

МИНИСТЕРСТВО НАУКИ И ВЫСШЕГО ОБРАЗОВАНИЯ РОССИЙСКОЙ ФЕДЕРАЦИИ  
ПОВОЛЖСКИЙ ГОСУДАРСТВЕННЫЙ ТЕХНОЛОГИЧЕСКИЙ УНИВЕРСИТЕТ



УТВЕРЖДАЮ  
Декан ФГО

УТВЕРЖДАЮ /А.В. Артамонова/  
(Ф.И.О. декана (директора института))

11.03.2024 г.

**РАБОЧАЯ ПРОГРАММА ДИСЦИПЛИНЫ (МОДУЛЯ)**

М.1.2.3 Корпусная лингвистика и инструменты

(код и наименование дисциплины по учебному плану)

Направление подготовки  
(специальность)

45.04.02 Лингвистика

Квалификация выпускника

Магистр

(бакалавр/магистр/специалист)

Программа магистратуры

Цифровая лингвистика

Курс 1, 2

Семестр 2, 3

**Распределение учебного времени**

Трудоемкость по учебному плану	144 / 4	часов/зачетных единиц
Лекции	8	часов
Лабораторные работы	-	часов
Практические занятия	30	часов
Иная контактная работа	-	часов
Всего контактной работы (без учета экз.)	38	часов
Контактная работа по экзамену	-	часов
Курсовой проект (работа)	-	семестр
Самостоятельная работа обучающихся (без учета экз.)	106	часов
Самостоятельная работа по подготовке к экзамену	-	часов
Экзамен	-	семестр
Зачет	2	семестр
БРК, ДЗ	3	семестр

(год)

Программа составлена в соответствии с требованиями ФГОС ВО направления подготовки (специальности) 45.04.02 Лингвистика

Программу составили:

старший преподаватель	ИЯиЛ	СОГЛАСОВАНО	С.В. Манукянц
(должность)	(кафедра)		(И.О. Фамилия)

РАССМОТРЕНА и ОДОБРЕНА на заседании кафедры, за которой закреплена дисциплина  
Кафедра иностранных языков и лингвистики

		(наименование кафедры)	
31.01.2024	протокол №	5	
(дата)			
Заведующий кафедрой	СОГЛАСОВАНО	О.В. Филипчук	
		(И.О. Фамилия)	

Рабочая программа СОГЛАСОВАНА с факультетом (институтом), выпускающей(ими) кафедрой(ами).

СООТВЕТСТВУЕТ действующей ОП.

Заведующий кафедрой	СОГЛАСОВАНО	О.В. Филипчук
		(И.О. Фамилия)

Председатель методической комиссии факультета (института), в который входит выпускающая кафедра

СОГЛАСОВАНО	А.В. Артамонова
	(И.О. Фамилия)

Эксперт(ы): Лазарева Е.А. , директор Школы изучения иностранных языков «Лингва»

Рабочая программа проверена и зарегистрирована в УМЦ 12.03.2024 г.

Специалист учебно-методического центра СОГЛАСОВАНО /Т.А. Смирнова/

## Раздел 1. ЦЕЛЬ ОСВОЕНИЯ ДИСЦИПЛИНЫ

Целью освоения дисциплины является достижение планируемых результатов обучения, соответствующих установленным в ОПОП индикаторам достижения компетенций:

