

МИНИСТЕРСТВО НАУКИ И ВЫСШЕГО ОБРАЗОВАНИЯ РОССИЙСКОЙ ФЕДЕРАЦИИ
Федеральное государственное автономное образовательное учреждение высшего образования
"Казанский (Приволжский) федеральный университет"
Институт международных отношений
Отделение Высшая школа иностранных языков и перевода



УТВЕРЖДАЮ

Проректор по образовательной деятельности КФУ
проф. Таюрский Д.А.

"__" _____ 20__ г.

Программа дисциплины

Введение в корпусную лингвистику

Направление подготовки: 45.04.02 - Лингвистика

Профиль подготовки: Теория перевода, межкультурная и межъязыковая коммуникация

Квалификация выпускника: магистр

Форма обучения: очное

Язык обучения: русский

Год начала обучения по образовательной программе: 2018

Содержание

1. Перечень планируемых результатов обучения по дисциплине (модулю), соотнесенных с планируемыми результатами освоения ОПОП ВО
2. Место дисциплины (модуля) в структуре ОПОП ВО
3. Объем дисциплины (модуля) в зачетных единицах с указанием количества часов, выделенных на контактную работу обучающихся с преподавателем (по видам учебных занятий) и на самостоятельную работу обучающихся
4. Содержание дисциплины (модуля), структурированное по темам (разделам) с указанием отведенного на них количества академических часов и видов учебных занятий
 - 4.1. Структура и тематический план контактной и самостоятельной работы по дисциплине (модулю)
 - 4.2. Содержание дисциплины (модуля)
5. Перечень учебно-методического обеспечения для самостоятельной работы обучающихся по дисциплине (модулю)
6. Фонд оценочных средств по дисциплине (модулю)
7. Перечень литературы, необходимой для освоения дисциплины (модуля)
8. Перечень ресурсов информационно-телекоммуникационной сети "Интернет", необходимых для освоения дисциплины (модуля)
9. Методические указания для обучающихся по освоению дисциплины (модуля)
10. Перечень информационных технологий, используемых при осуществлении образовательного процесса по дисциплине (модулю), включая перечень программного обеспечения и информационных справочных систем (при необходимости)
11. Описание материально-технической базы, необходимой для осуществления образовательного процесса по дисциплине (модулю)
12. Средства адаптации преподавания дисциплины (модуля) к потребностям обучающихся инвалидов и лиц с ограниченными возможностями здоровья
13. Приложение №1. Фонд оценочных средств
14. Приложение №2. Перечень литературы, необходимой для освоения дисциплины (модуля)
15. Приложение №3. Перечень информационных технологий, используемых для освоения дисциплины (модуля), включая перечень программного обеспечения и информационных справочных систем

Программу дисциплины разработал(а)(и) доцент, к.н. Рахимбирдиева И.М. (Кафедра теории и практики перевода, Высшая школа иностранных языков и перевода), IMRahimbirdieva@kpfu.ru

1. Перечень планируемых результатов обучения по дисциплине (модулю), соотнесенных с планируемыми результатами освоения ОПОП ВО

Обучающийся, освоивший дисциплину (модуль), должен обладать следующими компетенциями:

Шифр компетенции	Расшифровка приобретаемой компетенции
ПК-16	владением методикой предпереводческого анализа текста, способствующей точному восприятию исходного высказывания, подготовки к выполнению перевода, включая поиск информации в справочной, специальной литературе и компьютерных сетях
ПК-35	владением методиками экспертной оценки программных продуктов лингвистического профиля
ПК-37	владением основами современной информационной и библиографической культуры

Обучающийся, освоивший дисциплину (модуль):

Должен знать:

- определение понятия 'корпусная лингвистика';
- основные преимущества практического применения достижений корпусной лингвистики;
- основные корпусы русского и английского языков.

Должен уметь:

- пользоваться существующими корпусами русского и английского языков;
- использовать корпуса в повседневной работе.

