



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

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

«СОГЛАСОВАНО»
Руководитель ОП
Филология

Спицына Н.А.

«11» июля 2019 г.

«УТВЕРЖДАЮ»
Заведующая кафедрой
лингвистики и межкультурной коммуникации

Ловцевич Г.Н.

«11» июля 2019 г.

РАБОЧАЯ ПРОГРАММА УЧЕБНОЙ ДИСЦИПЛИНЫ

Технологии корпусной лингвистики

**Направление подготовки 45.03.03 Фундаментальная и прикладная лингвистика
профиль «Фундаментальная и прикладная лингвистика (английский и французский)»**

Форма подготовки очная

курс 3 семестр 5
лекции 18 час.
практические занятия час.
лабораторные работы час.
в том числе с использованием МАО лек. 6 /пр. /лаб. час.
в том числе в электронной форме лек. /пр. /лаб. час
всего часов аудиторной нагрузки 18 часа.
в том числе с использованием МАО 6 час.
самостоятельная работа 126 час.
в том числе на подготовку к экзамену 54 час.
курсовая работа / курсовой проект семестр
зачет семестр
экзамен 5 семестр

Рабочая программа составлена в соответствии с требованиями образовательного стандарта, самостоятельно устанавливаемого ДВФУ, утвержденного приказом ректора от 18.02.2016 № 12-13-235.

Рабочая программа дисциплины обсуждена на заседании кафедры лингвистики и межкультурной коммуникации, протокол № 11 от «10» июля 2019 г.

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

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)

Аннотация к рабочей программе 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 (английский и французский)».

Общая трудоемкость освоения дисциплины составляет 4 зачётные единицы, 144 часа. Учебным планом предусмотрены лекции (18 часов), самостоятельная работа (126 часов, том числе 54 часа на подготовку к экзамену). Дисциплина реализуется в 5 семестре, промежуточная форма отчетности - экзамен.

Дисциплина «Технологии корпусной лингвистики» является дисциплиной по выбору вариативной части Блока «Дисциплины (модули)».

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

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)», «Современные информационные технологии», «Компьютерная лексикография и информационные технологии».

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;

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 компетенций (профессиональные компетенции (элементы компетенций)):

Код и формулировка компетенции	Этапы формирования компетенции	
ПК – 3 Владение методами сбора и документации лингвистических данных	Знает	базовые методы сбора и документации лингвистических данных при помощи корпусных менеджеров
	Умеет	использовать методику сбора и документации лингвистических данных для анализа и оценивания различных фактов и явлений на основе корпусного менеджера AntConc
	Владеет	навыками анализа проблем, возникающих при решении исследовательских и практических задач на основе сбора и документации лингвистических данных в SketchEngine (WebBootCat)
ПК – 7: Владение параметрами разнообразия естественных языков и их ареальной,	Знает	Методы проведения классификации языков
	Умеет	Проводить научное исследование в соответствии с поставленной целью и

типологической и генеалогической классификации		задачами, определять логику проведения научного исследования учитывая новизну собственных концептуальных решений
	Владеет	Инструментами и методами проведения научных исследований, методами анализа и обоснования эффективности дискурсивного анализа,
ПК-8: владение навыками оформления и представления результатов научного исследования	Знает	методы описания языка через создание собственных учебных корпусов
	Умеет	применять современные текстовые, корпусы; словари, и осуществлять отбор и использование материала для своего корпуса
	Владеет	навыками использования современных ресурсов автоматической обработки естественного языка
ПК-11: владение принципами создания электронных языковых ресурсов (текстовых, речевых и мультимодальных корпусов; словарей, тезаурусов, онтологии; фонетических, лексических, грамматических и иных баз данных и баз знаний) и умением пользоваться такими ресурсами	Знает	функциональные свойства корпусных менеджеров 1-4 поколений
	Умеет	осуществлять изучение структурных и функциональных свойств языков при помощи корпусных менеджеров и систем автоматической разметки
	Владеет	навыками исследования структурных и функциональных свойств корпусных менеджеров и систем автоматической разметки
ПК-13: способность проводить квалифицированное тестирование лингвистически ориентированных программных продуктов, электронных ресурсов, лингвистически ориентированных систем и лингвистических компонентов интеллектуальных и информационных электронных систем	Знает	основные положения лингвистически ориентированных программных продуктов, электронных ресурсов, лингвистически ориентированных систем и лингвистических компонентов интеллектуальных и информационных электронных систем;
	Умеет	устанавливать отличительные компоненты интеллектуальных и информационных электронных систем;

