



МИНИСТЕРСТВО НАУКИ И ОБРАЗОВАНИЯ И РОССИЙСКОЙ ФЕДЕРАЦИИ
Федеральное государственное автономное образовательное учреждение высшего образования
«Дальневосточный федеральный университет»
(ДВФУ)

**ВОСТОЧНЫЙ ИНСТИТУТ - ШКОЛА РЕГИОНАЛЬНЫХ И МЕЖДУНАРОДНЫХ
ИССЛЕДОВАНИЙ**

«СОГЛАСОВАНО»
Руководитель ОП
Фундаментальная и прикладная лингвистика

«УТВЕРЖДАЮ»
Заведующая кафедрой
лингвистики и межкультурной коммуникации

Спицына Н.А.

Ловцевич Г.Н.

«21» июня 2019 г.

«21» июня 2019 г.

РАБОЧАЯ ПРОГРАММА ДИСЦИПЛИНЫ
Технологии корпусной лингвистики
Направление подготовки 45.03.03 Фундаментальная и прикладная лингвистика
Форма подготовки очная
Наименование программы «Фундаментальная и прикладная лингвистика»

курс 3 семестр 5
лекции 18 час.
практические занятия 18 час.
лабораторные работы час.
в том числе с использованием МАО лек. 6 /пр. /лаб. час.
всего часов аудиторной нагрузки 18 часа.
в том числе с использованием МАО 6 час.
самостоятельная работа 90 час.
в том числе на подготовку к экзамену 45 час.
курсовая работа / курсовой проект семестр
зачет семестр
экзамен 5 семестр

Рабочая программа составлена в соответствии с требованиями Федерального государственного образовательного стандарта по направлению подготовки 45.03.03 **Фундаментальная и прикладная лингвистика** утвержденного приказом Министерства образования и науки РФ от 24.04.2018 г. № 323.

Рабочая программа обсуждена на заседании кафедры лингвистики и межкультурной коммуникации протокол № 10 от «21» июня 2019 г.

Заведующая кафедрой д.ф.н., профессор Ловцевич Г.Н.
Составитель: канд. филол. наук, доцент Н.А. Спицына

Оборотная сторона титульного листа РПУД

I. Рабочая программа пересмотрена на заседании кафедры:

Протокол от « _____ » _____ 20__ г. № _____

Заведующий кафедрой _____
(подпись) (И.О. Фамилия)

II. Рабочая программа пересмотрена на заседании кафедры:

Протокол от « _____ » _____ 20__ г. № _____

Заведующий кафедрой _____
(подпись) (И.О. Фамилия)

III. Рабочая программа пересмотрена на заседании кафедры:

Протокол от « _____ » _____ 20__ г. № _____

Заведующий кафедрой _____
(подпись) (И.О. Фамилия)

IV. Рабочая программа пересмотрена на заседании кафедры:

Протокол от « _____ » _____ 20__ г. № _____

Заведующий кафедрой _____
(подпись) (И.О. Фамилия)

Аннотация дисциплины «Технологии корпусной лингвистики»

Учебная дисциплина «Технологии корпусной лингвистики» предназначена для студентов, обучающихся по направлению подготовки 45.03.03 Фундаментальная и прикладная лингвистика, по профилю «Фундаментальная и прикладная лингвистика».

Общая трудоемкость освоения дисциплины составляет 3 зачётные единицы, 108 часов. Учебным планом предусмотрены лекционные занятия (18 часов, из них 6 часов – в интерактивной форме), лабораторные работы (18 часов), самостоятельная работа (72 часа, том числе 45 часов на подготовку к экзамену). Дисциплина реализуется в 5 семестре, промежуточная форма отчетности - экзамен.

Дисциплина входит в базовую часть учебного плана и входит в вариативную, формируемую участниками образовательных отношений.

При разработке рабочей программы учебной дисциплины использованы Федеральный государственный образовательный стандарт высшего образования по направлению подготовки 45.03.03 Фундаментальная и прикладная лингвистика, профиль «Фундаментальная и прикладная лингвистика».

Содержание дисциплины охватывает следующий круг проблем: она способствует формированию лингвистического мировоззрения, т.е. понимания законов существования языка, практическому освоению одного из наиболее перспективных направлений в современной теоретической и прикладной лингвистике, основанного на использовании языковых электронных корпусов и корпусных технологий. Курс также призван сообщить сведения о новых направлениях прикладной лингвистики, связанных с автоматической обработкой текста (речи), с широким использованием ЭВМ, а также, рассмотреть традиционные проблемы прикладного языкознания в их связи с новыми направлениями. Курс является междисциплинарным, привлекает данные из нескольких дисциплин, изучающих лингвистические проблемы.

Дисциплина Технологии корпусной лингвистики логически и содержательно связана с такими курсами, как «Введение в языкознание (европейские языки)», «Семантика и аналитическая грамматика», «Методы семантического анализа», «Риторика и академическое письмо», «Компьютерная лексикография и информационные технологии».

Цели и задачи дисциплины:

Цель курса: познакомить обучающихся с основами корпусной лингвистики в рамках эмпирического подхода к изучению языка, с концепциями корпусной лингвистики, дать им возможность освоить основы корпусных технологий, приобрести навыки работы с корпусами, определить место дисциплины в ряду дисциплин лингвистического цикла, показать

теоретическое и практическое значение корпусов для проведения подлинно научных лингвистических исследований;

Задачи:

- ознакомить с новой парадигмой в лингвистических исследованиях;
- ознакомить с предметами, задачами, методами и понятийным (терминологическим) аппаратом корпусной лингвистики;
- ознакомить с историей корпусных исследований;
- сформировать знания о компьютерном инструментарии корпусной лингвистики;
- ознакомить с практикой составления корпусов, сферами применения корпусной лингвистики;
- способствовать освоению обучающимися методики работы с корпусными базами данных в целях лингвистического анализа;
- способствовать изучению языковых и программных средств корпусной лингвистики;
- сформировать навыки работы с программными средствами и информационными ресурсами корпусной лингвистики.

