми и экспертными системами, системами представления знаний и обработки вербальной информации

Структура компетенции

Компетенция	Формируемые ЗУНы
ОПК-7 Способен работать с основными информационно-поисковыми и экспертными системами, системами представления знаний и обработки вербальной информации	З. Знать информационные и экспертные системы и средства представления знаний и обработки вербальной информации У. Уметь выбирать релевантные информационно-поисковые и экспертные системы и средства представления знаний и обработки вербальной информации Н. Владеть навыками работы с информационными и экспертными системами и средствами поиска и обработки вербальной информации

3. Место дисциплины (модуля) в структуре образовательной программы

Принадлежность дисциплины - БЛОК 1 ДИСЦИПЛИНЫ (МОДУЛИ):
Обязательная часть.

4. Объем дисциплины (модуля) в зачетных единицах с указанием количества академических часов, выделенных на контактную работу обучающихся с преподавателем (по видам учебных занятий) и на самостоятельную работу обучающихся

Общая трудоемкость дисциплины составляет 3 зач. ед., 108 часов.

Вид учебной работы	Количество часов
Контактная(аудиторная) работа	
Лекции	14
Практические (сем, лаб.) занятия	14
Самостоятельная работа, включая подготовку к экзаменам и зачетам	80
Всего часов	108

5. Содержание дисциплины (модуля), структурированное по темам (разделам) с указанием отведенного на них количества академических часов и видов учебных занятий

5.1. Содержание разделов дисциплины

№ п/п	Раздел и тема дисциплины	Семестр	Лекции	Семинар Лаборат. Практич.	Самостоят. раб.	В интерактивной форме	Формы текущего контроля успеваемости
		21					
1	Корпусы и корпусная лингвистика как раздел языкознания. Основные понятия	21	2	2	10		Задание 1
2	Основные задачи и направления корпусной лингвистики. Взаимодействие корпусной лингвистики и компьютерной лингвистики	21	2	2	10		Задание 2
3	Классификация и характеристики корпусов	21	2	2	12		Задание 3
4	Принципы создания корпуса	21	2	2	12		Задание 4
5	Современные корпусы русского языка. Национальный корпус русского языка (НКРЯ)	21	4	4	24		Задание 5
6	Пользование корпусом. Лингвистические исследования на базе НКРЯ	21	2	2	12		Задание 6
	ИТОГО		14	14	80		

5.2. Лекционные занятия, их содержание

№ п/п	Наименование разделов и тем	Содержание
1	Корпусная лингвистика как раздел языкознания.	Основные понятия корпусной лингвистики. История создания лингвистических корпусов. Основные характеристики корпусов
2	Задачи и направления корпусной лингвистики. Корпусная и компьютерная лингвистика	Задачи, решаемые с помощью компьютерных ресурсов. Обзор различных языковых и текстовых ресурсов, доступных в Интернете.
3	Основные характеристики корпусов	Устные – письменные. Статические – динамические. Одноязычные — многоязычные
4	Создание корпуса	Что такое разметка? История систем разметки. Тэги как лингвистический инструмент. Текст с разных сторон. Автоматическая разметка текстов
5	Современные корпуса русского языка. Национальный корпус русского языка	Методы извлечения информации из корпуса. Типы извлекаемой информации.
6	Лингвистические исследования	Описания лексики. Исследования частей речи. Исследования синтаксических процессов на уровне предложения. Исследования прагматики и устной речи. Исследования вариаций языка

5.3. Семинарские, практические, лабораторные занятия, их содержание

№ раздела и темы	Содержание и формы проведения
1	Корпусная лингвистика как раздел языкознания.. Понятие лингвистического корпуса. История корпусной лингвистики.
2	Задачи и направления корпусной лингвистики. Корпусная и компьютерная лингвистика. Задачи и основные направления корпусной лингвистики. Корпусная лингвистика и компьютерная лингвистика.
3	Типы корпусов. Типы корпусов: устные и письменные, одноязычные и многоязычные
4	Типы корпусов. Типы корпусов: аннотированные и неаннотированные. Лингвистическая аннотация и метаданные
5	Национальный корпус русского языка. Методы извлечения информации из корпуса. Типы извлекаемой информации
6	Лингвистические исследования. Описания лексики. Исследования частей

№ раздела и темы	Содержание и формы проведения
	речи. Исследования синтаксических процессов на уровне предложения . Исследования прагматики и устной речи. Исследования вариаций языка