Код и наименование компетенции	Код и наименование индикатора достижения компетенции	Результаты обучения
1. УК-5 Способен анализировать и учитывать разнообразие культур в процессе межкультурного взаимодействия	ИД-1 ук-5 Анализирует современное состояние поликультурного общества на основе знаний межкультурной коммуникации и базы лингвистических данных (лингвистические корпуса, словари, тезаурусы)	<b>знания:</b> Знает возможности инструментов корпусной лингвистики, применимые для анализа современного состояния поликультурного общества <b>умения:</b> Умеет подбирать инструменты корпусной лингвистики для анализа современного состояния поликультурного общества <b>навыки:</b> Способен анализировать современное состояние поликультурного общества на основе инструментов корпусной лингвистики
2. ПК-4 Способен осуществлять ведение процедуры медиации и разрабатывать методику выполнения аналитических работ	ИД-1 ПК-4 использует корпуса языков и инструменты их обработки, системы машинного перевода, технологию искусственного интеллекта для анализа текстовой информации	<b>знания:</b> Знает возможности корпусов языков и инструментов их обработки для анализа текстовой информации <b>умения:</b> Умеет подбирать лингвистические корпуса и корпусные инструменты для анализа текстовой информации в соответствии с поставленной задачей <b>навыки:</b> Способен анализировать текстовую информацию с помощью корпусных инструментов в соответствии с поставленной
	ИД-3 ПК-4 использует инструменты ИТ, лингвистические корпуса и базы данных для выявления, анализа и описания мировых практик и проведения аналитических работ	<b>знания:</b> Использует лингвистические корпуса для выявления, анализа и описания мировых практик проведения аналитических работ <b>умения:</b> Умеет подбирать корпусные инструменты для выявления, анализа и описания мировых практик проведения аналитических работ <b>навыки:</b> Способен использовать лингвистические корпуса для выявления, анализа и описания мировых практик проведения аналитических работ

## Раздел 2. МЕСТО ДИСЦИПЛИНЫ В СТРУКТУРЕ ОПОП

Дисциплина относится к части, формируемой участниками образовательных отношений ОПОП.

Дисциплина является обязательной

Для продолжения формирования заявленных компетенций необходимы знания предшествующих дисциплин: Основы межкультурной коммуникации (УК-5), Практическая грамматика (УК-5), Практический курс первого иностранного языка (УК-5), Теория и практика перевода (УК-5), Общее языкознание и история лингвистических учений (УК-5), Программирование лингвистических задач на языке Python (УК-5), Теория и практика

перевода (ПК-4), Программирование лингвистических задач на языке Python (ПК-4), Методы искусственного интеллекта в компьютерной лингвистике (ПК-4); практик: Учебная практика. Переводческая практика (распределенная) (ПК-4)

Изучаемая дисциплина является основой для продолжения формирования указанных компетенций в следующих дисциплинах: Практический курс первого иностранного языка (УК-5), Язык презентаций (УК-5), Методы искусственного интеллекта в компьютерной лингвистике (ПК-4), Практикум по культуре общения на иностранном языке (УК-5), Конфликтология (ПК-4); практиках: Преддипломная практика (УК-5); государственной итоговой аттестации в форме: Подготовка к процедуре защиты и защита выпускной квалификационной работы (УК-5), Подготовка к процедуре защиты и защита выпускной квалификационной работы (ПК-4)

### Раздел 3. ОПИСАНИЕ ОБРАЗОВАТЕЛЬНЫХ ТЕХНОЛОГИЙ

Для формирования заявленных компетенций используются методологические технологии, реализующие деятельностный, личностно-ориентированный, практико-ориентированный подходы.

Основными стратегическими технологиями являются: дискуссионные, имитационное моделирование, исследовательские, практические занятия, процедуры самообучения, лекционные занятия

На достижение конкретных целей обучения направлены применяемые тактические технологии: case-study, деловая игра, задания, мини-проекты, классическая лекция, лекция с элементами мозгового штурма, лекция-провокация, проблемная лекция

