

МИНИСТЕРСТВО ОБРАЗОВАНИЯ И НАУКИ РОССИЙСКОЙ ФЕДЕРАЦИИ
Федеральное государственное автономное учреждение
высшего профессионального образования
"Казанский (Приволжский) федеральный университет"
Отделение русской и зарубежной филологии



УТВЕРЖДАЮ

Проректор
по образовательной деятельности КФУ
Проф. Минзарипов Р.Г.

"__" _____ 20__ г.

Программа дисциплины

Корпусные технологии в лингвистике Б1.В.ОД.11

Направление подготовки: 45.04.01 - Филология
Профиль подготовки: Прикладное языкознание
Квалификация выпускника: магистр
Форма обучения: очное
Язык обучения: русский

Автор(ы):

Каримуллина Г.Н.

Рецензент(ы):

-

СОГЛАСОВАНО:

Заведующий(ая) кафедрой: Галиуллин К. Р.

Протокол заседания кафедры No ___ от "___" _____ 201__ г

Учебно-методическая комиссия Института филологии и межкультурной коммуникации
(отделение русской и зарубежной филологии):

Протокол заседания УМК No ___ от "___" _____ 201__ г

Регистрационный No

Казань
2014

Содержание

1. Цели освоения дисциплины
2. Место дисциплины в структуре основной образовательной программы
3. Компетенции обучающегося, формируемые в результате освоения дисциплины /модуля
4. Структура и содержание дисциплины/ модуля
5. Образовательные технологии, включая интерактивные формы обучения
6. Оценочные средства для текущего контроля успеваемости, промежуточной аттестации по итогам освоения дисциплины и учебно-методическое обеспечение самостоятельной работы студентов
7. Литература
8. Интернет-ресурсы
9. Материально-техническое обеспечение дисциплины/модуля согласно утвержденному учебному плану

Программу дисциплины разработал(а)(и) доцент, к.н. Каримуллина Г.Н. кафедра прикладной лингвистики отделение русской и зарубежной филологии им. Л.Н.Толстого ,
Guzel.Karimullina@kpfu.ru

1. Цели освоения дисциплины

В результате освоения дисциплины студенты должны обладать теоретическими знаниями в области корпусной лингвистики, знать ключевые понятия, связанные с данным направлением прикладного языкознания; ориентироваться в проблемах различных направлений изучаемой дисциплины; овладеть навыками лингвистического поиска в сети Интернет; анализа корпусов текстов, а также их применения в лингвистических исследованиях.

2. Место дисциплины в структуре основной образовательной программы высшего профессионального образования

Данная учебная дисциплина включена в раздел " Б1.В.ОД.11 Дисциплины (модули)" основной образовательной программы 45.04.01 Филология и относится к обязательные дисциплины. Осваивается на 1 курсе, 1 семестр.

Для изучения дисциплины необходимы знания, умения, компетенции, полученные обучающимися в средней общеобразовательной школе, в результате изучения дисциплин математического и естественнонаучного цикла ("Основы лингвографии", "Инновационные технологии в образовании / в науке" и др.) и дисциплин профессионального цикла ("Введение в прикладную филологию").

Данная учебная дисциплина носит обобщающий характер. Изучение данного курса позволяет систематизировать и углубить ранее полученные знания в области прикладной лингвистики.

3. Компетенции обучающегося, формируемые в результате освоения дисциплины /модуля

В результате освоения дисциплины формируются следующие компетенции:

Шифр компетенции	Расшифровка приобретаемой компетенции
ОК-1 (общекультурные компетенции)	способностью к абстрактному мышлению, анализу, синтезу (ОК-1);
ОК-2 (общекультурные компетенции)	готовностью действовать в нестандартных ситуациях, нести социальную и этическую ответственность за принятые решения (ОК-2);
ОК-4 (общекультурные компетенции)	способностью самостоятельно приобретать, в том числе с помощью информационных технологий и использовать в практической деятельности новые знания и умения, в том числе в новых областях знаний, непосредственно не связанных со сферой деятельности (ОК-4).
ПК-10 (профессиональные компетенции)	способностью к трансформации различных типов текстов (например, изменению стиля, жанра, целевой принадлежности текста) (ПК-10);
ПК-9 (профессиональные компетенции)	способностью к созданию, редактированию, реферированию и систематизированию всех типов текстов официально-делового и публицистического стиля (ПК-9);
ПК-15 (профессиональные компетенции)	умением планировать комплексное информационное воздействие и осуществлять руководство им (ПК-15);
ПК-16 (профессиональные компетенции)	владением навыками организации и проведения учебных занятий и практик, семинаров, научных дискуссий и конференций (ПК-16);