6. Фонд оценочных средств для проведения промежуточной аттестации по дисциплине (полный текст приведен в приложении к рабочей программе)

6.1. Текущий контроль

№ п/п	Этапы формирования компетенций (Тема из рабочей программы дисциплины)	Перечень формируемых компетенций по ФГОС ВО	(ЗУНы: (З.1...З.п, У.1...У.п, Н.1...Н.п))	Контрольные задания или иные материалы, необходимые для оценки знаний, умений, навыков и (или) опыта деятельности, характеризующих этапы формирования компетенций в процессе освоения образовательной программы (Наименование оценочного средства)	Описание показателей и критериев оценивания компетенций на различных этапах их формирования, описание шкал оценивания (по 100-балльной шкале)
1	1. Корпусы и корпусная лингвистика как раздел языкознания. Основные понятия	ОПК-7	З.Знать информационные и экспертные системы и средства представления знаний и обработки вербальной информации У.Уметь выбирать релевантные информационно-поисковые и экспертные системы и средства представления знаний и обработки вербальной информации Н.Владеть навыками работы с информационными и экспертными системами и средствами поиска и обработки вербальной информации	Задание 1	5 баллов за каждый правильный ответ (20)
2	2. Основные задачи и направления корпусной лингвистики. Взаимодействие корпусной лингвистики и компьютерной лингвистики	ОПК-7	З.Знать информационные и экспертные системы и средства представления знаний и обработки вербальной информации У.Уметь выбирать релевантные информационно-поисковые и экспертные системы и средства	Задание 2	5 баллов за правильный ответ (10)

№ п/п	Этапы формирования компетенций (Тема из рабочей программы дисциплины)	Перечень формируемых компетенций по ФГОС ВО	(ЗУНы: (З.1...З.п, У.1...У.п, Н.1...Н.п)	Контрольные задания или иные материалы, необходимые для оценки знаний, умений, навыков и (или) опыта деятельности, характеризующих этапы формирования компетенций в процессе освоения образовательной программы (Наименование оценочного средства)	Описание показателей и критериев оценивания компетенций на различных этапах их формирования, описание шкал оценивания (по 100-балльной шкале)
			представления знаний и обработки вербальной информации Н. Владеть навыками работы с информационными и экспертными системами и средствами поиска и обработки вербальной информации		
3	3. Классификация и характеристики корпусов	ОПК-7	З. Знать информационные и экспертные системы и средства представления знаний и обработки вербальной информации У. Уметь выбирать релевантные информационно-поисковые и экспертные системы и средства представления знаний и обработки вербальной информации Н. Владеть навыками работы с информационными и экспертными системами и средствами поиска и обработки вербальной информации	Задание 3	5 баллов за правильный ответ (10)
4	4. Принципы создания корпуса	ОПК-7	З. Знать информационные и экспертные системы и средства представления знаний и обработки вербальной информации У. Уметь выбирать релевантные информационно-поисковые и экспертные системы и средства	Задание 4	5 баллов за правильный ответ (10)

№ п/п	Этапы формирования компетенций (Тема из рабочей программы дисциплины)	Перечень формируемых компетенций по ФГОС ВО	(ЗУНы: З.1...З.п, У.1...У.п, Н.1...Н.п)	Контрольные задания или иные материалы, необходимые для оценки знаний, умений, навыков и (или) опыта деятельности, характеризующих этапы формирования компетенций в процессе освоения образовательной программы (Наименование оценочного средства)	Описание показателей и критериев оценивания компетенций на различных этапах их формирования, описание шкал оценивания (по 100-балльной шкале)
			представления знаний и обработки вербальной информации Н. Владеть навыками работы с информационными и экспертными системами и средствами поиска и обработки вербальной информации		
5	5. Современные корпусы русского языка. Национальный корпус русского языка (НКРЯ)	ОПК-7	З. Знать информационные и экспертные системы и средства представления знаний и обработки вербальной информации У. Уметь выбирать релевантные информационно-поисковые и экспертные системы и средства представления знаний и обработки вербальной информации Н. Владеть навыками работы с информационными и экспертными системами и средствами поиска и обработки вербальной информации	Задание 5	5 баллов за правильный ответ (20)
6	6. Пользование корпусом. Лингвистические исследования на базе НКРЯ	ОПК-7	З. Знать информационные и экспертные системы и средства представления знаний и обработки вербальной информации У. Уметь выбирать релевантные информационно-поисковые и экспертные системы и средства	Задание 6	5 баллов за правильный ответ (30)