### Раздел 4. СОДЕРЖАНИЕ ДИСЦИПЛИНЫ

#### 2 семестр

Виды и темы занятий	Количество часов	Формируемые компетенции
<b>Лингвистический корпус</b>	<b>72</b>	ПК-4, УК-5
Лекция. Что такое корпус?	2	
Практическое занятие. История корпусной лингвистики	2	
Практическое занятие. Самые известные корпуса русского языка	2	
Лекция. Классификация корпусов	2	
Практическое занятие. Самые известные корпуса английского языка	2	
Практическое занятие. Классификация корпусов	2	
Лекция. Типы лингвистического аннотирования	2	
Практическое занятие. Типы лингвистического аннотирования	2	
Практическое занятие. Морфологическая разметка	4	
Лекция. Проблемы морфологического аннотирования	2	
Практическое занятие. Синтаксическая разметка	2	
Задания для самостоятельной работы, в том числе выполнение Подготовка докладов и сообщений, выполнение заданий на электронном курсе (в т.ч. проведение микроисследований на основе имеющихся корпусов)	48	
Иная контактная работа:	0	

#### 3 семестр

Виды и темы занятий	Количество часов	Формируемые компетенции
<b>Инструменты корпусной лингвистики</b>	<b>72</b>	ПК-4, УК-5
Практическое занятие. Мультимодальный корпус	1	
Практическое занятие. Разметка мультимодальных корпусов	1	
Практическое занятие. Многоязычные корпуса	2	
Практическое занятие. Корпуса второго языка	1	
Практическое занятие. Интернет как корпус	1	
Практическое занятие. Инструменты создания лингвистических корпусов	2	
Практическое занятие. Форматы представления корпусных данных	2	
Практическое занятие. Методы и методика корпусных исследований	2	
Практическое занятие. Сферы применения корпусной лингвистики	2	
Задания для самостоятельной работы, в том числе выполнение Подготовка докладов и сообщений, выполнение заданий на электронном курсе (в т.ч. проведение микроисследований на основе имеющихся корпусов и собственного корпуса)	58	
Иная контактная работа:	0	

## Раздел 5. МЕТОДИЧЕСКИЕ УКАЗАНИЯ ДЛЯ ОБУЧАЮЩИХСЯ ПО ОСВОЕНИЮ ДИСЦИПЛИНЫ

Изучение дисциплины рекомендуется начать с ознакомления с рабочей программой, ее структурой и содержанием разделов. Учебный материал структурирован, изучение дисциплины осуществляется в тематической последовательности.

Занятия **лекционного типа** дают систематизированные знания по дисциплине (модулю), концентрируют внимание на наиболее сложных и важных вопросах. Во время лекционных занятий рекомендуется вести конспектирование учебного материала; обращать внимание на формулировки и категории, раскрывающие суть проблемы, явления или процесса; зафиксировать выводы и практические рекомендации.

Подготовка к занятиям **семинарского типа** включает ознакомление с планом практического занятия; работу с конспектом лекций, выполнение домашнего задания, работу с учебной и учебно-методической литературой, научными изданиями и электронными образовательными ресурсами, рекомендованными рабочей программой дисциплины.

Содержание **самостоятельной работы** определяется рабочей программой дисциплины, оценочными и методическими материалами, заданиями и указаниями преподавателя. Самостоятельная работа может осуществляться в аудиторной и внеаудиторной формах. Эффективным средством осуществления самостоятельной работы является электронная информационно-образовательная среда университета, которая обеспечивает доступ к образовательной программе, рабочей программе дисциплины, к электронным библиотечным системам, профессиональным базам данных и информационным справочным системам.

Периодичность проведения, формы текущего контроля успеваемости, система оценивания хода освоения дисциплин представлены в рабочей программе.

Формой промежуточной аттестации по дисциплине (модулю) является зачёт в 2-м семестре, БРК в 3-м семестре.

