Учреждение образования

«Гомельский государственный университет имени Франциска Скорины»

**УТВЕРЖДАЮ**

Проректор по учебной работе

ГГУ имени Ф. Скорины

\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_ И. В. Семченко

(подпись)

\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_

(дата утверждения)

Регистрационный № УД-\_\_\_\_\_\_\_\_\_\_\_/уч.

**КОРПУСНЫЕ МЕТОДЫ**

**ИЗУЧЕНИЯ ТЕКСТА**

Учебная программа учреждения высшего образования

по учебной дисциплине для специальности

1-21 05 02-01 Русская филология (литературно-редакционная деятельность)

специализации 1-21 05 02-01 01 Языкознание

2019

Учебная программа составлена на основе Образовательного стандарта ОСВО 1-21 05 02-2013, учебного плана учреждения образования «Гомельский государственный университет имени Ф. Скорины», регистрационный номер D 21-02-16/ин., утвержденного 11.06.2016.

**СОСТАВИТЕЛЬ:**

И. Г. Гомонова − доцент кафедры русского, общего и славянского языкознания УО «ГГУ имени Ф. Скорины», кандидат филологических наук.

**РЕКОМЕНДОВАНА К УТВЕРЖДЕНИЮ:**

Кафедрой русского, общего и славянского языкознания ГГУ имени Ф. Скорины

(протокол № 9 от 16.05.2019);

Научно-методическим советом ГГУ имени Ф. Скорины

(протокол № 8 от 17.05.2019)

**ПОЯСНИТЕЛЬНАЯ ЗАПИСКА**

**Актуальность изучения дисциплины**

Дисциплина компонента учреждения образования из цикла дисциплин специализации «Корпусные методы изучения текста» знакомит студентов с корпусными технологиями, с использованием которых в настоящее время проводится большинство лингвистических исследований, и тем самым позволяет повысить компетентность студентов в сфере их профессиональной деятельности. Изучение данной дисциплины способствует овладению корпусными методами, необходимыми специалисту-филологу для исследования и преподавания языка.

Учебная программа разработана на основе компетентностного подхода, требований к формированию компетенций, сформулированных в образовательном стандарте ОСВО 1-21 05 02-2013.

**Цель и задачи учебной дисциплины**

Цель дисциплины «Корпусные методы изучения текста» − формирование у студентов системного представления о методах и технологиях корпусной лингвистики, а также о направлениях применения методов и инструментария корпусной лингвистики для решения профессиональных задач в области филологии.

Основные задачи дисциплины включают:

– выработку представлений о типологии лингвистических корпусов, технологиях их создания и об областях применения различных видов корпусов;

– выработку практических навыков работы с программными средствами и информационными ресурсами корпусной лингвистики.

В результате изучения дисциплины студенты должны

**знать**:

– терминологический аппарат изучаемой предметной области;

– классы лингвистических задач, решаемых с применением корпусных технологий;

– характерные особенности доступных on-line корпусов языков специальности;

**уметь**:

– решать профессиональные задачи с обращением к методам и инструментам корпусной лингвистики;

– работать с изучаемыми в пределах курса инструментами корпусной лингвистики;

– работать с доступными on-line корпусами различных языков;

– применять лингвистические корпуса как исследовательский инструмент для изучения текста.

**Требования к уровню освоения содержания учебного материала**

В результате изучения дисциплины студент должен закрепить и развить академические (АК), социально-личностные (СЛК) и профессиональные (ПК) компетенции, предусмотренные в образовательном стандарте ОСВО 1-21 05 02-2013:

1. академические компетенции; студент должен:

АК-1 – уметь применять базовые научно-теоретические знания для решения теоретических и практических задач;

АК-2 – владеть системным и сравнительным анализом;

АК-3 – владеть исследовательскими навыками;

АК-4 – уметь работать самостоятельно;

АК-5 – быть способным порождать новые идеи, обладать креативностью;

АК-6 – владеть междисциплинарным подходом при решении проблем;

АК-7 – иметь навыки, связанные с использованием технических устройств;

АК-8 – обладать навыками устной и письменной коммуникации;

АК-9 – уметь учиться, повышать свою квалификацию в течение жизни.