	Владеет	основными способами квалифицированного тестирования лингвистически ориентированных программных продуктов;
--	---------	---

Обучение корпусным технологиям, основаны на принципах, именуемых MAO 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

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 MAO)

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 MAO)

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

практические занятия не предусмотрены учебным планом

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.

I. 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	базовые методы сбора и документации лингвистических данных при помощи корпусных менеджеров	собеседование (УО-1)	Вопросы № 1,2 экзаменационного билета, билеты 1-12
			Умеет	использовать методику сбора и документации лингвистических данных для анализа и оценивания различных фактов и явлений на основе корпусного менеджера AntConc	собеседование (УО-1)	
			Владеет	навыками анализа проблем, возникающих при решении исследовательских и практических задач на основе сбора и документации	собеседование (УО-1)	

				лингвистических данных в SketchEngine (WebBootCat)		
		ПК-7	Знает	методы описания языка через его системное сравнение в программном продукте SketchEngine	собеседование (УО-1)	
			Умеет	применять современные научные парадигмы прикладной лингвистики (SketchEngine)	собеседование (УО-1)	
			Владеет	навыками использования современных методов прикладной лингвистики (SketchEngine)	реферат (ПР-4)	
2.	корпусные лингвистические исследования	ПК-8	Знает	методы описания языка через создание собственных учебных корпусов	собеседование (УО-1)	Вопросы № 1,2 экзаменационного билета, билеты 1-12
			Умеет	применять современные текстовые, корпусы; словари, и осуществлять отбор и использование материала для своего корпуса	реферат (ПР-4)	
			Владеет	навыками использования современных ресурсов автоматической обработки естественного языка	собеседование (УО-1)	
		ПК-11, ПК-13	Знает	функциональные свойства корпусных менеджеров 1-4 поколений	собеседование (УО-1)	
			Умеет	осуществлять изучение структурных и функциональных свойств языков при помощи корпусных менеджеров и систем	собеседование (УО-1)	

				автоматической разметки		
			Владеет	навыками исследования структурных и функциональных свойств корпусных менеджеров и систем автоматической разметки	собеседование (УО-1)	

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.

КОНТРОЛЬ ДОСТИЖЕНИЯ ЦЕЛЕЙ КУРСА

Контроль осуществляется посредством экзамена (5 семестр).

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 корпусы.

19. Специальные корпуса.
20. Корпусные исследования.
21. Пользователи корпусов.
22. Способы использования корпусов.
23. Лексикографические исследования, основанные на корпусах.
24. Понятие гранулярности.

II. УЧЕБНО-МЕТОДИЧЕСКОЕ ОБЕСПЕЧЕНИЕ ДИСЦИПЛИНЫ

Основная литература

(печатные и электронные издания)

1. Корпусная лингвистика [Электронный ресурс] / Грудева Е.В. - М. : ФЛИНТА, 2017. - <http://www.studentlibrary.ru/book/ISBN9785976514973.html>
2. 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>
3. 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. - Красноярск: 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>.

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. 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)

IV 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.

V. 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»
(ДФУ)

**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 (английский и французский)»

Форма подготовки очная

**Владивосток
2017**

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 нижеприведенных 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 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 (английский и французский)»
Форма подготовки очная