Должен владеть:

- методикой предпереводческого анализа текста, способствующей точному восприятию исходного текста.

Должен демонстрировать способность и готовность:

- применять полученные знания в области теории и истории основного изучаемого языка (языков) и литературы (литератур), теории коммуникации, филологического анализа и интерпретации текста.

2. Место дисциплины (модуля) в структуре ОПОП ВО

Данная дисциплина (модуль) включена в раздел "Б1.В.ДВ.7 Дисциплины (модули)" основной профессиональной образовательной программы 45.04.02 "Лингвистика (Теория перевода, межкультурная и межъязыковая коммуникация)" и относится к дисциплинам по выбору.

Осваивается на 1, 2 курсах в 2, 3 семестрах.

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

Общая трудоемкость дисциплины составляет 3 зачетных(ые) единиц(ы) на 108 часа(ов).

Контактная работа - 34 часа(ов), в том числе лекции - 0 часа(ов), практические занятия - 34 часа(ов), лабораторные работы - 0 часа(ов), контроль самостоятельной работы - 0 часа(ов).

Самостоятельная работа - 38 часа(ов).

Контроль (зачёт / экзамен) - 36 часа(ов).

Форма промежуточного контроля дисциплины: зачет во 2 семестре; экзамен в 3 семестре.

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

4.1 Структура и тематический план контактной и самостоятельной работы по дисциплине (модулю)

N	Разделы дисциплины / модуля	Семестр	Виды и часы контактной работы, их трудоемкость (в часах)			Самостоятельная работа
			Лекции	Практические занятия	Лабораторные работы	
N	Разделы дисциплины / модуля	Семестр	Виды и часы контактной работы, их трудоемкость (в часах)			Самостоятельная работа
			Лекции	Практические занятия	Лабораторные работы	
1.	Тема 1. Интернет и развитие гуманитарного знания	2	0	2	0	2
2.	Тема 2. Определение понятия "корпус". Его основные характеристики.	2	0	4	0	4
3.	Тема 3. Возникновение корпусной лингвистики	2	0	4	0	4
4.	Тема 4. Виды корпусов. Их характерные особенности	2	0	4	0	4
5.	Тема 5. Поиск информации в корпусе	2	0	2	0	2
6.	Тема 6. Разметка корпусов	2	0	2	0	2
7.	Тема 7. Исследование языка с помощью лингвистического корпуса: общие принципы	3	0	4	0	4
8.	Тема 8. Исследование языка с помощью лингвистического корпуса: направления исследования и практическое применение результатов	3	0	4	0	6
9.	Тема 9. Корпусное преподавание	3	0	4	0	6
4.2 Содержание дисциплины (модуля)						
10.	Тема 10. Корпус и междисциплинарные исследования лингвистики. Соотношение корпусной лингвистики и компьютерной лингвистики. Становление корпусной лингвистики как раздела компьютерной лингвистики. Ее цели, задачи, предмет, объект исследования. Основные направления корпусной лингвистики: лексикографические исследования, квантитативная лингвистика, машинный перевод, лингводидактика, лингвистика текста и др.	3	0	4	0	4

Тема 1. Интернет и развитие гуманитарного знания

Цели и основные направления корпусной лингвистики. Соотношение корпусной лингвистики и компьютерной лингвистики. Становление корпусной лингвистики как раздела компьютерной лингвистики. Ее цели, задачи, предмет, объект исследования. Основные направления корпусной лингвистики: лексикографические исследования, квантитативная лингвистика, машинный перевод, лингводидактика, лингвистика текста и др.

Тема 2. Определение понятия "корпус". Его основные характеристики.

Предмет исследования. Общая характеристика лингвистических корпусов. Машинные фонды, базы данных, корпус. Основные характеристики корпусов. Понятие репрезентативности корпусов. Понятие сбалансированности корпусов. Машинные фонды, базы данных, корпус. Общие черты и отличия. Предпосылки создания корпусов на основе машинных фондов, баз данных. Машинный фонд vs Корпус. База данных vs корпус.