№ п/п	Этапы формирования компетенций (Тема из рабочей программы дисциплины)	Перечень формируемых компетенций по ФГОС ВО	(ЗУНы: (З.1...З.п, У.1...У.п, Н.1...Н.п))	Контрольные задания или иные материалы, необходимые для оценки знаний, умений, навыков и (или) опыта деятельности, характеризующих этапы формирования компетенций в процессе освоения образовательной программы (Наименование оценочного средства)	Описание показателей и критериев оценивания компетенций на различных этапах их формирования, описание шкал оценивания (по 100-балльной шкале)
			представления знаний и обработки вербальной информации Н. Владеть навыками работы с информационными и экспертными системами и средствами поиска и обработки вербальной информации		
				Итого	100

6.2. Промежуточный контроль (зачет, экзамен)

Рабочим учебным планом предусмотрен Зачет в семестре 21.

ВОПРОСЫ ДЛЯ ПРОВЕРКИ ЗНАНИЙ:

1-й вопрос билета (40 баллов), вид вопроса: Тест/проверка знаний. Критерий: 40 баллов: всесторонние и глубокие знания программного материала, способности в понимании и изложении материала (1-2 ошибки при выполнении теста); 20-34 баллов: стабильные с незначительными ошибками знания программного материала, способность к самостоятельному применению и обновлению знаний в ходе ответа на вопрос (выполнения теста); 10-19 баллов: знания основного программного материала в объеме, необходимом для последующего обучения и предстоящей практической деятельности, неточности в ответе, но умение их устранить при корректировке со стороны экзаменатора; 0-9 баллов: существенные пробелы в знании основного программного материала, принципиальные ошибки при выполнении теста.

Компетенция: ОПК-7 Способен работать с основными информационно-поисковыми и экспертными системами, системами представления знаний и обработки вербальной информации

Знание: Знать информационные и экспертные системы и средства представления знаний и обработки вербальной информации

1. 1. _Определение корпуса.
2. 10. Словообразование в семантической разметке НКРЯ
3. 2. Задачи, решаемые с использованием корпусов.
4. 3. _Параметры классификации корпусов
5. 4. Виды разметки корпусов
6. 5. Основные методы морфологической разметки в корпусе
7. 6. Семантическая разметка
8. 7. Основные этапы синтаксической разметки
9. 8. Семантические пометы в НКРЯ?

10. 9. Принцип семантической разметки? Его проявление в семантической разметке НКРЯ?

ТИПОВЫЕ ЗАДАНИЯ ДЛЯ ПРОВЕРКИ УМЕНИЙ:

2-й вопрос билета (30 баллов), вид вопроса: Задание на умение. Критерий: 30 баллов: всесторонние и глубокие знания программного материала, способности в понимании и изложении материала; 16-24 баллов: стабильные с незначительными ошибками знания программного материала, способность к самостоятельному применению и обновлению знаний в ходе ответа на вопрос (выпо; 8-15 баллов: знания основного программного материала в объеме, необходимом для последующего обучения и предстоящей практической деятельности, неточности в ответе, но умение их устранить при корректировке со стороны экзаменатора; 0-7 баллов: существенные пробелы в знании основного программного материала, принципиальные ошибки.

Компетенция: ОПК-7 Способен работать с основными информационно-поисковыми и экспертными системами, системами представления знаний и обработки вербальной информации

Умение: Уметь выбирать релевантные информационно-поисковые и экспертные системы и средства представления знаний и обработки вербальной информации

Задача № 1. На примере анализа слова продемонстрируйте умение работать с инструментами НКРЯ

ТИПОВЫЕ ЗАДАНИЯ ДЛЯ ПРОВЕРКИ НАВЫКОВ:

3-й вопрос билета (30 баллов), вид вопроса: Задание на навыки. Критерий: 30 баллов: всесторонние и глубокие знания программного материала, способности в понимании и изложении материала; 16-24 баллов: стабильные с незначительными ошибками знания программного материала, способность к самостоятельному применению и обновлению знаний в ходе ответа на вопрос; 8-15 баллов: знания основного программного материала в объеме, необходимом для последующего обучения и предстоящей практической деятельности, неточности в ответе, но умение их устранить при корректировке со стороны экзаменатора; 0-7 баллов: существенные пробелы в знании основного программного материала, принципиальные ошибки.