Владивосток
2017

ПАСПОРТ ФОС

Код и формулировка компетенции	Этапы формирования компетенции	
ПК – 3 Владение методами сбора и документации лингвистических данных	Знает	базовые методы сбора и документации лингвистических данных при помощи корпусных менеджеров
	Умеет	использовать методику сбора и документации лингвистических данных для анализа и оценивания различных фактов и явлений на основе корпусного менеджера AntConc
	Владеет	навыками анализа проблем, возникающих при решении исследовательских и практических задач на основе сбора и документации лингвистических данных в SketchEngine (WebBootCat)
ПК – 7: Владение параметрами разнообразия естественных языков и их ареальной, типологической и генеалогической классификации	Знает	Методы проведения классификации языков
	Умеет	Проводить научное исследование в соответствии с поставленной целью и задачами, определять логику проведения научного исследования учитывая новизну собственных концептуальных решений
	Владеет	Инструментами и методами проведения научных исследований, методами анализа и обоснования эффективности дискурсивного анализа,
ПК-8: владение навыками оформления и представления результатов научного исследования	Знает	методы описания языка через создание собственных учебных корпусов
	Умеет	применять современные текстовые, корпусы; словари, и осуществлять отбор и использование материала для своего корпуса
	Владеет	навыками использования современных ресурсов автоматической обработки естественного языка
ПК-11: владение принципами создания электронных языковых ресурсов (текстовых, речевых и мультимодальных корпусов);	Знает	функциональные свойства корпусных менеджеров 1-4 поколений
	Умеет	осуществлять изучение структурных и функциональных свойств языков при помощи корпусных менеджеров и систем автоматической разметки

словарей, тезаурусов, онтологии; фонетических, лексических, грамматических и иных баз данных и баз знаний) и умением пользоваться такими ресурсами	Владеет	навыками исследования структурных и функциональных свойств корпусных менеджеров и систем автоматической разметки
ПК-13: способность проводить квалифицированное тестирование лингвистически ориентированных программных продуктов, электронных ресурсов, лингвистически ориентированных систем и лингвистических компонентов интеллектуальных и информационных электронных систем	Знает	основные положения лингвистически ориентированных программных продуктов, электронных ресурсов, лингвистически ориентированных систем и лингвистических компонентов интеллектуальных и информационных электронных систем;
	Умеет	устанавливать отличительные компоненты интеллектуальных и информационных электронных систем;
	Владеет	основными способами квалифицированного тестирования лингвистически ориентированных программных продуктов;

№ п/п	Контролируемые разделы / темы дисциплины	Коды и этапы формирования компетенций		Оценочные средства	
				текущий контроль	промежуточная аттестация
1	Введение в корпусную лингвистику	ПК-3 владение методами сбора и документации лингвистических данных	Знает	базовые методы сбора и документации лингвистических данных при помощи корпусных менеджеров	собеседование (УО-1) Вопросы № 1,2 экзаменационного билета,

			Умеет	использовать методику сбора и документации лингвистических данных для анализа и оценивания различных фактов и явлений на основе корпусного менеджера AntConc	собеседование (УО-1)	билеты 1-12
			Владеет	навыками анализа проблем, возникающих при решении исследовательских и практических задач на основе сбора и документации лингвистических данных в SketchEngine (WebBootCat)	собеседование (УО-1)	
		ПК-8 владение навыками оформления и представления результатов научного исследования	Знает	методы описания языка через его системное сравнение в программном продукте SketchEngine	собеседование (УО-1)	
			Умеет	применять современные научные парадигмы прикладной лингвистики (SketchEngine)	собеседование (УО-1)	
			Владеет	навыками использования современных методов прикладной лингвистики (SketchEngine)	реферат (ПР-4)	
2.	корпусные лингвистические исследования	ПК-10 владение принципами создания электронных языковых ресурсов (текстовых, речевых и мультимодальных корпусов; словарей, тезаурусов,	Знает	методы описания языка через создание собственных учебных корпусов	собеседование (УО-1)	Вопросы № 1,2 экзаменационного билета, билеты 1-12
			Умеет	применять современные текстовые, корпусы; словари, и осуществлять отбор и использование материала для своего корпуса	реферат (ПР-4)	

	онтологии; фонетических, лексических, грамматических и иных баз данных и баз знаний) и умением пользоваться такими ресурсами	Владеет	навыками использования современных ресурсов автоматической обработки естественного языка	собеседование (УО-1)
	ПК-12 способность проводить квалифицированное тестирование лингвистически ориентированных программных продуктов, электронных ресурсов, лингвистически ориентированных систем и лингвистических компонентов интеллектуальных и информационных электронных систем	Знает	функциональные свойства корпусных менеджеров 1-4 поколений	собеседование (УО-1)
		Умеет	осуществлять изучение структурных и функциональных свойств языков при помощи корпусных менеджеров и систем автоматической разметки	собеседование (УО-1)
		Владеет	навыками исследования структурных и функциональных свойств корпусных менеджеров и систем автоматической разметки	собеседование (УО-1)