Шифр компетенции	Расшифровка приобретаемой компетенции
ПК-17 (профессиональные компетенции)	владением навыками организации и управления научно-исследовательскими и производственными работами при решении конкретных задач в соответствии с направленностью (профилем) программы магистратуры (ПК-17).

В результате освоения дисциплины студент:

1. должен знать:

основные теоретические положения корпусной лингвистики; предмет, задачи, методы понятийный аппарат изучаемой дисциплины.

2. должен уметь:

анализировать корпуса текстов; применять современные информационные технологии в исследовательской работе

3. должен владеть:

навыками лингвистического поиска в сети Интернет, ориентироваться в проблемах составления различных корпусов данных.

4. должен демонстрировать способность и готовность:

применять компьютерные технологии на языковом материале, грамотно пользоваться материалами различных корпусов текстов и т.д.

4. Структура и содержание дисциплины/ модуля

Общая трудоемкость дисциплины составляет 2 зачетных(ые) единиц(ы) 72 часа(ов).

Форма промежуточного контроля дисциплины зачет в 1 семестре.

Суммарно по дисциплине можно получить 100 баллов, из них текущая работа оценивается в 50 баллов, итоговая форма контроля - в 50 баллов. Минимальное количество для допуска к зачету 28 баллов.

86 баллов и более - "отлично" (отл.);

71-85 баллов - "хорошо" (хор.);

55-70 баллов - "удовлетворительно" (удов.);

54 балла и менее - "неудовлетворительно" (неуд.).

4.1 Структура и содержание аудиторной работы по дисциплине/ модулю

Тематический план дисциплины/модуля

N	Раздел Дисциплины/ Модуля	Семестр	Неделя семестра	Виды и часы аудиторной работы, их трудоемкость (в часах)			Текущие формы контроля
				Лекции	Практические занятия	Лабораторные работы	
1.	Тема 1. Основные понятия корпусной лингвистики. Типология корпусов текстов.	1		2	2	0	устный опрос

N	Раздел Дисциплины/ Модуля	Семестр	Неделя семестра	Виды и часы аудиторной работы, их трудоемкость (в часах)			Текущие формы контроля
				Лекции	Практические занятия	Лабораторные работы	
2.	Тема 2. Корпусы разговорной речи и аудио-архивы. Корпусы русской разговорной речи. Мультимедийные корпусы.	1		2	4	0	письменная работа
3.	Тема 3. Фундаментальные корпусы английского, французского, немецкого, китайского и др. языков.	1		0	2	0	домашнее задание
4.	Тема 4. Корпусы текстов русского языка: история и перспективы развития.	1		0	4	0	контрольная работа
	Тема . Итоговая форма контроля	1		0	0	0	зачет
	Итого			4	12	0	

4.2 Содержание дисциплины

Тема 1. Основные понятия корпусной лингвистики. Типология корпусов текстов.

лекционное занятие (2 часа(ов)):

Основные понятия корпусной лингвистики: проблемная область, корпус данных и корпус текстов, единица хранения корпуса данных / текстов.

практическое занятие (2 часа(ов)):

Типология корпусов текстов: исследовательские корпусы, иллюстративные корпусы, статистические и динамические (мониторные) корпусы, синхронические и диахронические корпусы и др. Способы представления и хранения корпуса данных.

Тема 2. Корпусы разговорной речи и аудио-архивы. Корпусы русской разговорной речи. Мультимедийные корпусы.

лекционное занятие (2 часа(ов)):

Корпусы разговорной речи и аудио-архивы. Корпусы русской разговорной речи. Мультимедийные корпусы.

практическое занятие (4 часа(ов)):

Особенности Электронного корпуса русских народных говоров

Тема 3. Фундаментальные корпусы английского, французского, немецкого, китайского и др. языков.

практическое занятие (2 часа(ов)):

История создания корпусов английского, немецкого, французского языков.