Планируемые результаты обучения по данной дисциплине (знания, умения, владения), соотнесенные с планируемыми результатами освоения образовательной программы, характеризуют этапы формирования следующих общекультурных компетенций:

Код и формулировка компетенции	Этапы формирования компетенции	
ПК-3. Владеет методами сбора и документации лингвистических данных	Знает	основные методы документирования естественного языка; принципы полевой работы с носителями естественных языков; технологии проведения сбора языковых данных; базовые принципы передачи и хранения собранной информации.
	Умеет	производить запись речевого сигнала на цифровой носитель; аннотировать, обрабатывать и анализировать собранные данные; собирать методические данные, в частности, социолингвистические данные о носителе языка.
	Владеет	Имеет практический опыт документирования лингвистических данных: записи речевого сигнала с последующей

		обработкой и анализом; обработки, структурирования и хранения метаинформации о собранных данных
ПК-4. Способен спланировать и провести лингвистический эксперимент, описать его результаты и сформулировать выводы	Знает	общие положения о специфике и архитектуре эксперимента; базовые принципы проведения экспериментов с носителями языка; принципы составления анкет и вопросников; методы создания лингвистических экспериментов с привлечением достижений современных технологий; имеет базовые представления о методах математической статистики, используемых при обработке результатов эксперимента.
	Умеет	разработать и провести базовый лингвистический эксперимент; отобрать испытуемых; обобщить и проанализировать полученные данные; сформулировать результат.
	Владеет	Имеет практический опыт проведения базовых лингвистических экспериментов; разработки их архитектуры; поиска испытуемых; обработки результатов.
ПК-8. Владеет навыками оформления и представления результатов научного исследования	Знает	общие представления о способе подачи научных материалов; методы оформления результатов научного исследования в виде статей, квалификационных работ, постеров, презентаций, докладов на конференциях и семинарах.
	Умеет	грамотно и в соответствии с требованиями оформить научную работу, раздаточный материал или презентацию к докладу; сделать устный доклад в соответствии с регламентом.
	Владеет	Имеет практический опыт выступлений перед аудиторией, а также создания письменных работ для представления результатов научного исследования
ПК-9.Способен пользоваться лингвистически ориентированными программными продуктами	Знает	основные типы систем, использующих модули лингвистического анализа; основные принципы и методы компьютерного моделирования лингвистических задач.
	Умеет	Анализировать работу различных систем обработки текста и звучащей речи для выявления основных лингвистических компонентов и основных типов обработки

		<p>текста, используемых в данных системах; подбирать необходимые лингвистические ресурсы для различных задач лингвистического обеспечения систем (например, лексикографических, задач морфологического анализа и т.п.).</p>
<p>ПК-10. Владеет принципами создания электронных языковых ресурсов (текстовых, речевых и мультимодальных корпусов; словарей, тезаурусов, онтологий; фонетических, лексических, грамматических и иных баз данных и баз знаний) и умеет пользоваться такими ресурсами</p>	<p>Владеет</p>	<p>Имеет практический опыт работы с различными системами автоматической и экспертной обработки текста и звучащей речи.</p>
	<p>Знает</p>	<p>Основные принципы обработки информации; базовые принципы корпусной лингвистики, лексикографии, математической статистики; базовые представления о языковом разнообразии; наиболее полные и значимые лингвистические корпуса, электронные словари и базы данных.</p>
	<p>Умеет</p>	<p>пользоваться основными методами, способами и средствами получения, хранения, переработки информации; пользоваться лингвистически ориентированными программными продуктами.</p>
<p>Владеет</p>	<p>основными принципами обработки информации; базовыми принципами корпусной лингвистики, лексикографии, математической статистики; базовые представления о языковом разнообразии; наиболее полные и значимые лингвистические корпуса, электронные словари и базы данных. пользоваться основными методами, способами и средствами получения, хранения, переработки информации; пользоваться лингвистически ориентированными программными продуктами.</p>	

Обучение корпусным технологиям, основаны на принципах, именуемых МАО по определению, поскольку анализ проводится на основе сопоставления знаковых структур исходных единиц и включает консультирование. Использование лекции-беседы. Мозговой штурм – для порождения собственного высказывания в рамках отдельных заданий.

I. СТРУКТУРА И СОДЕРЖАНИЕ ТЕОРЕТИЧЕСКОЙ ЧАСТИ КУРСА (18 ЧАС.)

РАЗДЕЛ 1. ВВЕДЕНИЕ В КОРПУСНУЮ ЛИНГВИСТИКУ (8 ЧАС.)

Тема 1: предмет, цели и задачи корпусной лингвистики (4 час., в том числе 4 часа с применением МАО)

Введение в корпусную лингвистику. Понятие лингвистического корпуса. История корпусной лингвистики. Корпусный (эмпирический) подход в сравнении с хомскианской лингвистикой. Задачи и основные направления корпусной лингвистики. Корпусная лингвистика и компьютерная лингвистика. Предмет исследования.

Тема 2: методика лингвистического корпуса (2 час.)