Шкала оценивания уровня сформированности компетенций

Код и формулировка компетенции	Этапы формирования компетенции		критерии	показатели
ПК – 3 Владение методами сбора и документации и лингвистических данных	знает (пороговый уровень)	базовые методы сбора и документации лингвистических данных при помощи корпусных менеджеров	знание основных понятий по методам сбора и документации лингвистических данных;	- способность перечислить и раскрыть суть методов сбора и документации лингвистических данных;; -способность самостоятельно сформулировать проблематику проекта; - способность обосновать

				актуальность проектного решения; -способность перечислить источники информации по методам и подходам к разработке проекта
	умеет (продвинутый)	использовать методику сбора и документации лингвистических данных для анализа и оценивания различных фактов и явлений на основе корпусного менеджера AntConc	Умение работать с информацией из электронных и библиотечных источников, умение применять известные методы разработки проектных идей, умение представлять результаты существующих проектных решений по изучаемой проблеме и собственных исследований на их основе, умение применять творческий подход к решению задачи сбора и документации лингвистических данных;	способность работать с информацией из электронных и библиотечных источников;; - способность найти проектные разработки по похожей проблематике и обосновать объективность применения их результатов; - способность изучить основные этапы проектирования; - способность применять методы концептуальных решений при решении проектных задач
	владеет (высокий)	навыками анализа проблем, возникающих при решении исследовательских и практических задач на основе сбора и документации лингвистических данных в SketchEngine (WebBootCat)	Владение терминологией проблемного проектирования, владение способностью сформулировать задание сбора и документации лингвистических данных;	- способность бегло и точно применять терминологический аппарат при обосновании своих предложений при разработке сбора и документации лингвистических данных;
ПК – 7: Владение параметрами разнообразия естественных языков и их ареальной, типологической и генеалогической классификации	знает (пороговый уровень)	Методы проведения классификации языков	знание основных понятий по методам научных исследований; знание методов научных исследований; знает источники информации по методам и подходам к проведению исследований	- способность перечислить и раскрыть суть методов научного исследования; -способность самостоятельно сформулировать объект предмет и научного исследования; - способность обосновать актуальность выполняемого

				задания или исследования; -способность перечислить источники информации по методам и подходам к проведению исследований
	умеет (продвинутой)	Проводить научное исследование в соответствии с поставленной целью и задачами, определять логику проведения научного исследования учитывая новизну собственных концептуальных решений	Умение работать с электронными базами данных и библиотечными каталогами, умение применять известные методы научных исследований, умение представлять результаты проектных исследований по изучаемой проблеме и собственных исследований, умение применять методы научных исследований учитывая новизну собственных концептуальных решений	- способность работать с данными, каталогов для исследования; - способность найти труды учёных и обосновать объективность применения изученных результатов проектных решений; - способность изучить научные определения относительно объекта и предмета исследования; - способность применять методы научных исследований учитывая новизну собственных концептуальных решений
	знает (пороговый уровень)	Методы проведения классификации языков	знание основных понятий по методам научных исследований; знание методов научных исследований; знает источники информации по методам и подходам к проведению исследований	- способность перечислить и раскрыть суть методов научного исследования; -способность самостоятельно сформулировать объект предмет и научного исследования; - способность обосновать актуальность выполняемого задания или исследования; -способность перечислить источники информации по методам и подходам к проведению исследований
ПК-8: владение навыками	знает (пороговый уровень)	методы описания языка через создание собственных учебных	может дать определения основных понятий	способен перечислить и раскрыть