Тема 4. Корпусы текстов русского языка: история и перспективы развития.

практическое занятие (4 часа(ов)):

Корпусы текстов русского языка: история и перспективы развития. Тюбингенский корпус русского языка. Хельсинкский аннотированный корпус (ХАНКО). Корпус газетных текстов русского языка конца XX века (КГТ). Особенности НКРЯ.

4.3 Структура и содержание самостоятельной работы дисциплины (модуля)

N	Раздел Дисциплины	Семестр	Неделя семестра	Виды самостоятельной работы студентов	Трудоемкость (в часах)	Формы контроля самостоятельной работы
1.	Тема 1. Основные понятия корпусной лингвистики. Типология корпусов текстов.	1		подготовка к устному опросу	14	устный опрос
2.	Тема 2. Корпусы разговорной речи и аудио-архивы. Корпусы русской разговорной речи. Мультимедийные корпусы.	1		подготовка к письменной работе	14	письменная работа
3.	Тема 3. Фундаментальные корпусы английского, французского, немецкого, китайского и др. языков.	1		подготовка домашнего задания	14	домашнее задание
4.	Тема 4. Корпусы текстов русского языка: история и перспективы развития.	1		подготовка к контрольной работе	14	контрольная работа
	Итого				56	

5. Образовательные технологии, включая интерактивные формы обучения

Лекции, практические занятия, самостоятельная работа студентов, интернет-технологии, информационно-коммуникативный метод.

6. Оценочные средства для текущего контроля успеваемости, промежуточной аттестации по итогам освоения дисциплины и учебно-методическое обеспечение самостоятельной работы студентов

Тема 1. Основные понятия корпусной лингвистики. Типология корпусов текстов.

устный опрос , примерные вопросы:

Примерные вопросы для устного опроса 1. Корпусная лингвистика ? это ?. 2. Приведите пример обновления собранного языкового материала (без использования информационных средств). 3. Корпус данных ? это? Единица хранения ? это? Приведите пример единицы хранения определенного корпуса данных. 4. Что представляет собой звуковой тип данных в корпусе? 5. Чем исследовательский корпус отличается от иллюстративного? 6. Основные свойства корпуса текста. 7. Какие данные включает в себя морфологическая разметка? 8. Чем динамический корпус отличается от статического? 9. Какие данные включает в себя семантическая разметка?

Тема 2. Корпусы разговорной речи и аудио-архивы. Корпусы русской разговорной речи. Мультимедийные корпусы.

письменная работа , примерные вопросы:

Схема описания корпуса 1. Время создания корпуса (время появления в Интернете). 2. Общая характеристика корпуса (цели корпуса, создатели и участники проекта, объем корпуса). 3. Материалы корпуса (источники): а) что включено, б) что планируется включить. 4. Назначение корпуса (для кого предназначен). 5. Структура корпуса (какие разделы включает; краткая характеристика разделов). 6. Тип корпуса. 7. "Жизнь" корпуса (как развивается; см., например, раздел "Новости"). 8. Что создано на основе материалов данного корпуса (справочники, грамматики и т.п.).

Тема 3. Фундаментальные корпуса английского, французского, немецкого, китайского и др. языков.

домашнее задание , примерные вопросы:

6.2.2. Конспекты научных статей: 1. Гришина Е.А. Два новых проекта для национального корпуса: мультимедийный подкорпус и подкорпус названий - <http://ruscorpora.ru/sbornik2005/14grishina.pdf> 2. Гришина Е. А. Корпус ?История русского ударения? // Национальный корпус русского языка: 2006?2008. Новые результаты и перспективы. СПб.: Нестор-История, 2009, 150?174. - <http://ruscorpora.ru/sbornik2008/08.pdf> 3. Захаров В.П. Чешский национальный корпус текстов: организация и способы использования // Доклады научной конференции ?Корпусная лингвистика т лингвистические базы данных? / Под ред. А.С.Герда.- СПб., 2002.- С. 72 - 80. 4. Костыркин А. В. Корпус японской разговорной речи // Национальный корпус русского языка: 2006?2008. Новые результаты и перспективы. СПб.: Нестор-История, 2009, 474?500. - <http://ruscorpora.ru/sbornik2008/23.pdf> 5. Летучий А. Б. Корпус диалектных текстов: задачи и проблемы // Национальный корпус русского языка: 2003?2005. М.: Индрик, 2005, 215?232.<http://ruscorpora.ru/sbornik2005/13letuchy.pdf> 6. Марчук Ю.Н. Корпус текстов и сверхбольшие базы лингвистических данных // Доклады научной конференции ?Корпусная лингвистика т лингвистические базы данных? / Под ред. А.С.Герда.- СПб., 2002.- С. 102 - 108. 7. Орехов Б. В. Параллельный корпус переводов ?Слова о полку Игореве?: итоги и перспективы // Национальный корпус русского языка: 2006?2008. Новые результаты и перспективы. СПб.: Нестор-История, 2009, 462?473. <http://ruscorpora.ru/sbornik2008/22.pdf> 8. Резникова Т. И. Славянская корпусная лингвистика: современное состояние ресурсов // Национальный корпус русского языка: 2006?2008. Новые результаты и перспективы. СПб.: Нестор-История, 2009, 402?461. - <http://ruscorpora.ru/sbornik2008/21.pdf> 9. Резникова Т. И., Копотев М. В. Лингвистически аннотированные корпуса русского языка (обзор общедоступных ресурсов)// Национальный корпус русского языка: 2003?2005. М.: Индрик, 2005, 31?61.- <http://ruscorpora.ru/sbornik2005/04reznikova.pdf> 10. Скрелин П.А. Звуковые базы данных: архивы и речевые корпуса // Доклады научной конференции ?Корпусная лингвистика т лингвистические базы данных? / Под ред. А.С.Герда.- СПб., 2002.- С. 151 - 157.

Тема 4. Корпусы текстов русского языка: история и перспективы развития.

контрольная работа , примерные вопросы:

II. 1. Когда начались работы по созданию Машинного фонда русского языка? Кто был инициатором? 2. Кто был руководителем отдела Машинного фонда русского языка в Институте русского языка РАН? 3. Что включает в себя коллекция Машинного фонда? 4. Какой период охватывает подкорпус русских газет Машинного фонда русского языка? Соответствует ли название подкорпуса его содержанию? 5. Словарные продукты Машинного фонда (тип или название). 6. Каковы основные достоинства и недостатки Машинного фонда? III. 1. Где (страна), когда и для чего был создан Тюбингенский корпус? 2. Руководитель проекта. 3. Опишите подкорпуса Тюбингенского корпуса (характер входящих текстов, количество словоупотреблений, типы разметок). 4. Какие подкорпуса содержат текстовые фрагменты? 5. Как организован поиск в Тюбингенском корпусе? 6. Что такое текстоформа? 7. Каковы основные достоинства и недостатки Тюбингенского корпуса?

Тема . Итоговая форма контроля

Примерные вопросы к зачету:

Вопросы к зачету

1. Корпусная лингвистика - это ?.

2. Приведите пример обновления собранного языкового материала (без использования информационных средств).
3. Корпус данных - это?
4. Единица хранения - это?
5. Приведите пример единицы хранения определенного корпуса данных.
6. Что представляет собой звуковой тип данных в корпусе?
7. Лемма - это ?.
8. Лемматизация - это ?
9. Чем динамический корпус отличается от статического?
10. Основные свойства корпуса текста.
11. Какие данные включает в себя семантическая разметка?
12. Корпусная лингвистика и традиционная лингвистика.
13. Классификация корпусов текстов.
14. Звуковые базы данных.
15. Звуковые базы данных и аудио-архивы.
16. Звуковые базы данных русского языка.
17. Национальный корпус русского языка: особенности и перспективы развития.
18. Машинный фонд русского языка: краткая характеристика.
19. Охарактеризуйте основные корпуса текстов английского языка.
20. Охарактеризуйте основные корпуса текстов немецкого языка

7.1. Основная литература:

Корпусная лингвистика, Захаров, Виктор Павлович;Богданова, Светлана Юрьевна, 2013г.
Информационные технологии в лингвистике, Зубов, Александр Васильевич;Зубова, И.И., 2004г.