Задачи: создание корпусов и лингвистические исследования на их базе, объективное лингвистическое описание языковой системы. Направления: лексикографические исследования, исследования лексического состава языков, исследования относительных частот употребления тех или иных слов, изучение изменений в лексическом составе языков, различных его вариаций, изучение грамматики естественных языков. Корпусная лингвистика занимается в основном описанием и объяснением сущности, структуры и использования языка. Компьютерная лингвистика выступает для корпусной в качестве «поставщика» инструментов анализа и обработки корпусов. Понятие лингвистического корпуса. Создание и типология корпусов. Развитие лингвистических корпусов в мире. Первое и второе поколение корпусов. Разметка. Виды разметок.

Тема 3: типы корпусов (2 час., в том числе 2 часа с применением МАО)

История лингвистических корпусов: от картотеки к корпусу. Корпусная лингвистика в России. История электронных лингвистических корпусов. Четыре группы корпусных лингвистов. Первое поколение корпусов: The Brown Corpus, Lancaster-Oslo/Bergen (LOB) Corpus, London-Lund Corpus (LLC). Второе поколение корпусов: The Cobuild Project / The Bank of English, The Longman Corpus Network, British National Corpus (BNC), The International Corpus of English (ICE), American National Corpus, Gigaword corpora. Типология корпусов. Типы корпусов: устные и

письменные, одноязычные и многоязычные, аннотированные и неаннотированные.

РАЗДЕЛ 2. КОРПУСНЫЕ ЛИНГВИСТИЧЕСКИЕ ИССЛЕДОВАНИЯ (10 час.)

Тема 4: лексические корпусные лингвистические исследования (2 час.)

Понятие разметки. Типы разметки. Автоматический морфологический и синтаксический анализ. Металингвистическая разметка. Параллельные корпуса. Проблема выравнивания. Языковые средства представления размеченных текстов Экстралингвистическая разметка. Методы и средства разметки и создания корпусов. Лингвистические исследования на базе корпуса: изучение лексики. Лингвистические исследования на базе корпуса: изучение грамматики.

Тема 5: узус в корпусе (2 час.)

Лингвистические исследования на базе корпуса: изучение лексики, изучение грамматики. Распределение и функции номинализаций. Распределение грамматических категорий. Обзор существующих корпусов различных типов. Сравнительный анализ корпусов и корпусных менеджеров. Брауновский корпус. Британский национальный корпус. Корпус современного американского английского. Русско-английский корпус С. Шарова. Технология создания корпусов. Классификация (типология) корпусов по различным основаниям. Типы корпусов по задачам. Типы корпусов по формальным признакам. Корпус первого типа. Корпус второго типа. Лингвистические исследования на базе корпусов. Исследование нормы/узуса. Социолингвистические исследования. Возможности преподавания иностранного языка на основе корпусной лингвистики

Тема 6: практические сферы применения корпусов: определения и основные понятия (4 час.)

Лингвистические исследования на базе корпусов. Стилистические исследования. Лингвистические исследования на базе корпусов. Практические сферы применения: установление плагиата.

Тема 7: Программы для работы с корпусом (2 ч.)

Семантическое наполнение Британского национального корпуса. Методы извлечения информации из корпуса. Типы извлекаемой информации. Конкорданс. Программы для работы с корпусом: конкордансеры и другие. Параллельные корпуса. Использование корпусов в других науках. Статистические меры вычисления совместной встречаемости. Методы извлечения информации из корпуса. Типы извлекаемой информации.

Конкорданс. Программы для работы с корпусом: конкордансеры и другие.
Параллельные корпуса.

II. СТРУКТУРА И СОДЕРЖАНИЕ ПРАКТИЧЕСКОЙ ЧАСТИ КУРСА

Отсутствует по плану.

III. УЧЕБНО-МЕТОДИЧЕСКОЕ ОБЕСПЕЧЕНИЕ САМОСТОЯТЕЛЬНОЙ РАБОТЫ ОБУЧАЮЩИХСЯ

Учебно-методическое обеспечение самостоятельной работы обучающихся по дисциплине представлено в Приложении 1 и включает в себя:

план-график выполнения самостоятельной работы по дисциплине, в том числе примерные нормы времени на выполнение по каждому заданию;

характеристика заданий для самостоятельной работы обучающихся и методические рекомендации по их выполнению;

требования к представлению и оформлению результатов самостоятельной работы;

критерии оценки выполнения самостоятельной работы.

VI. КОНТРОЛЬ ДОСТИЖЕНИЯ ЦЕЛЕЙ КУРСА

№ п/п	Контролируемые разделы / темы дисциплины	Коды и этапы формирования компетенций		Оценочные средства	
				текущий контроль	промежуточная аттестация
1	Введение в корпусную лингвистику	ПК-3	Знает	собеседование (УО-1)	Вопросы № 1,2 экзаменационного билета, билеты 1-12
			Умеет	собеседование (УО-1)	
			Владеет	собеседование (УО-1)	
		ПК-4	Знает	собеседование (УО-1)	
			Умеет	собеседование (УО-1)	
			Владеет	реферат (ПР-4)	
2.	корпусные лингвистические исследования	ПК-9	Знает	собеседование (УО-1)	Вопросы № 1,2 экзаменационного билета, билеты 1-12
			Умеет	реферат (ПР-4)	
			Владеет	собеседование (УО-1)	

		ПК-8	Знает	собеседовани е (УО-1)	
			Умеет	собеседовани е (УО-1)	
			Владеет	собеседовани е (УО-1)	
3	корпусные технологии	ПК-10	Знает	собеседовани е (УО-1)	Вопросы № 1,2 экзаменационн ого билета, билеты 1-12
			Умеет	собеседовани е (УО-1)	Вопросы № 1,2 экзаменационн ого билета, билеты 1-12
			Владеет	собеседовани е (УО-1)	Вопросы № 1,2 экзаменационн ого билета, билеты 1-12

Типовые контрольные задания, методические материалы, определяющие процедуры оценивания знаний, умений и навыков и (или) опыта деятельности, а также критерии и показатели, необходимые для оценки знаний, умений, навыков и характеризующие этапы формирования компетенций в процессе освоения образовательной программы, представлены в Приложении 2.

