

Документ подписан простой электронной подписью
Информация о владельце:
ФИО: Наумова Наталия Александровна
Должность: Ректор
Дата подписания: 24.10.2024 14:21:41
Уникальный программный ключ:
6b5279da4e034bff679172803da5b7b5559fc69e2

МИНИСТЕРСТВО ПРОСВЕЩЕНИЯ РОССИЙСКОЙ ФЕДЕРАЦИИ
Федеральное государственное автономное образовательное учреждение высшего образования
«ГОСУДАРСТВЕННЫЙ УНИВЕРСИТЕТ ПРОСВЕЩЕНИЯ»
(ГОСУДАРСТВЕННЫЙ УНИВЕРСИТЕТ ПРОСВЕЩЕНИЯ)

Лингвистический факультет
Кафедра переводоведения и когнитивной лингвистики

Согласовано
деканом лингвистического факультета
« 14 » марта 2024 г.
/ Вековищева С.Н. /

Рабочая программа дисциплины

Корпусные исследования в лингвистике

Направление подготовки
45.03.02 Лингвистика

Профиль:
Перевод и переводоведение (английский язык + немецкий или французский языки)

Квалификация
Бакалавр

Форма обучения
Очная

Согласовано учебно-методической комиссией
лингвистического факультета
Протокол «14» марта 2024 г. № 5
Председатель УМКом / Горбачева О.А. /

Рекомендовано кафедрой
переводоведения и когнитивной
лингвистики
Протокол от «11» марта 2024 г. № 7
Зав. кафедрой / Ахренова.Н.А. /

Мытищи
2024

Автор-составитель:

Ахренова Наталья Александровна, профессор, доктор филологических наук

Рабочая программа дисциплины «Корпусные исследования в переводе» составлена в соответствии с требованиями Федерального Государственного образовательного стандарта высшего образования по направлению подготовки 45.05.01 Лингвистика, утвержденного приказом МИНОБРНАУКИ России от 12.08.2020 г. № 969.

Дисциплина относится к Блоку ФТД. Факультативные дисциплины (модули) и является факультативной дисциплиной

Реализуется в формате электронного обучения с применением дистанционных образовательных технологий.

Год начала подготовки (по учебному плану) 2024

СОДЕРЖАНИЕ

1. Планируемые результаты обучения.....	4
2. Место дисциплины в структуре образовательной программы.....	4
3. Объем и содержание дисциплины.....	5
4. Учебно-методическое обеспечение самостоятельной работы обучающихся	7
5. Фонд оценочных средств для проведения текущей и промежуточной аттестации по дисциплине.....	8
6. Учебно-методическое и ресурсное обеспечение дисциплины.....	16
7. Методические указания по освоению дисциплины.....	17
8. Информационные технологии для осуществления образовательного процесса по дисциплине	17
9. Материально-техническое обеспечение дисциплины.....	17

1. ПЛАНИРУЕМЫЕ РЕЗУЛЬТАТЫ ОБУЧЕНИЯ

1.1. Цель и задачи дисциплины

Целями освоения дисциплины «Корпусная лингвистика» являются ознакомление обучающихся с основными понятиями и возможностями корпусной лингвистики, освоение на практике технологически несложных методов корпусной лингвистики, формирование практических навыков работы с наиболее известными мировыми лингвистическими ресурсами, использование лингвистических корпусов для решения конкретных переводческих и исследовательских задач, а также приобретение опыта работы с корпусными средствами для построения своего собственного переводческого и исследовательского корпусов.

Задачи дисциплины:

1. представить возникновение и развитие корпусной лингвистики как науки;
2. сформировать представление о предмете корпусной лингвистики, о категориях ее понятийного аппарата;
3. ознакомить студентов с существующими корпусами русского языка, их возможностями;
4. выработать методику анализа параметров корпуса.

Планируемые результаты дисциплины

В результате освоения данной дисциплины у обучающихся формируется следующая компетенция:

СПК-1. Способен осуществлять письменный и устный перевод текстов различной тематической направленности на иностранном (иностранных) языках с соблюдением стилистических норм в соответствии с коммуникативной ситуацией.

