

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



УТВЕРЖДАЮ

Проректор по образовательной деятельности КФУ

Проф. Таюрский Д.А.

\_\_\_\_\_ 20\_\_ г.

*подписано электронно-цифровой подписью*

### Программа дисциплины

Корпусное программное обеспечение в лингвистике Б1.В.ДВ.4

Направление подготовки: 45.04.01 - Филология

Профиль подготовки: Деловой английский язык

Квалификация выпускника: магистр

Форма обучения: заочное

Язык обучения: русский

**Автор(ы):**

Каюмова Д.Ф.

**Рецензент(ы):**

Давлетбаева Д.Н.

### **СОГЛАСОВАНО:**

Заведующий(ая) кафедрой: Давлетбаева Д. Н.

Протокол заседания кафедры No \_\_\_\_ от " \_\_\_\_ " \_\_\_\_\_ 201\_\_ г

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

Протокол заседания УМК No \_\_\_\_ от " \_\_\_\_ " \_\_\_\_\_ 201\_\_ г

Регистрационный No 902270417

Казань  
2017

## 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) профессор, д.н. (профессор) Каюмова Д.Ф. кафедра германской филологии Высшая школа русской и зарубежной филологии ,  
Diana.Kajumova@kpfu.ru

### 1. Цели освоения дисциплины

Цели освоения дисциплины: предлагаемый курс углубляет знания магистрантов об основных направлениях корпусной лингвистики и лексикографии.

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

Данная учебная дисциплина включена в раздел " Б1.В.ДВ.4 Дисциплины (модули)" основной образовательной программы 45.04.01 Филология и относится к дисциплинам по выбору. Осваивается на 1, 2 курсах, 2, 3 семестры.

Дисциплина является дисциплиной по выбору ОП (М1.В.ДВ.1). Дисциплина относится к вариативной части профессионального цикла дисциплин ОП (М1.В.ДВ.1).

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

В результате освоения дисциплины формируются следующие компетенции:

Шифр компетенции	Расшифровка приобретаемой компетенции
(ОК-7) (общекультурные компетенции)	владеет культурой мышления, способен к анализу, обобщению информации, постановке целей и выбору путей их достижения, владеет культурой устной и письменной речи
(ПК-21) (профессиональные компетенции)	умеет работать с основными информационно-поисковыми и экспертными системами, системами представления знаний, синтаксического и морфологического анализа, автоматического синтеза, распознавания и понимания речи, обработки лексикографической информации и автоматизированного перевода, автоматизированными системами идентификации и верификации личности
(ПК-22) (профессиональные компетенции)	владеет методами когнитивного и формального моделирования естественного языка и методами создания метаязыков
(ПК-23) (профессиональные компетенции)	владеет современными методиками сбора, хранения и представления баз данных и знаний в интеллектуальных системах различного назначения с учетом достижений корпусной лингвистики
(СК-1)	умеет работать с основными информационно-поисковыми и экспертными системами, создавать лингвистические информационные ресурсы

В результате освоения дисциплины студент:

4. должен демонстрировать способность и готовность:

В результате изучения дисциплины студент - магистр должен:

знать

- основные типы корпусов (23.1);

- сущность процедуры разметки и основные стандарты корпусной разметки.

уметь

- давать определения базовым терминам;
- выражать и обосновывать свою точку зрения на проблемные вопросы изучаемого курса;
- пользоваться поисковыми средствами;
- осуществлять поиск и исследования на базе корпусов;
- делать самостоятельные обобщения и выводы из наблюдений над конкретным языковым материалом;

владеть

- навыками анализа и сопоставления, оценки информации из предложенных источников;
- практическими навыками лингвистического анализа конкретного языкового материала.
- основами корпусных технологий;
- навыками работы с программными средствами и информационными ресурсами корпусной лингвистики;
- методикой работы с корпусными и лексикографическими базами данных в целях лингвистического анализа;
- понятийной базой и терминологическим минимумом корпусной лингвистики

#### 4. Структура и содержание дисциплины/ модуля

Общая трудоемкость дисциплины составляет зачетных(ые) единиц(ы) 144 часа(ов).