Компетенция: ОПК-7 Способен работать с основными информационно-поисковыми и экспертными системами, системами представления знаний и обработки вербальной информации

Навык: Владеть навыками работы с информационными и экспертными системами и средствами поиска и обработки вербальной информации

Задание № 1. Результат самостоятельного исследования лексических единиц по материалам своей магистерской диссертации с использованием инструментов НКРЯ.

ОБРАЗЕЦ БИЛЕТА

Министерство науки и высшего образования
Российской Федерации
Федеральное государственное бюджетное
образовательное учреждение
высшего образования
**«БАЙКАЛЬСКИЙ ГОСУДАРСТВЕННЫЙ
УНИВЕРСИТЕТ»
(ФГБОУ ВО «БГУ»)**

Направление - 45.04.02 Лингвистика
Профиль - Теория и методика
преподавания русского языка как
иностранного
Кафедра русского языка и
межкультурной коммуникации
Дисциплина - Корпусная лингвистика

БИЛЕТ № 1

1. Тест (40 баллов).
2. На примере анализа слова продемонстрируйте умение работать с инструментами НКРЯ (30 баллов).
3. Результат самостоятельного исследования лексических единиц по материалам своей магистерской диссертации с использованием инструментов НКРЯ. (30 баллов).

Составитель _____ А.А. Федосеев

Заведующий кафедрой _____ А.М. Литовкина

7. Перечень основной и дополнительной учебной литературы, необходимой для освоения дисциплины (модуля)

а) основная литература:

1. Захаров В. П., Богданова С. Ю. Корпусная лингвистика. Учебник для студентов гуманитарных вузов/ В.П. Захаров.- Иркутск: Иркутский государственный лингвистический университет, 2011.-161 с.
2. [Захаров В.П. Корпусная лингвистика : учебник для студентов гуманитарных вузов / Захаров В.П., Богданова С.Ю.. — Иркутск : Иркутский государственный лингвистический университет, 2011. — 161 с. — ISBN 978-5-88267-316-0. — Текст : электронный // Электронно-библиотечная система IPR BOOKS : \[сайт\]. — URL: <http://www.iprbookshop.ru/21088.html>](#)

б) дополнительная литература:

1. [Современные проблемы лингвистики и лингводидактики : материалы итоговой научно-практической конференции факультета славянской и западноевропейской филологии / И.В. Дегтева \[и др.\].. — Москва : Прометей, 2011. — 216 с. — ISBN 978-5-4263-0009-5. — Текст : электронный // Электронно-библиотечная система IPR BOOKS : \[сайт\]. — URL: <http://www.iprbookshop.ru/8402.html>](#)

8. Перечень ресурсов информационно-телекоммуникационной сети «Интернет», необходимых для освоения дисциплины (модуля), включая профессиональные базы данных и информационно-справочные системы

Для освоения дисциплины обучающемуся необходимы следующие ресурсы информационно-телекоммуникационной сети «Интернет»:

- Сайт Байкальского государственного университета, адрес доступа: <http://bgu.ru/>, доступ круглосуточный неограниченный из любой точки Интернет
- Блог по изучению русского языка и культуры, адрес доступа: <http://www.learnrussianweb.com/>. доступ неограниченный
- Время говорить по-русски! Материалы по РКИ, адрес доступа: http://www.speak-russian.cie.ru/time_new/. доступ неограниченный
- Единое окно доступа к информационным ресурсам, адрес доступа: <http://window.edu.ru/>. доступ неограниченный
- КиберЛенинка, адрес доступа: <http://cyberleninka.ru>. доступ круглосуточный, неограниченный для всех пользователей, бесплатное чтение и скачивание всех научных публикаций, в том числе пакет «Юридические науки», коллекция из 7 журналов по правоведению