СПК-4. Способен применять основные современные методы научного исследования, в том числе и в смежных областях, в самостоятельных исследованиях

2. МЕСТО ДИСЦИПЛИНЫ В СТРУКТУРЕ ОБРАЗОВАТЕЛЬНОЙ ПРОГРАММЫ

Дисциплина относится к Блоку ФТД. Факультативные дисциплины (модули) и является факультативной дисциплиной

Данная дисциплина ориентирована на формирование гуманистического мировоззрения студентов, расширения их общекультурной, языковой и практической профессиональной подготовки, а также совершенствования навыков работы с корпусами текстов для решения профессиональных переводческих задач.

В методическом плане дисциплина является теоретическим курсом, получаемым студентами на третьем курсе и погружающих обучающихся в специфику профессиональной деятельности наравне с учебными дисциплинами *«Практический курс первого иностранного языка (английский язык)»* и *«Теория перевода»*, *«Информационные технологии в профессиональной деятельности»*. Изучая дисциплину *«Корпусные исследования в переводе»*, студенты приобретают знания, необходимые для таких дисциплин, как *«Синхронный перевод»*, *«Письменный перевод»*, *«Судебный перевод»*, *«Последовательный и двусторонний перевод»*.

Полученные в процессе обучения знания будут использованы при изучении таких дисциплин как *«Письменный перевод»*, *«Аудиовизуальный перевод»*, *«Художественный перевод»*, *«Судебный перевод»*.

Требования к «входным» знаниям, умениям и готовностям обучающегося, необходимые при освоении данной дисциплины и приобретенные в результате освоения предшествующих дисциплин. Обучающийся, освоивший дисциплину (модуль) должен:

знать:

- ключевые понятия, актуальные проблемы, перспективные направления развития корпусной лингвистики;
- основные типы корпусов, наиболее известные и широко применяемые иноязычные и русскоязычные корпуса;
- способы использования корпусов для различных целей;
- технологию создания корпусов, основные форматы и способы организации корпусов;

уметь:

- определять тип и способы применения незнакомых корпусов;
- осваивать новые компьютерные инструменты;
- интерпретировать данные, полученные из корпусов;
- сравнивать и оценивать функциональность корпусов и инструментов по результатам работы с ними;

владеть навыками

- работы с несколькими существующими иноязычными и русскоязычными корпусами и корпусными менеджерами;
- формулирования сложных лингвистических запросов с использованием специальных обозначений (тэгов) и регулярных выражений, разметки текстов и чтения разметки;
- проектирования собственного корпуса;
- создания небольших специализированных корпусов.

3. ОБЪЁМ И СОДЕРЖАНИЕ ДИСЦИПЛИНЫ.

3.1. Объём дисциплины

Показатель объёма дисциплины	Форма обучения
	Очная
Объём дисциплины в зачетных единицах	1
Объём дисциплины в часах	36
Контактная работа	20,2(10) ¹
Лекции	-
Практические занятия	20(10) ²
Контактные часы на промежуточную аттестацию	0.2
Зачет	0.2
Самостоятельная работа	8
Контроль	7.8

Форма промежуточной аттестации – зачет в 5 семестре

3.2. Содержание дисциплины

Наименование разделов (тем) дисциплины с	Количество часов
--	------------------

¹ Реализуется в формате электронного обучения с применением дистанционных образовательных технологий.

² Реализуется в формате электронного обучения с применением дистанционных образовательных технологий.

кратким содержанием	Лекции	Практические занятия
Тема 1. Введение в корпусную лингвистику. Корпус как поисковая система. Основные понятия и определения. Разметка. Репрезентативность. Метаданные. Корпусный менеджер. Treebank. Лемматизация. Конкорданс. Параллельный корпус. Лингвистические (языковые) и нелингвистические корпуса. История лингвистических корпусов: от картотеки к корпусу. Корпусная лингвистика: современное состояние. Корпусная лингвистика в России. Корпусоподобные интерфейсы между лингвистом и поисковыми системами Интернета. Лингвистические исследования, базирующиеся на корпусах		2(2) ³
Тема 2. Практика по работе в наиболее известных корпусах. Создание электронной хрестоматии по корпусной лингвистике. Исследование механизмов взаимодействия корпуса текстов и электронной картотеки (корпусы цитат). Создание веб-сайта по корпусной лингвистике. Предварительные работы по созданию корпуса. Проблемы репрезентативности. Отбор источников. Внешние и внутренние критерии отбора. Нормализация файлов. Графематический анализ. Средства создания и разметки корпусов. Понятие разметки. Типы разметки. Автоматический морфологический и синтаксический анализ. Металингвистическая разметка. Параллельные корпуса. Проблема выравнивания. Стандартизация в корпусной лингвистике. Языковые средства представления размеченных текстов. Международные стандарты и проекты (TEI, EAGLES, CDIF, XCES)		8 (4) ⁴

³ Реализуется в формате электронного обучения с применением дистанционных образовательных технологий.