## Раздел 6. МАТЕРИАЛЬНО-ТЕХНИЧЕСКОЕ И УЧЕБНО-МЕТОДИЧЕСКОЕ ОБЕСПЕЧЕНИЕ ДИСЦИПЛИНЫ

### 6.1. Учебно-методическое обеспечение

№№ п/п	Список используемой литературы	Количество экземпляров печатных изданий, имеющих в библиотеке, или электронный адрес издания (ресурса) в сети Интернет
<b>УЧЕБНЫЕ, УЧЕБНО-МЕТОДИЧЕСКИЕ И НАУЧНЫЕ ИЗДАНИЯ</b>		
1.	Коломейченко, А. С. Информационные технологии [Текст] : Учебное пособие для вузов / Коломейченко А. С., Польшакова Н. В., Чеха О. В.; Польшакова Н. В., Чеха О. В. 3-е изд., стер. Санкт-Петербург: Лань, 2022. - 212 с. ISBN 978-5-507-45293-4.	<a href="https://e.lanbook.com/book/264086">https://e.lanbook.com/book/264086</a>
2.	Информационные технологии в образовании [Электронный ресурс] : учебник / Е. В. Баранова, М. И. Бочаров, С. С. Куликова, Т. Б. Павлова. Санкт-Петербург: Лань, 2022. - 296 с. ISBN 978-5-8114-2187-9.	<a href="https://e.lanbook.com/book/212435">https://e.lanbook.com/book/212435</a>
3.	Костюк, А. В. Информационные технологии. Базовый курс [Электронный ресурс] / Костюк А. В., Бобонец С. А., Флегонтов А. В., Черных А. К. 3-е изд., стер. Санкт-Петербург: Лань, 2021. - 604 с. ISBN 978-5-8114-8776-9.	<a href="https://e.lanbook.com/book/180821">https://e.lanbook.com/book/180821</a>
4.	Гарифуллина, Флюра Зинатовна. Подготовка специалистов в области лингвистики и межкультурной коммуникации к научно-исследовательской деятельности [Текст] : учебно-методическое пособие : по программе магистерской подготовки "Лингвистика" / Ф. З. Гарифуллина, С. П. Фирсова; Министерство науки и высшего образования Российской Федерации, ФГБОУ ВО "Поволжский государственный технологический университет". Йошкар-Ола: ПГТУ, 2020. - 80, [1] с. ISBN 978-5-8158-2183-5. Экземпляры: всего 15.	15 / <a href="https://portal.volgatech.net/books/Garifullina_Podgotovka_spetsialistov_v_oblasti_lingvistiki_i_mezhkulturnoy_kommunikatsii_uchebno_metodicheskoye_posobiye_2020.pdf">https://portal.volgatech.net/books/Garifullina_Podgotovka_spetsialistov_v_oblasti_lingvistiki_i_mezhkulturnoy_kommunikatsii_uchebno_metodicheskoye_posobiye_2020.pdf</a>
<b>ЭЛЕКТРОННЫЕ ОБРАЗОВАТЕЛЬНЫЕ РЕСУРСЫ</b>		
1.	Научная электронная библиотека eLIBRARY.RU	<a href="http://elibrary.ru">http://elibrary.ru</a>
2.	Научная электронная библиотека «Киберленинка»	<a href="http://cyberleninka.ru">http://cyberleninka.ru</a>
<b>ПРОФЕССИОНАЛЬНЫЕ БАЗЫ ДАННЫХ И ИНФОРМАЦИОННЫЕ СПРАВОЧНЫЕ СИСТЕМЫ</b>		
1.	Национальный корпус русского языка	<a href="https://ruscorpora.ru/new/index.html">https://ruscorpora.ru/new/index.html</a>

### 6.2. Материально-техническая база и программное обеспечение

№№ п/п	Аудитории для проведения учебных занятий, самостоятельной работы и проведения государственной итоговой аттестации	Перечень основного оборудования	Программное обеспечение
1.	527 (I)	Доска маркерная 120x240 см (1), ЖК- панель Sharp PN-L602B 60" (152 см) в комплекте (1), Колонки MICROLAB SOLO15 (1), Микрофон Shure WL185 (2), Микрофонная радиосистема SHURE BLX188E/SM35 (1), Микшерный пульт YAMAHA MG10XU (1), Ноутбук ASUS K53SC 15,6" (1), Персональный компьютер-Моноблок Dell Inspiron 3277 21,5" Full (10), ПК в составе: Сист.блок Intel Core j5,2*8Gb+ Монитор ASUS 23,6" + Монитор Dell 24" + клав, мышь (1), ПК Моноблок RAMEC GALE Custom 21,5"/i3-3240/H61M/4DDR3/500SATA3/клав.,мышь (12), Проектор мультимедийный Hitachi CP-EX250 с креплением (1), Комплект	Microsoft Windows Enterprise, Справочная правовая система "Консультант Плюс", Microsoft Office Standard, Агент Dr.Web, Комплект ГАРАНТ-Мастер, Microsoft Visio Professional, Комплект ПО для решения основных пользовательских задач