V. УЧЕБНО-МЕТОДИЧЕСКОЕ ОБЕСПЕЧЕНИЕ ДИСЦИПЛИНЫ

Основная литература

(печатные и электронные издания)

1. Информатика и лингвистика : учеб. пособие / Т.М. Волосатова, Н.В. Чичварин. — М. : ИНФРА-М, 2018. — 196 с. + Доп. материалы [Электронный ресурс; Режим доступа <http://www.znanium.com>]. — (Высшее образование: Бакалавриат). — www.dx.doi.org/10.12737/16175. - Режим доступа: <http://znanium.com/catalog/product/938009>
2. Чернявская, В. Е. Лингвистика текста. Лингвистика дискурса : учеб. пособие / В. Е. Чернявская. — М. : Флинта : Наука, 2013. — 208 с. - ISBN 978-5-9765-1454-6 (Флинта), ISBN 978-5-02-037800-1 (Наука) - Режим доступа: <http://znanium.com/catalog/product/457232>
3. Информатика и лингвистика: Учебное пособие / Волосатова Т.М., Чичварин Н.В. - М.: НИЦ ИНФРА-М, 2016. - 196 с.: 60x90 1/16. - (Высшее образование: Бакалавриат) (Переплёт 7БЦ) ISBN 978-5-16-010977-0 - Режим доступа: <http://znanium.com/catalog/product/508097>

4. Щипицина, Л. Ю. Информационные технологии в лингвистике [Электронный ресурс] : учеб. пособие / Л. Ю. Щипицина. — М. : ФЛИНТА, 2013. — 128 с. - ISBN 978-5-9765-1431-7 - Режим доступа: <http://znanium.com/catalog/product/462989>

4. Лексикология английского языка: Учебное пособие / Кругликова Е.А. - Краснояр.:СФУ, 2016. - 162 с.: ISBN 978-5-7638-3479-6 - Режим доступа: <http://znanium.com/catalog/product/978599>

Дополнительная литература
(печатные и электронные издания)

1. Баранов, А. Н. Лингвистическая экспертиза текста : теория и практика [Электронный ресурс] : учеб. пособие / А. Н. Баранов. - 4-е изд., стереотип. - М. : Флинта : Наука, 2012. - ISBN 978-5-9765-0083-9 (Флинта), ISBN 978-5-02-034561-4 (Наука) - Режим доступа: <http://znanium.com/catalog/product/454832>
2. Моисеева И.Ю. Квантитативная лингвистика и новые информационные технологии [Электронный ресурс]: учебное пособие/ Моисеева И.Ю.— Электрон. текстовые данные.— Оренбург: Оренбургский государственный университет, ЭБС АСВ, 2017.— 103 с.— Режим доступа: <http://www.iprbookshop.ru/71281.html>.
3. Беляева Л.Н. Лингвистические технологии в современном сетевом пространстве. Language worker в индустрии локализации [Электронный ресурс]: монография/ Беляева Л.Н.— Электрон. текстовые данные.— СПб.: Книжный дом, 2016.— 133 с.— Режим доступа: <http://www.iprbookshop.ru/71523.html>.
4. Кайсарова Д.В. Математическая лингвистика. Практикум [Электронный ресурс]/ Кайсарова Д.В., Коцюба И.Ю.— Электрон. текстовые данные.— СПб.: Университет ИТМО, 2016.— 67 с.— Режим доступа: <http://www.iprbookshop.ru/67815.html>.
5. Лайонз, Дж. Введение в теоретическую лингвистику / Дж. Лайонз ; пер. с англ. под ред. и с предисл. В. А. Звегинцева. – М. : Прогресс, 1978. – 543 с.
<http://lib.dvfu.ru:8080/lib/item?id=chamo:63721&theme=FEFU>
6. Мечковская, Н. Б. Общее языкознание: структурная и социальная типология языков: учебн. пособие / Н. Б. Мечковская. – М. : Флинта : Наука , 2007. – 312 с.
<http://lib.dvfu.ru:8080/lib/item?id=chamo:249150&theme=FEFU>
7. Соссюр, Ф. Курс общей лингвистики / Ф. де Соссюр. – Екатеринбург : Изд-во Уральского

университета, 1999. – 426 с.

<http://lib.dvfu.ru:8080/lib/item?id=chamo:366054&theme=FEFU>

8. Степанов, Ю. С. Основы языкознания : учебное пособие для аспирантов и студентов филологических факультетов педагогических институтов / Ю. С. Степанов. – М. : Просвещение, 1975. – 271 с.