оформления и представления результатов научного исследования		корпусов	и категорий современной лингвистики, знает, где найти информацию	суть основных понятий и категорий современной лингвистики, знает, где найти информацию
	умеет (продвинутой)	Проводить несложные научные эксперименты и экспертизу;	Умение работать с электронными базами данных и библиотечными каталогами, умение применять известные методы научных исследований, умение представлять результаты проектных исследований по изучаемой проблеме и собственных исследований, умение применять методы научных исследований учитывая новизну собственных концептуальных решений	способность работать с данными, каталогов для исследования; - способность найти труды учёных и обосновать объективность применения изученных результатов проектных решений; - способность изучить научные определения относительно объекта и предмета исследования; - способность применять методы научных исследований учитывая новизну собственных концептуальных решений
	владеет (высокий)	Инструментами и методами проведения научных исследований, методами анализа и обоснования эффективности	Владение терминологией предметной области знаний, владение способностью сформулировать проблематику научной статьи;	- способность бегло и точно применять терминологический аппарат предметной области исследования, - способность сформулировать задание по научному исследованию; - способность проводить самостоятельные исследования и представлять их результаты на обсуждение на круглых столах, семинарах, научных конференциях.
ПК-11: владение принципами создания электронных языковых ресурсов	знает (пороговый уровень)	основные положения прикладной филологической деятельности когнитивные технологии в области	лингвистические маркеры социальных отношений	способность применять методы выражения фактуальной информации в иноязычном тексте способность

(текстовых, речевых и мультимодальных корпусов; словарей, тезаурусов, онтологий; фонетических, лексических, грамматических и иных баз данных и баз знаний) и умением пользоваться такими ресурсами		прикладной филологической деятельности		применять маркеры социальных отношений
	умеет (продвинутой)	устанавливать отличительные лингвистические особенности языка; применять когнитивные технологии в прикладной филологической деятельности	определить тип социальных отношений адекватно использовать лингвистические маркеры социальных отношений	- способность перечислить и раскрыть основные ресурсы создания информационных электронных систем;
	владеет (высокий)	навыками исследования структурных и функциональных свойств корпусных менеджеров и систем автоматической разметки	владение способностью выполнять письменный перевод компьютерном текстовом редакторе	способность выполнить письменный перевод компьютерном текстовом редакторе; способность правильно пользоваться методикой подготовки к выполнению перевода, включая поиск информации в справочной, специальной литературе и компьютерных сетях.
ПК-13: способность проводить квалифицированное тестирование лингвистически ориентированных программных продуктов, электронных ресурсов, лингвистически ориентированных систем и лингвистических компонентов интеллектуальных и информационных электронных систем	знает (пороговый уровень)	основные положения лингвистически ориентированных программных продуктов, электронных ресурсов, лингвистически ориентированных систем и лингвистических компонентов интеллектуальных и информационных электронных систем;	системы автоматической обработки звучащей речи и письменного текста на естественном языке	способность применять методы работы в среде AntConc
	умеет (продвинутой)	устанавливать отличительные компонент интеллектуальных и информационных электронных систем;	Определить лингвистические компоненты интеллектуальных и информационных электронных систем	- способность перечислить и раскрыть основные ресурсы создания информационных электронных систем;
	владеет (высокий)	основными способами квалифицированного тестирования лингвистически ориентированных программных продуктов;	Способами проведения квалифицированного тестирования лингвистически ориентированных программных продуктов, электронных ресурсов,	- способность работать с данными, каталогов во время работы ; - способность изучить научные определения выделенных явлений с позиций

			лингвистически ориентированных систем и лингвистических компонентов интеллектуальных и информационных электронных систем;	фундаментальной лингвистики; - способность применять методы научных исследований для нестандартного решения проблемы терминологической многозначности;
--	--	--	---	---

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 семестре 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 ой оценки)	Оценка зачета/ экзамена (стандартная)	Требования к сформированным компетенциям <i>Дописать оценку в соответствии с компетенциями. Привязать к дисциплине</i>
86-100	«отлично»	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.

Дополнительные вопросы

1. Ручное аннотирование текста
2. Сравнение результатов аннотирования текста
3. Использование цепочки обрабатывающих ресурсов
4. Создание и применение словаря
5. Применение JARF 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

- ответ оценивается на «отлично», если студент демонстрирует глубокие знания теории в области корпусной лингвистики, 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 корпусных явлений.

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

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