## Раздел 7. ФОРМЫ КОНТРОЛЯ ОСВОЕНИЯ ДИСЦИПЛИНЫ/ ФОНД ОЦЕНОЧНЫХ СРЕДСТВ

Критерии оценивания индикаторов достижения компетенций направлены на:

- усвоение теоретического материала (объем знаний, глубина усвоения), предусмотренного рабочей программой;
- умение излагать материал (четкость, грамотность изложения материала, точность и полнота воспроизведения учебного материала);
- умение применять теоретические знания при решении практических заданий.

Шкала оценивания представлена ниже.

Уровень сформированности элементов компетенции	Критерии оценивания	Шкала оценивания
Пороговый уровень	Обучающийся имеет знания основного материала, проявляет умение логично его излагать, но может допускать неточности в изложении материала, недостаточно правильные формулировки, испытывает затруднения в выполнении практических заданий.	удовлетворительно
Продвинутый уровень	Обучающийся твердо знает программный материал, излагает его грамотно и по существу, не допускает существенных неточностей в ответе на вопрос, правильно применяет теоретические положения при решении практических вопросов и задач, владеет необходимыми	хорошо

	навыками и приемами их выполнения	
Высокий уровень	Обучающийся глубоко и прочно усвоил программный материал, грамотно и логически стройно его излагает, дает исчерпывающие ответы на поставленные вопросы. В ответе тесно увязывается теория с практикой, при этом обучающийся не затрудняется с ответом при видоизменении задания, свободно справляется с задачами, вопросами и другими видами применения знаний, показывает знакомство с монографической литературой, периодическими изданиями, правильно обосновывает принятые решения, свободно владеет разносторонними навыками, приемами выполнения практических работ	отлично

### 7.1. Текущий контроль успеваемости

Текущий контроль успеваемости обеспечивает оценивание хода освоения дисциплины (модуля) и производится с применением технологии рейтингового контроля в соответствии с технологической картой дисциплины. Порядок составления технологической карты и алгоритм проведения процедуры оценивания видов деятельности обучающихся, направленных на освоение знаний, умений, навыков и/или опыта деятельности, по накопительной системе в баллах устанавливается положением о системе РИТМ в ФГБОУ ВО «ПГТУ»

### 7.2. Промежуточная аттестация обучающихся

Промежуточная аттестация обучающихся направлена на оценивание результатов обучения по дисциплине (модулю) и проводится с использованием фондов оценочных средств.

Примеры типовых контрольных заданий из базы фонда оценочных средств по образовательной программе.

1. Название «математическая лингвистика» по отношению к названию «корпусная лингвистика» возникло:

Выберите один правильный ответ

одновременно с ним

позже

раньше

2. Термины «математическая лингвистика», «квантитативная лингвистика», «статистика речи (текста)», «лингвостатистика», «комбинаторная лингвистика», «корпусная лингвистика» являются:

Выберите один правильный ответ

антонимами

гипонимами

паронимами

семантическими синонимами

3. Термин «корпусная лингвистика» был введен:

Выберите все правильные ответы (один или несколько)



Ирвингом Лорджем

Чарльзом Филлмором

Яном Свартвиком

Джефри Нилом Личем

4. Раздел лингвистики, в котором решение прикладных и теоретических задач связано с использованием информационных технологий, называют:

Выберите один правильный ответ

корпусная лингвистика

диалектология

лингводидактика

компаративистская лингвистика

5. Расположите следующие разделы лингвистики в порядке их появления:

Расставьте в правильном порядке

квантитативная лингвистика

корпусная лингвистика

компьютерная лингвистика

математическая лингвистика

статистическая лингвистика

Распределите направления в лингвистике по познавательным моделям.