<http://lib.dvfu.ru:8080/lib/item?id=chamo:59469&theme=FEFU>

Перечень ресурсов информационно-телекоммуникационной сети «Интернет»

Студенты могут получить доступ к электронным образовательным ресурсам через сайт ДВФУ НАУЧНО-ИНФОРМАЦИОННЫЕ ЭЛЕКТРОННЫЕ РЕСУРСЫ

(доступ с сайта Научной библиотеки ДВФУ) URL:
http://www.dvfu.ru/web/library/rus_res

1. Научная электронная библиотека eLIBRARY.RU URL:
<http://www.elibrary.ru>
2. Базы данных компании East View Publications URL:
<http://www.ebiblioteka.ru/>
3. Базы данных компании "Интегрум"
4. Электронная библиотека диссертаций РГБ URL: <http://diss.rsl.ru>
5. Единое окно доступа к образовательным ресурсам URL:
<http://window.edu.ru/>
6. Электронно-библиотечная система Издательства "Лань" URL:
<http://e.lanbook.com>
7. Электронно-библиотечная система znanium.com НИЦ "ИНФРА-М" URL: <http://znanium.com>
8. Информационная система - ЕДИНОЕ ОКНО доступа к образовательным ресурсам <http://window.edu.ru/>

а так же в свободном доступе в Интернет:

1. Gaudeamus [Электронная библиотека учебной PDF-литературы и учебников для вузов. (бесплатные полнотекстовые учебники)] URL: <http://www.gaudeamus.omskcity.com/>

Перечень информационных технологий и программного обеспечения:

1. Microsoft Word
2. Microsoft PowerPoint
3. Microsoft Internet Explorer/ Mozilla Firefox/ Opera
4. SketchEngine, AntConc (в свободном доступе)

VI. МЕТОДИЧЕСКИЕ УКАЗАНИЯ ПО ОСВОЕНИЮ ДИСЦИПЛИНЫ

Методические рекомендации к освоению дисциплины

Описание последовательности действий обучающихся, алгоритм изучения дисциплины:

- Систематизировать литературные источники,
- проанализировать и обобщить представленные в них концепции
- Из всего материала выделить существующие точки зрения на проблему,
- Проанализировать их, сравнить, дать им оценку.
- Итогом этой работы должна стать логически выстроенная система сведений по существу исследуемого вопроса.

рекомендации по подготовке реферата и эссе

Задание: выбрать тему, подготовить, оформить и сдать реферативное сообщение объемом 5-7 стандартных страниц. При написании пользоваться различными словарями (толковыми, энциклопедическими, терминологическими и др.), любыми другими источниками, список которых представить в конце реферата. В тексте реферата сделать ссылки на использованные источники.

VII. МАТЕРИАЛЬНО-ТЕХНИЧЕСКОЕ ОБЕСПЕЧЕНИЕ ДИСЦИПЛИНЫ

1.	Технологии корпусной лингвистики	Компьютерный класс: 21 шт. моноблоки HP 3420 AIO; мультимедийный проектор OptimaEX542I – 1 шт; аудио усилитель QVC RMX 850 – 1 шт; колонки – 1 шт; ноутбук; ИБП – 1 шт; настенный экран; микрофон – 1 шт.	690001, Приморский край, г. Владивосток, о. Русский, кампус ДВФУ, корпус D, ауд.D604
		Лекционная аудитория: мультимедийный проектор OptimaEX542I – 1 шт; аудио усилитель QVC RMX 850 – 1 шт; колонки – 1 шт; ноутбук; ИБП – 1 шт; настенный экран; микрофон – 1 шт.	690001, Приморский край, г. Владивосток, о. Русский, кампус ДВФУ, корпус D, ауд.D446

В целях обеспечения специальных условий обучения инвалидов и лиц с ограниченными возможностями здоровья в ДВФУ все здания оборудованы пандусами, лифтами, подъемниками, специализированными местами, оснащенными туалетными комнатами, табличками информационно-навигационной поддержки.



МИНИСТЕРСТВО ОБРАЗОВАНИЯ И НАУКИ РОССИЙСКОЙ ФЕДЕРАЦИИ
Федеральное государственное автономное образовательное учреждение высшего образования
«Дальневосточный федеральный университет»
(ДВФУ)

**ВОСТОЧНЫЙ ИНСТИТУТ- ШКОЛА РЕГИОНАЛЬНЫХ И МЕЖДУНАРОДНЫХ
ИССЛЕДОВАНИЙ**

**УЧЕБНО-МЕТОДИЧЕСКОЕ ОБЕСПЕЧЕНИЕ САМОСТОЯТЕЛЬНОЙ
РАБОТЫ ОБУЧАЮЩИХСЯ**

по дисциплине « Технологии корпусной лингвистики»

**Направление подготовки 45.03.03 Фундаментальная и прикладная
лингвистика**

профиль «Фундаментальная и прикладная лингвистика»

Форма подготовки очная

**Владивосток
2019**

План-график выполнения самостоятельной работы по дисциплине

№ п/п	Дата/сроки выполнения	Вид самостоятельной работы	Примерные нормы времени на выполнение	Форма контроля
1	1.09.– 21.09.	Изучение темы: Предмет, цели и задачи корпусной лингвистики	3 нед.	УО-1 собеседование
2	22.09. – 15.10	Изучение темы: Методика лингвистического корпуса	3 нед.	УО-1 собеседование
3	2.10. – 6.11.	Изучение темы: Типы корпусов	3 нед.	УО-2 коллоквиум
4	7.11 - 30.11	Эссе на тему: Лексические корпусные лингвистические исследования .	3 нед.	эссе (ПР-3)
5	1.12 – 21.12.	Реферат на тему Корпусные лингвистические исследования	3 нед.	рефераты (ПР-4)

Характеристика заданий для самостоятельной работы обучающихся и методические рекомендации по их выполнению

Самостоятельная работа студентов предполагает изучение нижеследующих тем, используя нижеприведенную литературу. На основе изученного материала студентам предлагается ответить на вопросы, представленные после тем и выполнить письменный анализ проблемных вопросов в виде эссе и рефератов.