⁴ Реализуется в формате электронного обучения с применением дистанционных образовательных технологий.

Тема 3. Качественные и количественные данные в корпусе Способы использования корпусов в лингвистических исследованиях. Исследование способов использования корпусов в лексикографии. Изучение средств обработки корпусных данных, представленных на языке XML.		4 (2) ⁵
Тема 4. Принципы и способы составления корпусов Обзор существующих корпусов различных типов Классификация (типология) корпусов по различным основаниям. Типы корпусов по задачам. Типы корпусов по формальным признакам.		4(2) ⁶
Тема 5. Корпусные методы в применении к переводу Корпусная лексикология и лексикография. Корпусная грамматика. Исследования дискурса, основанные на корпусах		2
Итого:		20(10)⁷

4. УЧЕБНО-МЕТОДИЧЕСКОЕ ОБЕСПЕЧЕНИЕ САМОСТОЯТЕЛЬНОЙ РАБОТЫ ОБУЧАЮЩИХСЯ

Темы для самостоятельного обучения	Изучаемые вопросы	Кол -во часов	Формы самостоятельно й работы	Методическое обеспечение	Форма отчётности
Введение в корпусную лингвистику	1.Теоретическая значимость корпусных данных. 2.Практическое использование корпусных данных.	1	проработка конспектов лекций, изучение теоретического материала, подготовка сообщений к выступлению на	Основная и дополнительная литература, интернет-источники	Доклад , практическое задание

⁵ Реализуется в формате электронного обучения с применением дистанционных образовательных технологий.

⁶ Реализуется в формате электронного обучения с применением дистанционных образовательных технологий.

⁷ Реализуется в формате электронного обучения с применением дистанционных образовательных технологий.

			семинаре		
Практика по работе в наиболее известных корпусах.	1. Использование интерфейса и запросов, использование экспортированных данных. British National Corpus, НКРЯ, Intercorp, синтаксические корпуса. 2. Наиболее известные корпуса русского языка	2	проработка конспектов лекций, изучение теоретического материала, подготовка сообщений к выступлению на семинаре, подготовка проекта	Основная и дополнительная литература, интернет-источники	Доклад, практическое задание, проект
Качественные и количественные данные в корпусе	1.Получение корпусных примеров: как и зачем. 2. Корпусное изучение вариативности. 3.Проблема взаимодействия факторов. 4.Основы статистики.	1	проработка конспектов лекций, изучение теоретического материала, подготовка сообщений к выступлению на семинаре	Основная и дополнительная литература, интернет-источники	Доклад, практическое задание, проект
Принципы и способы составления корпусов	1.Принципы составления корпусов (репрезентативность, сбалансированность, металингвистическая разметка, типы корпусов). 2. Корпусные средства. 3. Лингвистическое аннотирование	2	проработка конспектов лекций, изучение теоретического материала, подготовка сообщений к выступлению на семинаре	Основная и дополнительная литература, интернет-источники	Доклад, практическое задание, проект
Корпусные методы в применении и к переводу	1.Корпусная фонетика и фонология. 2.Корпусная лексикология и лексикография. 3.Корпусный морфосинтаксис. 4.Использование корпуса для литературоведения	2	проработка конспектов лекций, изучение теоретического материала, подготовка сообщений к выступлению на семинаре	Основная и дополнительная литература, интернет-источники	Доклад, практическое задание, проектная работа
Итого		8			

5. ФОНД ОЦЕНОЧНЫХ СРЕДСТВ ДЛЯ ПРОВЕДЕНИЯ ТЕКУЩЕЙ И ПРОМЕЖУТОЧНОЙ АТТЕСТАЦИИ ПО ДИСЦИПЛИНЕ

5.1. Перечень компетенций с указанием этапов их формирования в процессе освоения образовательной программы:

<i>Код и наименование компетенции</i>	<i>Этапы формирования компетенции</i>
СПК-1. Способен осуществлять письменный и устный перевод текстов различной тематической направленности на иностранном (иностраннных) языках с соблюдением стилистических норм в соответствии с коммуникативной ситуацией.	1. Работа на учебных занятиях 2. Самостоятельная работа.
СПК-4. Способен применять основные современные методы научного исследования, в том числе и в смежных областях, в самостоятельных исследованиях	1. Работа на учебных занятиях 2. Самостоятельная работа.