- Научная электронная библиотека eLIBRARY.RU, адрес доступа: <http://elibrary.ru/>. доступ к российским журналам, находящимся полностью или частично в открытом доступе при условии регистрации
- Национальный цифровой ресурс «Руконт», адрес доступа: <http://www.rucont.ru>. доступ неограниченный
- Онлайн материалы по РКИ, адрес доступа: <http://russificateschool.com/>. доступ неограниченный
- Портал "Информационно-коммуникационные технологии в образовании", адрес доступа: <http://www.ict.edu.ru/lib/>. доступ неограниченный
- Портал «Образование на русском». Проект Государственного института русского языка имени А.С. Пушкина, адрес доступа: <https://pushkininstitute.ru/>. доступ неограниченный
- Русский язык за рубежом, адрес доступа: <http://www.russianedu.ru>. доступ неограниченный
- Словарь электронный, адрес доступа: <https://www.multitran.ru/>. доступ неограниченный
- Центр русского языка МГУ, адрес доступа: <http://www.mgu-russian.com/en/learn/test-online/>. доступ неограниченный
- Электронная библиотека книг, адрес доступа: <http://aldebaran.ru/>. доступ неограниченный
- Электронная библиотечная система «Юрайт» [biblio-online.ru](http://www.biblio-online.ru), адрес доступа: <http://www.biblio-online.ru/>. У тех изданий, на которые подписано учебное заведение, доступен полный текст с возможностью цитирования и создания закладок.
- Электронно-библиотечная система IPRbooks, адрес доступа: <https://www.iprbookshop.ru>. доступ неограниченный

9. Методические указания для обучающихся по освоению дисциплины (модуля)

Изучать дисциплину рекомендуется в соответствии с той последовательностью, которая обозначена в ее содержании. Для успешного освоения курса обучающиеся должны иметь первоначальные знания в области теории современного русского языка, культуры научной и деловой речи.

На лекциях преподаватель озвучивает тему, знакомит с перечнем литературы по теме, обосновывает место и роль этой темы в данной дисциплине, раскрывает ее практическое значение. В ходе лекций студенту необходимо вести конспект, фиксируя основные понятия и проблемные вопросы.

Практические (семинарские) занятия по своему содержанию связаны с тематикой лекционных занятий. Начинать подготовку к занятию целесообразно с конспекта лекций. Задание на практическое (семинарское) занятие сообщается обучающимся до его проведения. На семинаре преподаватель организует обсуждение этой темы, выступая в качестве организатора, консультанта и эксперта учебно-познавательной деятельности обучающегося.

Изучение дисциплины (модуля) включает самостоятельную работу обучающегося.

Основными видами самостоятельной работы студентов с участием преподавателей являются:

- текущие консультации;
- коллоквиум как форма контроля освоения теоретического содержания дисциплин: (в часы консультаций, предусмотренные учебным планом);
- прием и разбор домашних заданий (в часы практических занятий);
- прием и защита лабораторных работ (во время проведения занятий);
- выполнение курсовых работ в рамках дисциплин (руководство, консультирование и защита курсовых работ в часы, предусмотренные учебным планом) и др.

Основными видами самостоятельной работы студентов без участия преподавателей являются:

- формирование и усвоение содержания конспекта лекций на базе рекомендованной лектором учебной литературы, включая информационные образовательные ресурсы (электронные учебники, электронные библиотеки и др.);
- самостоятельное изучение отдельных тем или вопросов по учебникам или учебным пособиям;
- написание рефератов, докладов;
- подготовка к семинарам и лабораторным работам;
- выполнение домашних заданий в виде решения отдельных задач, проведения типовых расчетов, расчетно-компьютерных и индивидуальных работ по отдельным разделам содержания дисциплин и др.

10. Перечень информационных технологий, используемых при осуществлении образовательного процесса по дисциплине (модулю), включая перечень программного обеспечения

В учебном процессе используется следующее программное обеспечение:

- 7-Zip,
- Firefox Developer Edition,
- MS Office,
- АБИС (Автоматизированная библиотечная информационная система),

11. Описание материально-технической базы, необходимой для осуществления образовательного процесса по дисциплине (модулю):

В учебном процессе используется следующее оборудование:

- Помещения для самостоятельной работы, оснащенные компьютерной техникой с возможностью подключения к сети "Интернет" и обеспечением доступа в электронную информационно-образовательную среду вуза,
- Учебные аудитории для проведения: занятий лекционного типа, занятий семинарского типа, практических занятий, выполнения курсовых работ, групповых и индивидуальных консультаций, текущего контроля и промежуточной аттестации, укомплектованные специализированной мебелью и техническими средствами обучения,
- Мультимедийный класс,
- Компьютерный класс