Соотнесите элемент соответствующий категории выбрав вариант из списка

формальная грамматика

лингвопоэтика

корпусная лингвистика

экспериментальная фонетика

лексикология

функциональная грамматика

Соедините термины и их определения.

Соедините элементы попарно (неверно соединенную пару можно разбить, щелкнув на крестик)

реляционная модель

субстанциональная модель

корпусная лингвистика

раздел языкознания, который занимается организацией и исследованием лингвистических баз данных

базируется на том, что анализируемый в её рамках объект обладает качеством или их набором,

которые характеризуют объект как не существующий сам по себе, а являющийся следствием отношений между другими объектами

базируется на том, что анализируемый в её рамках объект обладает качеством или их набором, которые характеризуют объект как существующий сам по себе, самостоятельно, наряду с другими объектами, вне зависимости от наблюдателя

Количественный анализ языка и речи, предполагающий различные типы подсчётов количества категорий и единиц, называют:

Выберите один правильный ответ

квантификация

диглоссия

билингвизм

иерархизация

В каких науках традиционно более распространена реляционная модель?

Выберите все правильные ответы (один или несколько)

биология

физика

социология

математика

филология

химия

Выделение в языке определенных классов явлений, объединенных определенными качественными признаками (качественный анализ языка), называют:

Выберите один правильный ответ

категоризация

билингвизм

семиотика

стратификация

Особым образом организованная совокупность электронных текстов, над которой автоматически можно производить различные поисковые операции, называется:

Выберите один правильный ответ

картотека

библиотека

гипертекст

корпус

Специализированную систему, в которую входят программные средства для поиска данных в корпусе, называют:

Выберите один правильный ответ

корпусный менеджер

инверсия  
информационный ресурс  
база корпусных данных

Фиксация корпусом типичных явлений осуществляется за счёт:

Выберите один правильный ответ  
корпусного менеджера  
разметки  
больших объёмов информации  
сужения области изучения

Укажите верные суждения:

Выберите все правильные ответы (один или несколько)  
при фиксации данных в корпусе происходит ситуация вырывания из контекста  
корпусам свойственна высокая степень субъективности  
в корпусе фиксируются типичные, а не уникальные языковые явления  
внешним интерфейсом корпуса является поисковая программа  
корпус предполагает многократное использование

Какой советский и русский поэт придумал уникальный жанр поэзии на карточках? Укажите имя и фамилию полностью.

Введите ответ в виде текста (регистр не учитывается)

Производимый в процессе разметки процесс приписывания любой форме слова его начальной формы называется

Введите на месте пропуска текст (регистр не учитывается)

Обработка текстов в соответствии с задачами того корпуса, в который они помещены, называется:

Выберите один правильный ответ

токенизация  
разметка  
схематизация  
моделирование

Как называется разметка, которая позволяет выделять паузы, повторы, оговорки?

Выберите один правильный ответ

морфологическая разметка  
анафорическая разметка  
просодическая разметка  
дискурсивная разметка

Производимый в процессе разметки процесс выстраивания синтаксических зависимостей называется

Введите на месте пропуска текст (регистр не учитывается)

Информация об имени автора, названии произведения, а также род, вид, жанр текста, время его создания и т. п. называется

Введите на месте пропуска текст (регистр не учитывается)

Корпус, в котором представлено несколько языков, называется:

Выберите один правильный ответ

мультязычным  
двужычным  
однойзычным  
многоязычным

Корпус, в котором содержатся все ресурсы национального языка, называется корпусом.