Тема 3. Возникновение корпусной лингвистики

Понятие лингвистического корпуса. Определение лингвистического (языкового) корпуса. Лингвистические направления, предвосхитившие появление корпусной лингвистики: историческая лингвистика, написание грамматик, лексикография и обучение языку, социолингвистика, формализм в языке. Сопоставление корпусной и традиционной лингвистики. Корпусный подход и его специфика.

Тема 4. Виды корпусов. Их характерные особенности

Виды корпусов: устные и письменные, одноязычные и многоязычные. Виды корпусов: исследовательский, иллюстративный, статический, динамический, устный, письменный, одноязычный, многоязычный и др. Специфика каждого корпуса. Демонстрация разных типов корпусов. Создание формализованных описаний для разработки корпуса. Словарный и бессловарный способ. Языковой корпус как компьютерная база данных. Процессы токенизации, лемматизации, стемминга и парсинга: специфика, цели, результаты. Корпусные менеджеры.

Тема 5. Поиск информации в корпусе

Методы извлечения информации из корпуса. Языки запросов. Выходные интерфейсы. Типы извлекаемой информации. Автоматизированный поиск информации по жанровым разновидностям текстов, представленных в корпусе языковых данных. Конкорданс, коллокация, кластеры, чанки, коллигация. Аннотирование корпуса. Хранение, публикация и обновление корпусов. Единица хранения корпуса данных. Формат представления информации. Он-лайн корпуса. Офф-лайн корпуса. Программы аннотирования корпусов. Проблема обновления корпусов.

Тема 6. Разметка корпусов

Разметка. Виды разметок. Лингвистическая аннотация и метаданные. Понятие разметки. Программы автоматической разметки корпусов. Лингвистическая разметка: морфологическая, синтаксическая, семантическая, анафорическая, просодическая, дискурсная. Понятие метаданных. Экстралингвистическая разметка: библиографическая, типологическая, тематическая, социологическая, формальная структурная, технико-технологическая. Глубина разметки.

Тема 7. Исследование языка с помощью лингвистического корпуса: общие принципы

Лингвистические исследования на базе корпуса: изучение лексики. Частотный словарь. Структурно-вероятностная модель текста. Функции частотного словаря. Единица описания. Понятие токена, леммы, словоформы. Принципы составления частотных словарей. Лингвистические генеральная и выборочная совокупность. Словник. Абсолютная частота. Ранг. I_{pm} . Коэффициент D . Коэффициент R . Покрываемость текста. Лексическое разнообразие текста. Исследование нормы и узуса. Примеры частотных словарей.

Тема 8. Исследование языка с помощью лингвистического корпуса: направления исследования и практическое применение результатов

Лингвистические исследования на базе корпуса: составление словарей. Краткая характеристика основ идеографического описания языков. Понятие "семантической сети" как одного из способов отражения семантики предметной области как промежуточной формы представления текстовых данных. Классификации семантических сетей и их критика. Типы отношений в семантических сетях. Тезаурус как особая разновидность словарей. Перечисление значений слов в соответствии с их частотным распределением в корпусе текстов как один из базовых принципов построения wordnet-тезауруса.

Тема 9. Корпусное преподавание

Основы корпусного преподавания. Возможности использования корпусной лингвистики в преподавании: эффективность работы с обучающимися и с текстами; практические вопросы в умении составлять упражнения и задания, дидактизировать изучаемые тексты с использованием корпусных данных. Подготовка учебно-методических материалов для проведения занятий и внеклассных мероприятий на основе существующих методик.

Тема 10. Корпус и междисциплинарные исследования

лингвистические исследования на базе корпуса: изучение грамматических особенностей языка. Исследование нормы и узуса. Исследование морфологии и синтаксиса. Граммема. Частотность граммем. Частотность частей речи. Частотность синтаксических молекул. Структурно-вероятностная модель грамматики.