Форма промежуточного контроля дисциплины отсутствует во 2 семестре; экзамен в 3 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. Введение в корпусную лингвистику.	2		2	3	0	
2.	Тема 2. Типология корпусов. Корпусные и лексикографические исследования	2		0	3	0	
3.	Тема 3. Элементы прикладной лингвистики	3		3	2	0	

N	Раздел Дисциплины/ Модуля	Семестр	Неделя семестра	Виды и часы аудиторной работы, их трудоемкость (в часах)			Текущие формы контроля
				Лекции	Практические занятия	Лабораторные работы	
4.	Тема 4. Вероятностно-статистическое изучение языка и речи	3		3	2	0	
	Тема . Итоговая форма контроля	3		0	0	0	Экзамен
	Итого			8	10	0	

#### 4.2 Содержание дисциплины

##### Тема 1. Введение в корпусную лингвистику.

###### **лекционное занятие (2 часа(ов)):**

История лингвистических корпусов: от картотеки к корпусу. Корпусная лингвистика: современное состояние. Направления в лингвистике, предвосхитившие появление корпусной лингвистики: от картотеки к корпусу

###### **практическое занятие (3 часа(ов)):**

Корпусная лингвистика в России. Британский национальный корпус.

##### Тема 2. Типология корпусов. Корпусные и лексикографические исследования

###### **практическое занятие (3 часа(ов)):**

Классификация (типология корпусов), основания типологии. Корпус как поисковая система. Структура и характеристики корпусов. Методология компиляции корпусов. Понятие разметки.

##### Тема 3. Элементы прикладной лингвистики

###### **лекционное занятие (3 часа(ов)):**

Лингвистическая разметка. Экстралингвистическая разметка. Лексикографические исследования, основанные на корпусах.

###### **практическое занятие (2 часа(ов)):**

Основные области приложения структурно-вероятностной модели языка: лингвистический мониторинг функционирования языка; компьютерное моделирование языка и речи; дешифровка кодированного текста; авторизация/атрибуция текста. Прикладная лингвистика. Сравнение методов описательной, теоретической и прикладной лингвистики. Некоторые методы современной лингвистики, заимствованные у смежных наук.

##### Тема 4. Вероятностно-статистическое изучение языка и речи

###### **лекционное занятие (3 часа(ов)):**

Теоретическая и прикладная лингвистика. Методы классификации и моделирования в описательной, теоретической и прикладной лингвистике. Типы лингвистических моделей: а) компонентные, предсказывающие, имитирующие и диахронические; б) анализирующие и синтезирующие.

###### **практическое занятие (2 часа(ов)):**

Статистический подход к исследованию языковых структур. Основания и условия вероятностно-статистического изучения языка и речи. Этапы статистического эксперимента. Корпусная лингвистика.

#### 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. Введение в					

корпусную лингвистику.

устному опросу

N	Раздел Дисциплины	Семестр	Неделя семестра	Виды самостоятельной работы студентов	Трудоемкость (в часах)	Формы контроля самостоятельной работы
2.	Тема 2. Типология корпусов. Корпусные и лексикографические исследования	2		Подготовка к устному опросу	32	Устный опрос
3.	Тема 3. Элементы прикладной лингвистики	3		Подготовка к устному опросу	25	Устный опрос
4.	Тема 4. Вероятностно-статистическое изучение языка и речи	несвое		Подготовка к контрольной работе	28	Контрольная работа
	Итого				117	

## 5. Образовательные технологии, включая интерактивные формы обучения

Программы пакета Microsoft Office (Word, Power Point, Exell и др.), АСТ (тесты).

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

### Тема 1. Введение в корпусную лингвистику.

Устный опрос, примерные вопросы:

История лингвистических корпусов: от картотеки к корпусу.

### Тема 2. Типология корпусов. Корпусные и лексикографические исследования

Устный опрос, примерные вопросы:

Классификация (типология корпусов), основания типологии.

### Тема 3. Элементы прикладной лингвистики

Устный опрос, примерные вопросы:

Экстралингвистическая разметка. Лексикографические исследования, основанные на корпусах.