1. социально-личностные компетенции; студент должен:

СЛК-5 – быть способным к критике и самокритике;

СЛК-7 – логично, аргументированно и ясно строить устную и письменную речь, использовать навыки публичной речи, ведения дискуссии и полемики.

1. профессиональные компетенции; студент должен быть способен:

ПК-7 – планировать, организовывать и вести научно-исследовательскую деятельность в области филологии;

ПК-8 – выбирать необходимые методы исследования, модифицировать существующие и применять новые методы, исходя из задач конкретного исследования;

ПК-9 – использовать в работе новые компьютерные методы сбора, обработки и хранения информации;

ПК-10 – представлять итоги научной работы в соответствии с предъявляемыми требованиями;

ПК-11 – применять современную методику реферирования и редактирования текстов;

ПК-12 – пользоваться научной и справочной литературой на русском и белорусском языках;

ПК-13 – применять современную методологию лингвистических и литературоведческих исследований.

**Структура содержания учебной дисциплины**

Содержание дисциплины представлено в виде тем, которые характеризуются самостоятельными укрупненными дидактическими единицами содержания обучения.

Содержание тем опирается на приобретенные ранее студентами компетенции при изучении дисциплин «Введение в языкознание» и «Современный русский язык».

**Методы (технологии) обучения**

Основными методами (технологиями) обучения, отвечающими цели и задачам изучения дисциплины являются:

− модульное и дифференцированное обучение;

− элементы проблемного обучения (проблемное изложение, вариативное изложение, частично-поисковый метод), реализуемые на лекционных занятиях;

− элементы учебно-исследовательской деятельности, метод творческого задания, методики активной оценки, прогностической и ретроспективной оценки, используемые на практических занятиях и при самостоятельной работе.

**Диагностика компетенции студента**

Для оценки достижений студентов используется следующий диагностический инструментарий (в скобках какие компетенции проверяются):

− организация промежуточного и рубежного контроля по каждому модулю (АК-1, АК-2, ПК-8, ПК-13);

− выполнение индивидуальных заданий и выступление с докладами на практических занятиях (АК-2, АК-3, АК-4, АК-5, СЛК-7, ПК-12);

− организация итогового контроля (АК-1, АК-2, АК-3, АК-5, СЛК-5, СЛК-7, ПК-10).

Форма обучения – дневная, курс – 3, семестр – 5. Общее количество часов – 56 (1,5 зачетной единицы); аудиторное количество часов – 32, из них: лекции – 12 часов (в том числе управляемая самостоятельная работа – 2 часа), практические занятия – 22 часа. Форма отчетности – зачет в 5 семестре.

**СОДЕРЖАНИЕ УЧЕБНОГО МАТЕРИАЛА**

**Раздел 1. Лингвистические корпуса и их типология**

**Тема 1.1 Корпус текстов как особый лингвистический ресурс**

Определение лингвистического корпуса. Основные признаки корпуса. Проблема репрезентативности и полноты корпуса. Структура корпуса: корпус данных, корпус-менеджер. Особенности терминосистемы современной корпусной лингвистики.

**Тема 1.2 Типология лингвистических корпусов**

Типы лингвистических корпусов и основания их выделения. Особые типы корпусов. Параллельные корпуса: особенности разработки, области применения. Исторические, диалектные и другие типы корпусов, отражающие особенности исходного языкового материала. Мультимедийные корпуса.

**Тема 1.3 Корпуса языка специальности, доступные on-line**

Корпуса русского языка: виды разметки, выходные интерфейсы и возможности корпус-менеджеров. Национальный корпус русского языка: состав, структура, поисковые возможности. Открытые корпусные проекты (на примере проекта OpenCorpora).