5. Перечень учебно-методического обеспечения для самостоятельной работы обучающихся по дисциплине (модулю)

Самостоятельная работа обучающихся выполняется по заданию и при методическом руководстве преподавателя, но без его непосредственного участия. Самостоятельная работа подразделяется на самостоятельную работу на аудиторных занятиях и на внеаудиторную самостоятельную работу. Самостоятельная работа обучающихся включает как полностью самостоятельное освоение отдельных тем (разделов) дисциплины, так и проработку тем (разделов), осваиваемых во время аудиторной работы. Во время самостоятельной работы обучающиеся читают и конспектируют учебную, научную и справочную литературу, выполняют задания, направленные на закрепление знаний и отработку умений и навыков, готовятся к текущему и промежуточному контролю по дисциплине.

Организация самостоятельной работы обучающихся регламентируется нормативными документами, учебно-методической литературой и электронными образовательными ресурсами, включая:

Порядок организации и осуществления образовательной деятельности по образовательным программам высшего образования - программам бакалавриата, программам специалитета, программам магистратуры (утвержден приказом Министерства образования и науки Российской Федерации от 5 апреля 2017 года №301)

Письмо Министерства образования Российской Федерации №14-55-996ин/15 от 27 ноября 2002 г. "Об активизации самостоятельной работы студентов высших учебных заведений"

Устав федерального государственного автономного образовательного учреждения "Казанский (Приволжский) федеральный университет"

Правила внутреннего распорядка федерального государственного автономного образовательного учреждения высшего профессионального образования "Казанский (Приволжский) федеральный университет"

Локальные нормативные акты Казанского (Приволжского) федерального университета

6. Фонд оценочных средств по дисциплине (модулю)

Фонд оценочных средств по дисциплине (модулю) включает оценочные материалы, направленные на проверку освоения компетенций, в том числе знаний, умений и навыков. Фонд оценочных средств включает оценочные средства текущего контроля и оценочные средства промежуточной аттестации.

В фонде оценочных средств содержится следующая информация:

- соответствие компетенций планируемым результатам обучения по дисциплине (модулю);
- критерии оценивания сформированности компетенций;
- механизм формирования оценки по дисциплине (модулю);
- описание порядка применения и процедуры оценивания для каждого оценочного средства;
- критерии оценивания для каждого оценочного средства;
- содержание оценочных средств, включая требования, предъявляемые к действиям обучающихся, демонстрируемым результатам, задания различных типов.

Фонд оценочных средств по дисциплине находится в Приложении 1 к программе дисциплины (модулю).

7. Перечень литературы, необходимой для освоения дисциплины (модуля)

Освоение дисциплины (модуля) предполагает изучение основной и дополнительной учебной литературы. Литература может быть доступна обучающимся в одном из двух вариантов (либо в обоих из них):

- в электронном виде - через электронные библиотечные системы на основании заключенных КФУ договоров с правообладателями;
- в печатном виде - в Научной библиотеке им. Н.И. Лобачевского. Обучающиеся получают учебную литературу на абонементе по читательским билетам в соответствии с правилами пользования Научной библиотекой.

Электронные издания доступны дистанционно из любой точки при введении обучающимся своего логина и пароля от личного кабинета в системе "Электронный университет". При использовании печатных изданий библиотечный фонд должен быть укомплектован ими из расчета не менее 0,5 экземпляра (для обучающихся по ФГОС 3++ - не менее 0,25 экземпляра) каждого из изданий основной литературы и не менее 0,25 экземпляра дополнительной литературы на каждого обучающегося из числа лиц, одновременно осваивающих данную дисциплину.

Перечень основной и дополнительной учебной литературы, необходимой для освоения дисциплины (модуля), находится в Приложении 2 к рабочей программе дисциплины. Он подлежит обновлению при изменении условий договоров КФУ с правообладателями электронных изданий и при изменении комплектования фондов Научной библиотеки КФУ.