### Тема 4. Вероятностно-статистическое изучение языка и речи

Контрольная работа, примерные вопросы:

Статистический подход к исследованию языковых структур. Основания и условия вероятностно-статистического изучения языка и речи.

### Тема . Итоговая форма контроля

Примерные вопросы к экзамену:

По окончании дисциплины предусмотрено проведение зачета в форме защиты контрольных работ.

Темы контрольных работ:

1) Электронное тестирование по теоретической части курса с использованием сервера тестирования ТюмГУ "Ингрис" (<http://test.utmn.ru>). Пример тестового задания:

- Каким может быть количество языков в выравненном корпусе (aligned corpus)?

□ 1. Любое число больше одного.



- 2. Два языка.
  - 3. Не больше четырёх.
  - 2) Разметка собственного корпуса.
  - 3) Корпусное исследование (одноязычный корпус не менее 10 тысяч словоупотреблений), включающее следующие обязательных модули:
    - Обоснование корпуса, описание источников материала.
    - Выявление базовых параметров корпуса (соотношение types/tokens, индекс исключительности, индекс постоянства, относительная частота и коллокации слов, составление конкорданса для одного-двух частых слов) при помощи как минимум двух разных программ.
    - Краткое сравнение функциональности и удобства использованных программ.
    - Описание какой-либо лингвистической особенности корпуса, которая вскрылась в процессе выполнения проекта.
- Корпусное исследование сдаётся преподавателю в электронном виде.

### 7.1. Основная литература:

1. Баранов А.Н. Корпусная лингвистика // Баранов А.Н. Введение в прикладную лингвистику. - М., 2007
2. Гальперин И.Р. Текст как объект лингвистического исследования. - М.: Едиториал УРСС, 2005
3. Гвишиани Н.Б. Практикум по корпусной лингвистике: Учеб. пособие по английскому языку. - М.: Издательство Высшая школа, 2008.
4. Захаров В.П. Корпусная лингвистика. Учебно-метод. пособие. - СПб.: Издательство СПбГУ, 2006.
5. Захаров В.П. Информационно-поисковые системы: Учебно-метод. пособие. - СПб.: Издательство СПбГУ, 2005.
6. Марчук Ю.Н. Основы компьютерной лингвистики. - М.: Изд-во МГУ, 2000.
7. Kennedy G. An Introduction to Corpus Linguistics. - London: Longman, 1998.
8. Tony McEnery, Andrew Wilson. Corpus Linguistics. Edinburgh University Press, 2001.

### 7.2. Дополнительная литература:

1. Белозёрова Н.Н., Чуфистова Л.Е. Шекспир и компания, или использование электронных библиотек при лингвистическом исследовании: Учебное пособие. Тюмень: Издательство ТюмГУ, 2007.
2. Зубов А.В. Информационные технологии в лингвистике: Учебное пособие. - М.: Издательский центр "Академия", 2004.
3. Коваль С. А. Лингвистические проблемы компьютерной морфологии. - СПб.: Изд-во С.-Петербург. ун-та, 2005.
4. Мильруд Р.П. Введение в лингвистику. Introduction to Linguistics: Учебное пособие. - М.: Дрофа, 2005.
5. Рахилина Е. В. Корпус как творческий проект // Национальный корпус русского языка: 2006-2008. Новые результаты и перспективы. - СПб.: Издательство Нестор- История, 2009.
6. Corpus Linguistics Beyond the Word: Corpus Research from Phrase to Discourse / Ed. by E. Fitzpatrick. - Amsterdam, New York: Rodopi, 2007.
7. Corpus Linguistics: An International Handbook / Ed. by A. Lüdeling, M. Kytö. - Vols. 1-2.- Berlin et al: Walter de Gruyter, 2008.
8. Sinclair J.M. Corpus, Concordance, Collocation. - Oxford: Oxford University Press, 1991.