Методические рекомендации к самостоятельной работе:
Систематизировать литературные источники, проанализировать и обобщить представленные в них концепции. Из всего материала выделить существующие точки зрения на проблему, проанализировать их, сравнить, дать им оценку. Итогом этой работы должна стать логически выстроенная система сведений по существу исследуемого вопроса.

Выполнить письменный анализ проблемных вопросов по теме реферативного сообщения. Обосновать и применить к материалу одну из нескольких предложенных концепций – привести аргументы в пользу правильности выбранного варианта ответа и указать, в чем ошибочность других вариантов.

Аргументировать выбор адекватного способа действий, ведущего к решению задачи. При подготовке реферата указывать названия источников, авторов, год издания.

Темы:

1. Лексикографические исследования, основанные на корпусах
2. Пример одного лексикографического исследования
3. Выделение коллокаций статистическими методами
4. Лингвистические исследования на базе корпуса: изучение грамматики.
5. Распределение и функции номинализаций.
6. Распределение грамматических категорий. Семантическое наполнение Британского национального корпуса.
7. Способы использования корпусов.
8. Семантическое наполнение Британского национального корпуса.
9. Методы извлечения информации из корпуса. Типы извлекаемой информации.
10. Программы для работы с корпусом: конкордансеры и другие.
11. Статистические меры вычисления совместной встречаемости
12. Использование корпусов в социологии, истории.

Вопросы и задания для устного контроля самостоятельной работы студентов (3–5 вопросов)

Вариант 1.

1. Расскажите о становлении корпусной лингвистики как раздела компьютерной лингвистики, ее целях, задачах, предмете и объекте исследования.

2. Покажите двойственный характер корпусной лингвистики. 3. Какие лингвистические направления предвосхитили появление корпусной лингвистики.

3. Сопоставьте корпусную и традиционную лингвистику. В чем специфика корпусного подхода к языку

Вариант 2.

1. Каковы предпосылки создания и возможности использования корпусов в лингвистических исследованиях, лексикографии, переводческой деятельности и в обучении иностранным языкам. 4. Как развивались основные характеристики корпусов.

2. Классификации корпусов по разным основаниям.

3. Покажите основные процедуры обработки естественного языка в корпусе.

4.какие программы производят процессы токенизации, лемматизации, стемминга и парсинга: специфика, цели, результаты.

5. Сопоставьте корпусные менеджеры.

Перечень тем рефератов и эссе приведен в Приложении 2.

Требования к представлению и оформлению результатов самостоятельной работы

Оформить и сдать реферативное сообщение объемом 5-7 стандартных страниц. При написании пользоваться различными словарями (толковыми, энциклопедическими, терминологическими и др.), любыми другими источниками, список которых представить в конце реферата. В тексте реферата сделать ссылки на использованные источники.

Критерии оценки выполнения самостоятельной работы.

- Самостоятельная работа оценивается на «отлично», если студент демонстрирует глубокие знания теории в области типологии, излагает различные точки зрения и концепции с указанием авторов, владеет соответствующей научной терминологией, полно отвечает на дополнительные вопросы.
- Ответ оценивается на «хорошо», если студент, в целом владея теоретическим материалом, допускает отдельные неточности в интерпретации языковых и литературных явлений.
- Студент получает «удовлетворительно», если допускает значительные неточности и лакуны в изложении теоретического материала
- Оценка «неудовлетворительно» выставляется, если теоретические вопросы остаются нераскрытыми.



МИНИСТЕРСТВО ОБРАЗОВАНИЯ И НАУКИ РОССИЙСКОЙ ФЕДЕРАЦИИ
Федеральное государственное автономное образовательное учреждение высшего образования
«Дальневосточный федеральный университет»
(ДФУ)

**ВОСТОЧНЫЙ ИНСТИТУТ- ШКОЛА РЕГИОНАЛЬНЫХ И МЕЖДУНАРОДНЫХ
ИССЛЕДОВАНИЙ**

ФОНД ОЦЕНОЧНЫХ СРЕДСТВ
по дисциплине «Технологии корпусной лингвистики»
Направление подготовки 45.03.03 Фундаментальная и прикладная
лингвистика
профиль «Фундаментальная и прикладная лингвистика»
Форма подготовки очная

Владивосток
2019

ПАСПОРТ ФОС

Код и формулировка компетенции	Этапы формирования компетенции	
ПК-3. Владеет методами сбора и документации лингвистических данных	Знает	основные методы документирования естественного языка; принципы полевой работы с носителями естественных языков; технологии проведения сбора языковых данных; базовые принципы передачи и хранения собранной информации.
	Умеет	производить запись речевого сигнала на цифровой носитель; аннотировать, обрабатывать и анализировать собранные данные; собирать методанные, в частности, социолингвистические данные о носителе языка.
	Владеет	Имеет практический опыт документирования лингвистических данных: записи речевого сигнала с последующей обработкой и анализом; обработки, структурирования и хранения метаинформации о собранных данных
ПК-4. Способен спланировать и провести лингвистический эксперимент, описать его результаты и сформулировать выводы	Знает	общие положения о специфике и архитектуре эксперимента; базовые принципы проведения экспериментов с носителями языка; принципы составления анкет и вопросников; методы создания лингвистических экспериментов с привлечением достижений современных технологий; имеет базовые представления о методах математической статистики, используемых при обработке результатов эксперимента.
	Умеет	разработать и провести базовый лингвистический эксперимент; отобрать испытуемых; обобщить и проанализировать полученные данные; сформулировать результат.
	Владеет	Имеет практический опыт проведения базовых лингвистических экспериментов; разработки их архитектуры; поиска испытуемых; обработки результатов.