**Раздел 2. Технологии корпусной лингвистики и использование корпусов**

**Тема 2.1 Разметка как основной признак корпуса. Метаразметка**

Понятие разметки. Основные системы разметки. Типы разметки: экстралингвистическая разметка / метаразметка, лингвистическая разметка. Метаразметка и ее виды. Функции метаразметки.

**Тема 2.2 Лингвистическая разметка и ее виды**

Основные принципы лингвистической разметки. Виды лингвистической разметки: морфологическая, синтаксическая, семантическая, анафорическая, просодическая. Разметка ручная и автоматическая; их преимущества и недостатки.

**Тема 2.3 Корпусные технологии в исследовании текста**

Лингвистические задачи, решаемые с применением корпусных технологий. Корпуса как средство получения статистики и конкордансов. Использование метаинформации и лингвистической информации корпуса в исследовании текста.

**УЧЕБНО-МЕТОДИЧЕСКАЯ КАРТА УЧЕБНОЙ ДИСЦИПЛИНЫ**

|  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- |
| **Номер раздела,**  **темы** | **Название раздела, темы** | **Количество аудиторных часов** | | | | | **Количество часов УСР** | **Форма**  **контроля**  **знаний** |
| **Лекции** | **Практические занятия** | **Семинарские занятия** | **Лабораторные занятия** | **Иное** |
| 1 | 2 | 3 | 4 | 5 | 6 | 7 | 8 | 9 |
| **1** | **Лингвистические корпуса и их типология** | **6** | **12** | **−** | **−** | **−** | **−** |  |
| 1.1 | **Корпус текстов как особый лингвистический ресурс**  1. Определение лингвистического корпуса. Основные признаки корпуса.  2. Проблема репрезентативности и полноты корпуса.  3. Структура корпуса: корпус данных, корпус-менеджер.  4. Особенности терминосистемы современной корпусной лингвистики. | 2 | 2 | − | − | − | − | Терминологический диктант |
| 1.2 | **Типология лингвистических корпусов**  1. Типы лингвистических корпусов и основания их выделения.  2. Особые типы корпусов.  3. Параллельные корпуса: особенности разработки, области применения.  4. Исторические, диалектные и другие типы корпусов, отражающие особенности исходного языкового материала. Мультимедийные корпуса. | 2 | 4 | − | − | − | − | Выполнение практической работы с использованием корпусных технологий |
| 1.3 | **Корпуса языка специальности, доступные on-line**  1. Корпуса русского языка: виды разметки, выходные интерфейсы и возможности корпус-менеджеров.  2. Национальный корпус русского языка: состав, структура, поисковые возможности.  3. Открытые корпусные проекты (на примере проекта OpenCorpora). | 2 | 6 | **−** | − | − | − | Выполнение практической работы с использованием корпусных технологий |
| 2 | **Технологии корпусной лингвистики и использование корпусов** | **4** | **10** | **−** | **−** | **−** | **2** |  |
| 2.1 | **Разметка как основной признак корпуса. Метаразметка**  1. Понятие разметки.  2. Основные системы разметки.  3. Типы разметки: экстралингвистическая разметка / метаразметка, лингвистическая разметка.  4. Метаразметка и ее виды. Функции метаразметки. | 2 | 2 | **−** | **−** | **−** | **−** | Выполнение практической работы с использованием корпусных технологий |
| 2.2 | **Лингвистическая разметка и ее виды**  1. Основные принципы лингвистической разметки.  2. Виды лингвистической разметки: морфологическая, синтаксическая, семантическая, анафорическая, просодическая.  3. Разметка ручная и автоматическая; их преимущества и недостатки. | 2 | 6 | **−** | **−** | **−** | **−** | Выполнение практической работы с использованием корпусных технологий |
| 2.3 | **Корпусные технологии в исследовании текста**  1. Лингвистические задачи, решаемые с применением корпусных технологий. 2. Корпуса как средство получения статистики и конкордансов.  3. Использование метаинформации и лингвистической информации корпуса в исследовании текста. | − | 2 | − | **−** | − | 2 | Выполнение и защита индивидуального проекта |
|  | **Всего аудиторных часов за семестр:** | **10** | **22** | **−** | **−** | **−** | **2** | **Зачет в 5 семестре** |

Доцент кафедры русского, общего и славянского языкознания И. Г. Гомонова

**ИНФОРМАЦИОННО-МЕТОДИЧЕСКАЯ ЧАСТЬ**

**Методические рекомендации по организации и выполнению УСР**

Для самостоятельного изучения выделяется тема «Корпусные технологии в исследовании текста».