5.2. Описание показателей и критериев оценивания компетенций на различных этапах их формирования, описание шкал оценивания

<i>Оцениваемые компетенции</i>	<i>Уровень сформированности</i>	<i>Этап формирования</i>	<i>Описание показателей</i>	<i>Критерии оценивания</i>	<i>Шкала оценивания</i>
СПК-1	Пороговый	1. Работа на учебных занятиях 2. Самостоятельная работа.	Знать: основы осуществления письменного и устного перевода с использованием корпусов в соответствии с коммуникативной ситуацией Уметь: осуществлять письменный и устный перевод текстов различной тематической направленности на иностранном (иностраннных) языках с использованием корпусного подхода в соответствии с коммуникативной ситуацией	Устный опрос, проект, доклад, практическое задание	Шкала оценивания устного ответа Шкала оценивания проекта Шкала оценивания доклада шкала оценивания практического задания
	Продвинутый	1. Работа на учебных занятиях 2. Самостоятельная	Знать: основы применения стратегий и тактик осуществления письменного и	Устный опрос, проект, доклад, практическое задание	Шкала оценивания устного

		работа. устного перевода с использованием корпусов в соответствии с коммуникативной ситуацией Уметь: осуществлять письменный и устный перевод текстов различной тематической направленности на иностранном (иностраннных) языках с использованием корпусного подхода в соответствии с коммуникативной ситуацией Владеть: навыками осуществления письменного и устного перевода текстов различной тематической направленности на иностранном (иностраннных) языках с использованием корпусного подхода в соответствии с коммуникативной ситуацией		ответа Шкала оценивания проекта Шкала оценивания доклада шкала оценивания практическ ого задания	
СПК - 4	Пороговы й	1. Работа на учебных занятиях 2. Самостоятельная работа.	Знать: основные современные методы научного исследования, в том числе и в смежных областях, в самостоятельных исследованиях с учетом наработок корпусной лингвистики Уметь:	Устный опрос, проект, доклад, практическое задание	Шкала оценивания устного ответа Шкала оценивания проекта Шкала оценивания

			использовать на базовом уровне основные современные методы научного исследования, в том числе и в смежных областях, в самостоятельных исследованиях с учетом наработок корпусной лингвистики		доклада шкала оценивания практического задания
	Продвинутый	1. Работа на учебных занятиях 2. Самостоятельная работа.	Знать: возможности использования основных современных методов научного исследования, в том числе и в смежных областях, в самостоятельных исследованиях с позиций лингвопрагматики Уметь: использовать на продвинутом уровне основных современных методов научного исследования, в том числе и в смежных областях, в самостоятельных исследованиях с учетом наработок корпусной лингвистики Владеть: техниками использования основных современных методов научного исследования, в том числе и в смежных областях, в самостоятельных исследованиях с	Устный опрос, проект, доклад, практическое задание	Шкала оценивания устного ответа Шкала оценивания проекта Шкала оценивания доклада шкала оценивания практического задания

			учетом наработок корпусной лингвистики		
--	--	--	--	--	--

Шкала оценивания устного опроса

Критерии оценивания	Баллы
содержание вопроса раскрыто полностью: информация изложена логично с акцентом на ключевых особенностях литературного явления, предлагаются выводы/рекомендации по его оценке; терминология курса используется адекватно, нет подмены понятий и терминов; для практического анализа в качестве иллюстративных примеров используются адекватные примеры.	2
содержание вопроса раскрыто полностью: информация изложена логично с акцентом на ключевых особенностях описываемого явления, отсутствуют объективные выводы/рекомендации по его оценке; есть неточности в использовании терминологии курса; для практического анализа в качестве иллюстративных примеров используются адекватные примеры.	1,5
содержание вопроса раскрыто: информация изложена логично, однако не выделены ключевые положения, выводы/рекомендации носят субъективный характер; терминология курса используется в целом точно, представлены иллюстративные примеры адекватные рассматриваемым явлениям.	1
содержание вопроса раскрыто: информация изложена схематично, не выделены ключевые положения, выводы/рекомендации носят субъективный характер; терминология курса используется в целом точно, но не представлены иллюстративные примеры адекватные рассматриваемым явлениям	0,5