ПК-8. Владеет навыками оформления и представления результатов научного исследования	Знает	общие представления о способе подачи научных материалов; методы оформления результатов научного исследования в виде статей, квалификационных работ, постеров, презентаций, докладов на конференциях и семинарах.
	Умеет	грамотно и в соответствии с требованиями оформить научную работу, раздаточный материал или презентацию к докладу; сделать устный доклад в соответствии с регламентом.
	Владеет	Имеет практический опыт выступлений перед аудиторией, а также создания письменных работ для представления результатов научного исследования
ПК-9.Способен пользоваться лингвистически ориентированными программными продуктами	Знает	основные типы систем, использующих модули лингвистического анализа; основные принципы и методы компьютерного моделирования лингвистических задач.
	Умеет	Анализировать работу различных систем обработки текста и звучащей речи для выявления основных лингвистических компонентов и основных типов обработки текста, используемых в данных системах; подбирать необходимые лингвистические ресурсы для различных задач лингвистического обеспечения систем (например, лексикографических, задач морфологического анализа и т.п.).
	Владеет	Имеет практический опыт работы с различными системами автоматической и экспертной обработки текста и звучащей речи.
ПК-10. Владеет принципами создания электронных языковых ресурсов (текстовых, речевых и мультимодальных корпусов; словарей, тезаурусов, онтологий; фонетических, лексических,	Знает	Основные принципы обработки информации; базовые принципы корпусной лингвистики, лексикографии, математической статистики; базовые представления о языковом разнообразии; наиболее полные и значимые лингвистические корпуса, электронные словари и базы данных.

грамматических и иных баз данных и баз знаний) и умеет пользоваться такими ресурсами	Умеет	пользоваться основными методами, способами и средствами получения, хранения, переработки информации; пользоваться лингвистически ориентированными программными продуктами.
	Владеет	основными принципами обработки информации; базовыми принципами корпусной лингвистики, лексикографии, математической статистики; базовые представления о языковом разнообразии; наиболее полные и значимые лингвистические корпуса, электронные словари и базы данных. пользоваться основными методами, способами и средствами получения, хранения, переработки информации; пользоваться лингвистически ориентированными программными продуктами.

№ п/п	Контролируемые разделы / темы дисциплины	Коды и этапы формирования компетенций		Оценочные средства	
				текущий контроль	промежуточная аттестация
1	Введение в корпусную лингвистику	ПК-3	Знает	собеседование (УО-1)	Вопросы № 1,2 экзаменационного билета, билеты 1-12
			Умеет	собеседование (УО-1)	
			Владеет	собеседование (УО-1)	
		ПК-4	Знает	собеседование (УО-1)	
			Умеет	собеседование (УО-1)	
			Владеет	реферат (ПР-4)	
2.	корпусные лингвистические исследования	ПК-9	Знает	собеседование (УО-1)	Вопросы № 1,2 экзаменационного билета, билеты 1-12
			Умеет	реферат (ПР-4)	

			Владеет	собеседовани е (УО-1)	
		ПК-8	Знает	собеседовани е (УО-1)	
			Умеет	собеседовани е (УО-1)	
			Владеет	собеседовани е (УО-1)	
3	корпусные технологии	ПК-10	Знает	собеседовани е (УО-1)	Вопросы № 1,2 экзаменационн ого билета, билеты 1-12
			Умеет	собеседовани е (УО-1)	Вопросы № 1,2 экзаменационн ого билета, билеты 1-12
			Владеет	собеседовани е (УО-1)	Вопросы № 1,2 экзаменационн ого билета, билеты 1-12

Методические рекомендации, определяющие процедуры оценивания результатов освоения дисциплины

Текущая аттестация студентов по дисциплине проводится в соответствии с локальными нормативными актами ДВФУ и является обязательной. Текущая аттестация по дисциплине проводится в форме собеседования и коллоквиума и осуществляется ведущим преподавателем.

Объектами оценивания выступают:

- активность на занятиях, своевременность выполнения различных видов заданий, посещаемость всех видов занятий по аттестуемой дисциплине;
- степень усвоения теоретических знаний;
- уровень овладения практическими умениями и навыками по всем видам учебной работы;
- результаты самостоятельной работы.

Промежуточная аттестация студентов по дисциплине проводится в соответствии с локальными нормативными актами ДВФУ и является обязательной, в 5 семестрах предусмотрен экзамен

Оценочные средства для промежуточной аттестации

Содержание экзамена

Вопросы к экзамену

1. Британский национальный корпус.
2. История лингвистических корпусов: от картотеки к корпусу.

3. Классификация (типология) корпусов.
4. Корпусная лингвистика: современное состояние.
5. Основные понятия корпусной лингвистики.
6. Направления в лингвистике, предвосхитившие появление корпусной лингвистики: от картотеки к корпусу
7. История создания лингвистических корпусов.
8. Основные характеристики корпусов.
9. Репрезентативность корпусов.
10. Классификация корпусов по различным основаниям.
11. Особые типы корпусов.
12. Понятие разметки.
13. Разметка. Средства разметки корпусов
14. Лингвистическая разметка.
15. Экстралингвистическая разметка.
16. Стандартизация в корпусной лингвистике.
17. Обзор существующих корпусов различных типов.
18. Зарубежные национальные корпуса.
19. Специальные корпуса.
20. Корпусные исследования.
21. Пользователи корпусов.
22. Способы использования корпусов.
23. Лексикографические исследования, основанные на корпусах.
24. Понятие гранулярности.