Цели самостоятельного изучения данной темы:

− активизация учебно-познавательной деятельности обучающихся;

− формирование у обучающихся умений и навыков самостоятельного приобретения и обобщения знаний;

− формирование у обучающихся умений и навыков самостоятельного применения знаний на практике.

**Учебная программа УСР**

**Тема 2.3** Корпусные технологии в исследовании текста

Цели: сформировать достаточные знания по учебному материалу на уровне узнавания; сформировать компетенции на уровне воспроизведения; сформировать компетенции на уровне применения полученных знаний.

**Виды заданий УСР с учетом модулей сложности**

А) *Задания, формирующие знания по учебному материалу на уровне узнавания:*

1. Изучите электронный вариант лекции и дополнительную литературу по теме.

2. Составьте опорный конспект по материалам лекции и дополнительной литературы.

*Форма выполнения заданий* – индивидуальная.

*Форма контроля выполнения заданий* – проверка опорных конспектов.

Б) *Задания, формирующие компетенции на уровне воспроизведения:*

1. Решите на базе корпуса текстов предложенные лингвистические задачи.

*Форма выполнения заданий* – индивидуальная.

*Форма контроля выполнения заданий* – проверка решенных задач.

В) *Задания, формирующие компетенции на уровне применения полученных знаний:*

1. Сформулируйте самостоятельно и решите на базе корпуса текстов лингвистические задачи.

*Форма выполнения заданий –* индивидуальная.

*Форма контроля выполнения заданий –* защита индивидуального проекта.

*Учебно-методическое обеспечение:*

1. Захаров, В. П. Корпусная лингвистика : учебник для студентов гуманитарных вузов / В. П. Захаров, С. Ю. Богданова. – Иркутск : ИГЛУ, 2011. – 161 с.

2. Зубов, А. В. Информационные технологии в лингвистике : учеб. пособие / А. В. Зубов, И. И. Зубова. – М. : Академия, 2004. – 208 с.

3. Национальный корпус русского языка [Электронный ресурс]. – Режим доступа : http://ruscorpora.ru/.

4. Хроленко, А. Т. Современные информационные технологии для гуманитария : практич. руководство / А. Т. Хроленко, А. В. Денисов. – М. : Флинта : Наука, 2007. – 128 с.

**Перечень тем практических занятий**

1. Корпус текстов как особый лингвистический ресурс.

2. Типология лингвистических корпусов.

3. Особые типы корпусов.

4. Корпуса русского языка, доступные on-line.

5. Национальный корпус русского языка: состав и структура.

6. Национальный корпус русского языка: поисковые возможности.

7. Разметка как основной признак корпуса. Метаразметка.

8. Лингвистическая разметка в корпусе: морфологические признаки.

9. Лингвистическая разметка в корпусе: семантические признаки.

10. Лингвистическая разметка в корпусе: синтаксические признаки.

11. Корпусные технологии в исследовании текста.

**Рекомендуемые формы контроля знаний**

1. Устный опрос.

2. Терминологический диктант.

3. Анализ практической работы, выполненной с использованием корпусных технологий.

4. Защита индивидуального проекта, выполненного с использованием корпусных технологий.

**Рекомендуемые вопросы к зачету**

1. Корпус текстов как особый лингвистический ресурс. Основные признаки корпуса.

2. Проблема репрезентативности и полноты корпуса.

3. Структура корпуса: корпус данных, корпус-менеджер.