Шкала оценивания доклада

Критерии оценивания	Баллы
Представленный доклад свидетельствует о проведенном самостоятельном исследовании с привлечением различных источников информации; во время выступления продемонстрированы коммуникативные и рефлексивные умения.	28 баллов
Представленный доклад свидетельствует о проведенном исследовании с привлечением различных источников информации, однако исследование выполнено не самостоятельно, допущены ошибки в изложении материала	22 балла
Представленный доклад свидетельствует о проведенном исследовании, однако слабо	18 баллов

отражает знакомство с предметом/не соответствует теме исследования, исследование выполнено не самостоятельно, допущены ошибки в изложении материала	
Представленный доклад не отражает знакомство с предметом/не соответствует теме исследования, исследование выполнено не самостоятельно, допущены ошибки в изложении материала	12 баллов

Шкала оценивания проекта

Критерии оценивания	Баллы
Тема раскрыта точно и полностью. Под критериями точности и полноты понимаются отсутствие фактических ошибок, правильно приведенная цифровая и буквенная прецизионная информация.	10
Презентация качественно оформлена, соблюден баланс текстовой, графической и иллюстрационной информации	8
Соблюдены все технические требования в оформлении печатного варианта работы	7-4
Содержание темы раскрыто неполно и неверно, имеются фактические ошибки, искажающие смысл, информационные лакуны составляют более 40% текста, прецизионная информация передана неверно, русский язык изобилует речевыми ошибками.	0-3

Шкала оценивания практического задания

Критерии оценивания	Баллы
содержание вопроса раскрыто полностью: информация изложена логично с акцентом на ключевых особенностях литературного явления, предлагаются выводы/рекомендации по его оценке; терминология курса используется адекватно, нет подмены понятий и терминов; для практического анализа в качестве иллюстративных примеров используются адекватные примеры.	5 баллов
содержание вопроса раскрыто полностью: информация изложена логично с акцентом на ключевых особенностях описываемого явления, отсутствуют объективные выводы/рекомендации по его оценке; есть неточности в использовании терминологии курса; для практического анализа в качестве иллюстративных примеров используются адекватные примеры.	4 балла
содержание вопроса раскрыто: информация изложена логично, однако не выделены ключевые положения, выводы/рекомендации носят субъективный характер; терминология курса используется в целом точно, представлены иллюстративные примеры адекватные рассматриваемым явлениям.	3 балла
содержание вопроса раскрыто: информация изложена схематично, не выделены ключевые положения, выводы/рекомендации носят субъективный характер;	2 балла

терминология курса используется в целом точно, но не представлены иллюстративные примеры адекватные рассматриваемым явлениям	
содержание вопроса раскрыто неполностью: изложение материала не структурировано, отсутствуют выводы/рекомендации; представленные примеры недостаточно иллюстративны; не приводятся иллюстративные примеры адекватные рассматриваемым явлениям	1 балл
ответ неверный или нет ответа на вопрос задания/ не соответствует теме/ не представляет авторского вклада (отсутствуют иллюстративные примеры/выводы/рекомендации).	0 баллов

5.3. Типовые контрольные задания или иные материалы, необходимые для оценки знаний, умений и навыков и (или) опыта деятельности, характеризующих этапы формирования компетенций в процессе освоение образовательной программы

Примерные темы докладов

1. «Корпус как идеология» в представлении акад. Плунгяна.
2. Сравнительный анализ корпусов и корпусных менеджеров. Поиск в корпусах: конкордансер и корпусный менеджер как поисковая система. Поисковые системы
3. Интернет как корпусные менеджеры.
4. Сравнение корпусных менеджеров с архитектурой поисковых систем в сети Интернет.
5. Поисковые системы Интернет как корпусные менеджеры. Специализированные корпусные интерфейсы к индексам глобальных поисковых систем
6. Параллельные корпуса. Конкорданс. Многоязычие. Проблема выравнивания.
7. Примеры параллельных корпусов. Параллельные корпуса. Конкорданс. Анализ программных средств создания конкордансов и существующих параллельных корпусов.
8. Использование корпусов в обучении языкам и гуманитарных исследованиях.
9. Лингвистические исследования на базе корпуса: изучение лексики. Лингвистические исследования на базе корпуса НКРЯ и BNC: изучение грамматики, исследование семантики (метафора, этимология).
10. Способы использования корпусов в лингвистических исследованиях.
11. Исследование способов использования корпусов в лексикографии.
12. Изучение средств обработки корпусных данных, представленных на языке XML.
13. Создание электронной хрестоматии по корпусной лингвистике.
14. Исследование механизмов взаимодействия корпуса текстов и электронной картотеки (корпусы цитат).
15. Создание веб-сайта по корпусной лингвистике.
16. Британский национальный корпус
17. Национальный корпус русского языка: история создания, перспективы развития