Введите на месте пропуска текст (регистр не учитывается)

Корпус, в котором содержится только один тип языковых единиц, называется:

Выберите один правильный ответ

монотематическим  
специальным  
смешанным  
закрытым

С точки зрения назначения корпуса выделяются:

Выберите все правильные ответы (один или несколько)

специализированные корпуса  
иллюстративные корпуса  
синхронические корпуса  
исследовательские корпуса  
диахронические корпуса

Корпус, который постоянно пополняется, называют:

Выберите один правильный ответ

монотематическим  
статичным  
релятивистским  
динамическим

Закон, согласно которому, если все слова языка расположить по убыванию частоты их использования, то частота  $n$ -го слова в таком списке окажется приблизительно обратно пропорциональной его порядковому номеру  $n$ , открыл:

Выберите один правильный ответ

Джордж Кингсли Ципф  
Жан-Батист Эсту  
Феликс Ауэрбах  
Бенуа Мандельброт

Бенуа Мандельброт предложил объяснение закона рангового распределения с точки зрения:

Выберите один правильный ответ

информационной теории  
биологической теории  
алгебраической теории  
психологической теории

Выберите верное суждение:

Выберите один правильный ответ

рост популярности слов неограничен  
наиболее популярные слова часто меняют свой ранг

малое количество употреблений редких слов может значительно увеличить их популярность  
гипербола закона рангового распределения имеет длинный пик и длинный хвост

Во сколько раз частотность 97-го слова в списке частотности по убыванию будет ниже, чем частотность первого слова в этом списке?  
Введите ответ в виде числа

Положение слова в общем списке частотности слов по убыванию называют

Введите на месте пропуска текст (регистр не учитывается)

Перечень вопросов для проведения промежуточной аттестации

**2-й семестр (зачёт)**

1. Предмет корпусной лингвистики
2. Полевая лингвистика и корпусная лингвистика
3. История лингвистических корпусов: от картотеки к корпусу
4. Классификация (типология) корпусов
5. Корпусная лингвистика: современное состояние
6. Корпусная лингвистика в России
7. Обзор существующих корпусов различных типов
8. Корпус как поисковая система
9. Корпусоподобные интерфейсы между лингвистом и поисковыми системами Интернета
10. Лингвистические исследования, базирующиеся на корпусах
11. Проблемы репрезентативности корпусов
12. Проблемы хронологии в общеязыковых корпусах
13. Отбор текстов для корпусов
14. Графематический анализ

**3-й семестр (БРК)**

1. Предмет корпусной лингвистики
2. Полевая лингвистика и корпусная лингвистика
3. История лингвистических корпусов: от картотеки к корпусу
4. Классификация (типология) корпусов
5. Корпусная лингвистика: современное состояние
6. Корпусная лингвистика в России
7. Обзор существующих корпусов различных типов
8. Корпус как поисковая система

9. Корпусоподобные интерфейсы между лингвистом и поисковыми системами Интернета
10. Лингвистические исследования, базирующиеся на корпусах
11. Проблемы репрезентативности корпусов
12. Проблемы хронологии в общезыковых корпусах
13. Отбор текстов для корпусов
14. Графематический анализ
15. Понятие разметки
17. Типы разметки
18. Морфологическая разметка
19. Синтаксические корпуса (treebanks)
20. Семантическая разметка
21. Технология создания корпусов. Стадии работы
22. Понятие корпусоида
23. Автоматическая морфоразметка
24. Автоматический синтаксический анализ (parsing)
25. Языковые средства представления размеченных текстов (языки SGML, XML)
26. Международные стандарты (TEI, EAGLES, CDIF, XCES)
27. Способы использования корпусов в лингвистических исследованиях
28. Исследование способов использования корпусов в лексикографии
29. Изучение средств обработки корпусных данных, представленных на языке XML
30. Обзор существующих корпусов различных типов
31. Классификация (типология) корпусов по различным основаниям
32. Типы корпусов по задачам. Типы корпусов по формальным признакам