### 7.3. Интернет-ресурсы:

British National Corpus. American National Corpus. - <http://corpus.byu.edu>

ICAME (International Computer Archive of Modern and Medieval English) Journal.[ - <http://icame.uib.no/journal.html>

The DGT Multilingual Translation Memory of the Acquis Communautaire - <http://langtech.jrc.it/DGT-TM.html>

Tony McEnery, Andrew Wilson. Corpus Linguistics. - [http://books.google.com/books?id=nwmgdvN\\_akAC](http://books.google.com/books?id=nwmgdvN_akAC)

Коваль С.А. Роль корпуса в создании реалистичных моделей словоизменительной морфологии. [электронный ресурс] - [http://skowal.narod.ru/research/corpora2006/Koval\\_Corpora.2006.htm](http://skowal.narod.ru/research/corpora2006/Koval_Corpora.2006.htm)

## **8. Материально-техническое обеспечение дисциплины(модуля)**

Освоение дисциплины "Корпусное программное обеспечение 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.

Компьютерный класс, представляющий собой рабочее место преподавателя и не менее 15 рабочих мест студентов, включающих компьютерный стол, стул, персональный компьютер, лицензионное программное обеспечение. Каждый компьютер имеет широкополосный доступ в сеть Интернет. Все компьютеры подключены к корпоративной компьютерной сети КФУ и находятся в едином домене.

Лингафонный кабинет, представляющий собой универсальный лингафонно-программный комплекс на базе компьютерного класса, состоящий из рабочего места преподавателя (стол, стул, монитор, персональный компьютер с программным обеспечением SANAKO Study Tutor, головная гарнитура), и не менее 12 рабочих мест студентов (специальный стол, стул, монитор, персональный компьютер с программным обеспечением SANAKO Study Student, головная гарнитура), сетевого коммутатора для структурированной кабельной системы кабинета.

Лингафонный кабинет представляет собой комплекс мультимедийного оборудования и программного обеспечения для обучения иностранным языкам, включающий программное обеспечение управления классом и SANAKO Study 1200, которые дают возможность использования в учебном процессе интерактивные технологии обучения с использованием современных мультимедийных средств, ресурсов Интернета.

Программный комплекс SANAKO Study 1200 дает возможность инновационного ведения учебного процесса, он предлагает широкий спектр видов деятельности (заданий), поддерживающих как практики слушания, так и тренинги речевой активности: практика чтения, прослушивание, следование образцу, обсуждение, круглый стол, использование Интернета, самообучение, тестирование. Преподаватель является центральной фигурой процесса обучения. Ему предоставляются инструменты управления классом. Он также может использовать многочисленные методы оценки достижений учащихся и следить за их динамикой. SANAKO Study 1200 предоставляет учащимся наилучшие возможности для выполнения речевых упражнений и заданий, основанных на текстах, аудио- и видеоматериалах. Вся аудитория может быть разделена на подгруппы. Это позволяет организовать отдельную траекторию обучения для каждой подгруппы. Учащиеся могут работать самостоятельно, в автономном режиме, при этом преподаватель может контролировать их действия. В состав программного комплекса SANAKO Study 1200 также входит модуль Examination Module - модуль создания и управления тестами для проверки конкретных навыков и способностей учащегося. Гибкость данного модуля позволяет преподавателям легко варьировать типы вопросов в тесте и редактировать существующие тесты.

Также в состав программного комплекса SANAKO Study 1200 также входит модуль обратной связи, с помощью которых можно в процессе занятия провести экспресс-опрос аудитории без подготовки большого теста, а также узнать мнение аудитории по какой-либо теме.

Каждый компьютер лингафонного класса имеет широкополосный доступ к сети Интернет, лицензионное программное обеспечение. Все универсальные лингафонно-программные комплексы подключены к корпоративной компьютерной сети КФУ и находятся в едином домене.

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.

В лекциях используются мультимедийные презентации, которые выносятся на большой экран. На практических занятиях применяется программное обеспечение, установленное в компьютерном классе.

Для итогового контроля применяется сервер дистанционного электронного тестирования <http://test.utmn.ru>.

Программа составлена в соответствии с требованиями ФГОС ВПО и учебным планом по направлению 45.04.01 "Филология" и магистерской программе Деловой английский язык .

Автор(ы):

Каюмова Д.Ф. \_\_\_\_\_

"\_\_" \_\_\_\_\_ 201\_\_ г.

Рецензент(ы):

Давлетбаева Д.Н. \_\_\_\_\_

"\_\_" \_\_\_\_\_ 201\_\_ г.