8. Перечень ресурсов информационно-телекоммуникационной сети "Интернет", необходимых для освоения дисциплины (модуля)

Британский национальный корпус - www.corpus.byu.edu?bnc/

Национальный корпус русского языка - www.ruscorpora.ru

Открытое образование - <https://openedu.ru/course/hse/CORPUS/>

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

Вид работ	Методические рекомендации
практические занятия	Для обеспечения систематической и регулярной работы по изучению дисциплины и успешного прохождения промежуточных и итоговых контрольных испытаний студенту рекомендуется придерживаться следующего порядка обучения: посещать все практические занятия; вовремя выполнять письменные домашние задания; самостоятельно определить объем времени, который необходим для проработки каждой темы. При подготовке к практическим занятиям студентам необходимо изучить материал по теме занятия из источников основной и дополнительной литературы, выполнить все упражнения и задания для практических занятий.
самостоятельная работа	Самостоятельная работа проводится в часы внеаудиторной работы. Обучающиеся получают задания для проверки усвоения пройденного материала. Работа выполняется в письменном виде и сдаётся преподавателю. Оцениваются владение материалом по теме работы, аналитические способности, владение методами, умения и навыки, необходимые для выполнения заданий.

Вид работ	Методические рекомендации
зачет	Для успешной сдачи зачета по дисциплине необходимо проработать конспекты занятий по всем пройденным темам курса "корпусная лингвистика". Рекомендуется чтение основной и дополнительной литературы по темам, соответствующим вопросам, вынесенным на зачет. При выставлении зачета преподаватель учитывает также работы, выполненные студентом в течение семестра самостоятельно или в качестве эссе или контрольной работы.
экзамен	Для успешной сдачи экзамена по дисциплине необходимо проработать конспекты занятий по всем пройденным темам курса "корпусная лингвистика". Необходимо изучить основную и дополнительную литературу по темам, соответствующим вопросам, вынесенным на экзамен. При выставлении экзамена преподаватель учитывает также работы, выполненные студентом в течение семестра самостоятельно или в качестве эссе или контрольной работы.

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

Перечень информационных технологий, используемых при осуществлении образовательного процесса по дисциплине (модулю), включая перечень программного обеспечения и информационных справочных систем, представлен в Приложении 3 к рабочей программе дисциплины (модуля).

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

Материально-техническое обеспечение образовательного процесса по дисциплине (модулю) включает в себя следующие компоненты:

Помещения для самостоятельной работы обучающихся, укомплектованные специализированной мебелью (столы и стулья) и оснащенные компьютерной техникой с возможностью подключения к сети "Интернет" и обеспечением доступа в электронную информационно-образовательную среду КФУ.

Учебные аудитории для контактной работы с преподавателем, укомплектованные специализированной мебелью (столы и стулья).

Компьютер и принтер для распечатки раздаточных материалов.

Мультимедийная аудитория.

Компьютерный класс.

12. Средства адаптации преподавания дисциплины к потребностям обучающихся инвалидов и лиц с ограниченными возможностями здоровья

При необходимости в образовательном процессе применяются следующие методы и технологии, облегчающие восприятие информации обучающимися инвалидами и лицами с ограниченными возможностями здоровья:

- создание текстовой версии любого нетекстового контента для его возможного преобразования в альтернативные формы, удобные для различных пользователей;
- создание контента, который можно представить в различных видах без потери данных или структуры, предусмотреть возможность масштабирования текста и изображений без потери качества, предусмотреть доступность управления контентом с клавиатуры;
- создание возможностей для обучающихся воспринимать одну и ту же информацию из разных источников - например, так, чтобы лица с нарушениями слуха получали информацию визуально, с нарушениями зрения - аудиально;
- применение программных средств, обеспечивающих возможность освоения навыков и умений, формируемых дисциплиной, за счёт альтернативных способов, в том числе виртуальных лабораторий и симуляционных технологий;
- применение дистанционных образовательных технологий для передачи информации, организации различных форм интерактивной контактной работы обучающегося с преподавателем, в том числе вебинаров, которые могут быть использованы для проведения виртуальных лекций с возможностью взаимодействия всех участников дистанционного обучения, проведения семинаров, выступления с докладами и защиты выполненных работ, проведения тренингов, организации коллективной работы;
- применение дистанционных образовательных технологий для организации форм текущего и промежуточного контроля;
- увеличение продолжительности сдачи обучающимся инвалидом или лицом с ограниченными возможностями здоровья форм промежуточной аттестации по отношению к установленной продолжительности их сдачи:
- продолжительности сдачи зачёта или экзамена, проводимого в письменной форме, - не более чем на 90 минут;
- продолжительности подготовки обучающегося к ответу на зачёте или экзамене, проводимом в устной форме, - не более чем на 20 минут;

- продолжительности выступления обучающегося при защите курсовой работы - не более чем на 15 минут.

Программа составлена в соответствии с требованиями ФГОС ВО и учебным планом по направлению 45.04.02 "Лингвистика" и магистерской программе "Теория перевода, межкультурная и межъязыковая коммуникация".

Перечень литературы, необходимой для освоения дисциплины (модуля)

Направление подготовки: 45.04.02 - Лингвистика

Профиль подготовки: Теория перевода, межкультурная и межъязыковая коммуникация

Квалификация выпускника: магистр

Форма обучения: очное

Язык обучения: русский

Год начала обучения по образовательной программе: 2018

Основная литература:

1. Волосатова Т.М. 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/bookread2.php?book=508097>

Дополнительная литература:

1. Чурилина, Л. Н. Актуальные проблемы современной лингвистики [электронный ресурс] : учеб. пособие / сост. Л. Н. Чурилина. ? 6-е изд., стереотип. ? М. : Флинта : Наука, 2011. ? 416 с. - ISBN 978-5-89349-892-9 (Флинта), ISBN 978-5-02-033353-6 (Наука) -

Режим доступа: <http://znanium.com/bookread2.php?book=454683>

2. 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/bookread2.php?book=462989>

3. Актуальные проблемы современной лингвистики [Электронный ресурс] : сборник научных статей. - СПб : Издательство СПбГЭУ, 2014. - 137 с. - Режим доступа: <http://znanium.com/bookread2.php?book=527716>

4. Алпатов В. Языкознание: От Аристотеля до компьютерной лингвистики: Научно-популярное / Алпатов В. - М.: Альпина нон-фикшн, 2018. - 253 с.: 60x90 1/16. - (Научно-популярная литература) (Переплёт) ISBN 978-5-91671-804-1 -

Режим доступа: <http://znanium.com/bookread2.php?book=1003471>

Приложение 3
к рабочей программе дисциплины (модуля)
Б1.В.ДВ.7 Введение в корпусную лингвистику

Перечень информационных технологий, используемых для освоения дисциплины (модуля), включая перечень программного обеспечения и информационных справочных систем

Направление подготовки: 45.04.02 - Лингвистика

Профиль подготовки: Теория перевода, межкультурная и межъязыковая коммуникация

Квалификация выпускника: магистр

Форма обучения: очное

Язык обучения: русский

Год начала обучения по образовательной программе: 2018

Освоение дисциплины (модуля) предполагает использование следующего программного обеспечения и информационно-справочных систем:

Операционная система Microsoft Windows 7 Профессиональная или Windows XP (Volume License)

Пакет офисного программного обеспечения Microsoft Office 365 или Microsoft Office Professional plus 2010

Браузер Mozilla Firefox

Браузер Google Chrome

Adobe Reader XI или Adobe Acrobat Reader DC

Kaspersky Endpoint Security для Windows

Учебно-методическая литература для данной дисциплины имеется в наличии в электронно-библиотечной системе "ZNANIUM.COM", доступ к которой предоставлен обучающимся. 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, учебно-методические комплексы, 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 образования (ФГОС ВО) нового поколения.