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 (<http://znanium.com/catalog.php?bookinfo=462989>)

7.2. Дополнительная литература:

Новые информационные технологии и лингвистика, Потапова, Родмонга Кондратьевна, 2012г.
Грудева, Е. В. Корпусная лингвистика [Электронный ресурс] : учеб. пособие / Е. В. Грудева. - 2-е изд., стер. - М. : ФЛИНТА, 2012. - 165 с. - ISBN 978-5-9765-1497-3 (<http://znanium.com/catalog.php?bookinfo=455049>)

7.3. Интернет-ресурсы:

сайт - <http://www.bokrcorpora.narod.ru/>

сайт - ucnk.ff.cuni.cz

сайт - <http://rykov-cl.narod.ru/c.html>

сайт - www.natcorp.ox.ac.uk

сайт - www.ruscorporu.ru

8. Материально-техническое обеспечение дисциплины(модуля)

Освоение дисциплины "Корпусные технологии в лингвистике" предполагает использование следующего материально-технического обеспечения:

Мультимедийная аудитория, вместимостью более 60 человек. Мультимедийная аудитория состоит из интегрированных инженерных систем с единой системой управления, оснащенная современными средствами воспроизведения и визуализации любой видео и аудио информации, получения и передачи электронных документов. Типовая комплектация мультимедийной аудитории состоит из: мультимедийного проектора, автоматизированного проекционного экрана, акустической системы, а также интерактивной трибуны преподавателя, включающей тач-скрин монитор с диагональю не менее 22 дюймов, персональный компьютер (с техническими характеристиками не ниже Intel Core i3-2100, DDR3 4096Mb, 500Gb), конференц-микрофон, беспроводной микрофон, блок управления оборудованием, интерфейсы подключения: USB, audio, HDMI. Интерактивная трибуна преподавателя является ключевым элементом управления, объединяющим все устройства в единую систему, и служит полноценным рабочим местом преподавателя. Преподаватель имеет возможность легко управлять всей системой, не отходя от трибуны, что позволяет проводить лекции, практические занятия, презентации, вебинары, конференции и другие виды аудиторной нагрузки обучающихся в удобной и доступной для них форме с применением современных интерактивных средств обучения, в том числе с использованием в процессе обучения всех корпоративных ресурсов. Мультимедийная аудитория также оснащена широкополосным доступом в сеть интернет. Компьютерное оборудование имеет соответствующее лицензионное программное обеспечение.

Учебно-методическая литература для данной дисциплины имеется в наличии в электронно-библиотечной системе "БиблиоРоссика", доступ к которой предоставлен студентам. В ЭБС "БиблиоРоссика" представлены коллекции актуальной научной и учебной литературы по гуманитарным наукам, включающие в себя публикации ведущих российских издательств гуманитарной литературы, издания на английском языке ведущих американских и европейских издательств, а также редкие и малотиражные издания российских региональных вузов. ЭБС "БиблиоРоссика" обеспечивает широкий законный доступ к необходимым для образовательного процесса изданиям с использованием инновационных технологий и соответствует всем требованиям федеральных государственных образовательных стандартов высшего профессионального образования (ФГОС ВПО) нового поколения.

Учебно-методическая литература для данной дисциплины имеется в наличии в электронно-библиотечной системе "ZNANIUM.COM", доступ к которой предоставлен студентам. ЭБС "ZNANIUM.COM" содержит произведения крупнейших российских учёных, руководителей государственных органов, преподавателей ведущих вузов страны, высококвалифицированных специалистов в различных сферах бизнеса. Фонд библиотеки сформирован с учетом всех изменений образовательных стандартов и включает учебники, учебные пособия, УМК, монографии, авторефераты, диссертации, энциклопедии, словари и справочники, законодательно-нормативные документы, специальные периодические издания и издания, выпускаемые издательствами вузов. В настоящее время ЭБС ZNANIUM.COM соответствует всем требованиям федеральных государственных образовательных стандартов высшего профессионального образования (ФГОС ВПО) нового поколения.

Материально-техническое обеспечение дисциплины. Учебные аудитории, оборудованные учебной мебелью, компьютером, мультимедийным проектором для проведения практических занятий.

Программа составлена в соответствии с требованиями ФГОС ВПО и учебным планом по направлению 45.04.01 "Филология" и магистерской программе Прикладное языкознание .

Автор(ы):

Каримуллина Г.Н. _____

"__" _____ 201__ г.

Рецензент(ы):

"__" _____ 201__ г.