4. Типы лингвистических корпусов и основания их выделения.

5. Особые типы корпусов.

6. Корпуса русского языка, доступные on-line.

7. Национальный корпус русского языка: состав и структура.

8. Национальный корпус русского языка: поисковые возможности.

9. Открытые корпусные проекты.

10. Разметка как основной признак корпуса.

11. Основные системы разметки.

12. Типы разметки: экстралингвистическая и лингвистическая разметка.

13. Метаразметка и ее виды. Функции метаразметки.

14. Основные принципы лингвистической разметки.

15. Лингвистическая разметка в корпусе: морфологические признаки.

16. Лингвистическая разметка в корпусе: семантические признаки.

17. Лингвистическая разметка в корпусе: синтаксические признаки.

18. Разметка ручная и автоматическая; их преимущества и недостатки.

19. Лингвистические задачи, решаемые с применением корпусных технологий.

20. Использование метаинформации и лингвистической информации корпуса в исследовании текста.

**Рекомендуемая литература**

**Основная**

1. Баранов, А. Н. Введение в прикладную лингвистику : учеб. пособие / А. Н. Баранов. − 2-е изд., испр. – М. : Эдиториал УРСС, 2003. – 360 с.

2. Введение в электронные лингвистические ресурсы : учеб. пособие для студ. вузов / сост. В. Е. Гольдин, О. Ю. Крючкова. – Саратов, 2011. – 63 с.

3. Грудева, Е. В. Корпусная лингвистика [Электронный ресурс] : учеб. пособие / Е. В. Грудева. − 3-е изд., стер. − М. : Флинта, 2017 . − 165 с. − Режим доступа: https://rucont.ru/efd/244183.

4. Захаров, В. П. Корпусная лингвистика : учебно-метод. пособие / В. П. Захаров. – СПб. : Издательство СПбГУ, 2006. – 48 с.

5. Захаров, В. П. Корпусная лингвистика : учебник для студентов гуманитарных вузов / В. П. Захаров, С. Ю. Богданова. – Иркутск : ИГЛУ, 2011. – 161 с.

6. Зубов, А. В. Информационные технологии в лингвистике : учеб. пособие / А. В. Зубов, И. И. Зубова. − 2-е изд., испр. – М. : Академия, 2012. – 205 с.

7. Копотев, М. В. Введение в корпусную лингвистику : учеб. пособие / М. В. Копотев. – Прага : Animedia Company, 2014. − 230 с.

8. Хроленко, А. Т. Современные информационные технологии для гуманитария : практич. руководство / А. Т. Хроленко, А. В. Денисов. – М. : Флинта : Наука, 2007. – 128 с.

9. Щипицина, Л. Ю. Информационные технологии в лингвистике : учеб. пособие / Л. Ю. Щипицина. − М. : Флинта, 2013. − 128 с.

**Дополнительная**

1. Боярский, К. К. Введение в компьютерную лингвистику : учеб. пособие / К. К. Боярский. – СПб. : НИУ ИТМО, 2013. – 72 с.

2. Всеволодова, А. В. Компьютерная обработка лингвистических данных / А. В. Всеволодова. – М. : Флинта : Наука, 2007. – 96 с.

3. Захаров, В. П. Информационно-поисковые системы : учебно-метод. пособие. − СПб. : Издательство СПбГУ, 2005. – 48 с.

4. Зубов, А. В. Основы искусственного интеллекта для лингвистов / А. В. Зубов, И. И. Зубова. – М. : «Логос», 2007. – 320 с.

5. Корпусные исследования по русской грамматике. − М. : Пробел, 2009. − 516 с.

6. Национальный корпус русского языка: 2003–2005: результаты и перспективы : сб. ст. / Ин-т рус. яз. им. В. В. Виноградова Рос. Акад. наук. – М. : Индрик, 2005. – 343 с.

7. Национальный корпус русского языка: 2006–2008. Новые результаты и перспективы / отв. ред. В. А. Плунгян. – СПб. : Нестор-История, 2009.– 502 с.