Критерии выставления оценки студенту на экзамене

Баллы (рейтингов)	Оценка зачета/	Требования к сформированным компетенциям
------------------------------	---------------------------	---

ой оценки)	экзамена (стандартная)	
86-100	<i>«отлично»</i>	Оценка «отлично» выставляется студенту, если он глубоко и прочно усвоил программный материал, исчерпывающе, последовательно, четко и логически стройно его излагает, умеет тесно увязывать теорию с практикой, свободно справляется с задачами, вопросами и другими видами применения знаний, причем не затрудняется с ответом при видоизменении заданий, правильно обосновывает принятое решение, владеет разносторонними навыками и приемами выполнения практических задач.
76-85	<i>«хорошо»</i>	Оценка «хорошо» выставляется студенту, если он твердо знает материал, грамотно и по существу излагает его, не допуская существенных неточностей в ответе на вопрос, правильно применяет теоретические положения при решении практических вопросов и задач, владеет необходимыми навыками и приемами их выполнения.
61-75	<i>«удовлетворительно»</i>	Оценка «удовлетворительно» выставляется студенту, если он имеет знания только основного материала, но не усвоил его деталей, допускает неточности, недостаточно правильные формулировки, нарушения логической последовательности в изложении программного материала, испытывает затруднения при выполнении практических работ.
< 61	<i>«неудовлетворительно»</i>	Оценка «неудовлетворительно» выставляется студенту, который не знает значительной части программного материала, допускает существенные ошибки, неуверенно, с большими затруднениями выполняет практические работы. Как правило, оценка «неудовлетворительно» ставится студентам, которые не могут продолжить обучение без дополнительных занятий по соответствующей дисциплине.

Оценочные средства для текущей аттестации

Темы рефератов

1. Лексикографические исследования, основанные на корпусах
2. Пример одного лексикографического исследования
3. Выделение коллокаций статистическими методами
4. Лингвистические исследования на базе корпуса: изучение грамматики.
5. Распределение и функции номинализаций.
6. Распределение грамматических категорий. Семантическое наполнение Британского национального корпуса.
7. Способы использования корпусов.
8. Семантическое наполнение Британского национального корпуса.
9. Методы извлечения информации из корпуса. Типы извлекаемой информации.
10. Программы для работы с корпусом: конкордансеры и другие.

11. Статистические меры вычисления совместной встречаемости

12. Использование корпусов в социологии, истории.

Критерии оценивания реферата

- реферат оценивается на «отлично», если студент демонстрирует глубокие знания теории в области лингвистической типологии, излагает различные точки зрения и концепции с указанием авторов, владеет соответствующей научной терминологией, полно отвечает на дополнительные вопросы.

- реферат оценивается на «хорошо», если студент, в целом владея теоретическим материалом, допускает отдельные неточности в интерпретации языковых и литературных явлений.

- Студент получает «удовлетворительно», если допускает значительные неточности и лакуны в изложении теоретического материала

- Оценка «неудовлетворительно» выставляется, если теоретические вопросы остаются нераскрытыми.

Вопросы для собеседования

1. Чем занимается корпусная лингвистика? В чем заключается двойственность объекта ее исследования?
2. Какие направления послужили основой для развития корпусной лингвистики?
3. В чем состоят принципиальные различия между традиционной и корпусной лингвистикой?
4. Что такое языковой (лингвистический) корпус?
5. В чем состоят цели создания и использования языковых корпусов?
6. Что понимается под репрезентативностью корпуса?
7. Охарактеризуйте существующие классификации корпусов по разным критериям.
8. Что такое «токен» и в чем состоит процесс токенизации?
9. Что такое «лемма» и в чем состоит процесс лемматизации?
10. Чем отличается процесс лемматизации от стемминга?
11. Что такое парсинг и для чего он проводится?
12. Что такое разметка и для чего она проводится? Назовите типы лингвистических и экстралингвистических тэгов.
13. Дайте характеристику следующим типам разметки: морфологической, синтаксической, семантической, анафорической, просодической и дискурсной.
14. В чем заключается проблема морфологической неоднозначности и как она решается?

15. Назовите функции корпусных менеджеров.
16. Что такое конкорданс и для чего он используется?
17. Опишите общую модель языка запросов.
18. Дайте краткую характеристику Национального корпуса русского языка (состав и структура, виды разметки и т.д.)
19. Дайте краткую характеристику Корпуса современного американского варианта английского языка (Corpus of Contemporary American English).
20. Обобщите основные возможности использования языковых корпусов в научно-исследовательской и практической переводческой деятельности, а также в области обучения иностранным языкам.

Вопросы для коллоквиума

1. Ручное аннотирование текста
2. Сравнение результатов аннотирования текста
3. Использование цепочки обрабатывающих ресурсов
4. Создание и применение словаря
5. Применение JARE для автоматического аннотирования
6. Использование Searchable DataStore
7. Аннотирование текста с использованием JAVA
8. Извлечение информации из коллекции гомогенных текстов

Критерии оценивания ответа при собеседовании и на коллоквиуме

- ответ оценивается на «отлично», если студент демонстрирует глубокие знания теории в области лингвистической типологии, излагает различные точки зрения и концепции с указанием авторов, владеет соответствующей научной терминологией, полно отвечает на дополнительные вопросы.

- Ответ оценивается на «хорошо», если студент, в целом владея теоретическим материалом, допускает отдельные неточности в интерпретации языковых и литературных явлений.

- Студент получает «удовлетворительно», если допускает значительные неточности и лакуны в изложении теоретического материала

- Оценка «неудовлетворительно» выставляется, если теоретические вопросы остаются нераскрытыми.