8. Национальный корпус русского языка и проблемы гуманитарного образования : учебно-метод. пособие / отв. ред.-сост. Н. Р. Добрушина. – М. : ТЕИС, 2007. − 237 с.

9. Потапова, Р. К. Новые информационные технологии и лингвистика : учеб. пособие / Р. К. Потапова. – 4-е изд. – М. : КомКнига, 2005. – 368 с.

**Ресурсы электронного доступа**

1. Беларускі N-корпус [Электронный ресурс]. – Режим доступа : http://bnkorpus.info/.

2. Вавилонская башня – собрание ЛБД и других электронных языковых ресурсов, инициированное С. А. Старостиным [Электронный ресурс]. – Режим доступа : http://starling.rinet.ru.

3. Гомонова, И. Г. Введение в корпусную лингвистику : лекции / И. Г. Гомонова. − Гомель, [б.г]. − 41 с. [Электронный ресурс]. – Режим доступа : http://elib.gsu.by/handle/123456789/2116.

4. ДИАЛОГ: Международная конференция по компьютерной лингвистике [Электронный ресурс] – Режим доступа : http://www.dialog-21.ru.

5. Национальный корпус русского языка [Электронный ресурс]. – Режим доступа : http://ruscorpora.ru/.

6. Рыков, В. В. Курс лекций по корпусной лингвистике / В. В. Рыков [Электронный ресурс]. – Режим доступа : http://rykov-cl.narod.ru/c.html.

7. Список корпусов на сайте Национального корпуса русского языка [Электронный ресурс]. – Режим доступа : http://www.ruscorpora.ru/corpora-other.html.

8. Филология и лингвистика [Электронный ресурс]. – Режим доступа : http://filologia.su/freymy/.

9. Языкознание.ру [Электронный ресурс]. – Режим доступа : http://yazykoznanie.ru.

**ПРОТОКОЛ СОГЛАСОВАНИЯ УЧЕБНОЙ ПРОГРАММЫ**

**ПО ИЗУЧАЕМОЙ УЧЕБНОЙ ДИСЦИПЛИНЕ**

**«Корпусные методы изучения текста»**

**с другими дисциплинами специальности**

**1-21 05 02-01 Русская филология**

**(литературно-редакционная деятельность)**

**специализации 1-21 05 02-01 01 Языкознание**

|  |  |  |  |
| --- | --- | --- | --- |
| Название  дисциплины,  с которой  требуется согласование | Название  кафедры | Предложения  об изменениях  в содержании учебной программы  по изучаемой учебной  дисциплине | Решение, принятое кафедрой, разработавшей учебную программу  (с указанием даты и номера протокола) |
| Информационные технологии  в филологии | Кафедра русского, общего и славянского языкознания | Утверждена  без изменений | Рекомендовать  к утверждению учебную программу  в представленном варианте,  протокол № 9  от 16.05.2019 |

**ДОПОЛНЕНИЯ И ИЗМЕНЕНИЯ К УЧЕБНОЙ ПРОГРАММЕ**

**ПО ИЗУЧАЕМОЙ УЧЕБНОЙ ДИСЦИПЛИНЕ**

**«Корпусные методы изучения текста»**

**на 20 / 20 учебный год**

|  |  |  |
| --- | --- | --- |
| №№  пп | Дополнения и изменения | Основание |
|  |  |  |

Учебная программа пересмотрена и одобрена на заседании кафедры

русского, общего и славянского языкознания (протокол № \_\_\_\_ от \_\_\_\_\_\_\_\_ 20\_\_ г.)

Заведующий кафедрой

доктор филологических наук, доцент \_\_\_\_\_\_\_\_\_\_\_\_\_\_ Е. В. Ничипорчик

УТВЕРЖДАЮ

Декан факультета

кандидат филологических наук, доцент \_\_\_\_\_\_\_\_\_\_\_\_\_ Е. Н. Полуян