18. Технологии морфологической и синтаксической разметки НКРЯ
19. Программа MyStem Yandex: разработка, использование, перспективы

Примерные практические задания

1. Посчитайте соотношение частоты употребления слов «этот» и «сей» в следующие периоды: а) 1770-1799 б) 1800-1829 в) 1830-1869.
2. Когда в Корпусе появляются первые примеры глагола «реагировать»? Приведите первые два примера. Что они значат? Когда появляются первые примеры глагола «реагировать» с какой-либо приставкой? Приведите первые два примера.
3. Какая разметка из приведенных ниже является морфологической, а какая семантической?

```
<?xml version="1.0" encoding="windows-1251" ?>
<text>
<p>
<s>
<w>Звонили<ana lemma="ЗВОНИТЬ" pos="Г" gram="мн,нс,нп,дст,прш," /></w>
<w>к<ana lemma="К" pos="ПРЕДЛ" gram="" /></w>
<w>вечерне
<ana lemma="ВЕЧЕРНЯ" pos="С" gram="жр,ед,дт,пр,но," />
<ana lemma="ВЕЧЕРНИЙ" pos="П" gram="ср,ед,кр," /></w>
<pun>.</pun> </s>
<s><w>Торжественный<ana lemma="ТОРЖЕСТВЕННЫЙ" pos="П"
gram="мр,ед,им,вн," /></w>
<w>гул<ana lemma="ГУЛ" pos="С" gram="мр,ед,им,вн,но," /></w>
<w>колоколов
<ana lemma="КОЛОКОЛ" pos="С" gram="мр,мн,рд,но," />
<ana lemma="КОЛОКОЛОВ" pos="С" gram="мр,фам,ед,им,од," /></w>
.....<pun>.</pun> </s></p></text>
```

2)

```
<?xml version="1.0" encoding="cp866"?>
<МЕТА-ДОКУМЕНТ>
<ДОКУМЕНТ>
<НАЗВАНИЕ>Сообщение газеты "Мир Кино"</НАЗВАНИЕ>
<ДАТА>20.03.2000</ДАТА>
<УЧАСТНИК_ОБЩЕНИЯ Имя="Мир Кино" Тип="Автор">
<СОЦИАЛЬНЫЙ_СТАТУС>
<ХАРАКТЕРИСТИКА>Печатное издание</ХАРАКТЕРИСТИКА>
</СОЦИАЛЬНЫЙ_СТАТУС>
</УЧАСТНИК_ОБЩЕНИЯ>
</ДОКУМЕНТ>
```

30. Создайте морфологическую разметку предложения:

4. Создайте морфологическую разметку предложения: Модуль семантического анализа текстов предназначен для нормализации синтаксической структуры, распознавания терминов, классификации терминов по семантическим признакам.

Примерные вопросы для обсуждения на практических занятиях(устный опрос)

1. Понятие корпусной лингвистики. Роль и прикладное назначение корпусных исследований на современном этапе развития языкознания. Отечественные и зарубежные исследователи, ведущие разработки в русле корпусной лингвистики. Значение корпусной лингвистики для современного языкознания. Специфика корпусных баз данных как

лексикографического источника. Сбалансированные и несбалансированные корпуса.

2. Прикладные задачи переводоведения и лингводидактики через призму корпусных исследований. Развитие и ключевые современные направления корпусных исследований. Перспективы разработок в области корпусных баз данных. Преимущества и недостатки корпусных баз данных.
3. Первое и второе поколение корпусов. Типы корпусов: устные письменные, одноязычные и многоязычные, аннотированные и неаннотированные. Параллельные корпуса. Существующие корпуса: Brown Corpus, British National Corpus, Lancaster Corpus, Uppsala Corpus, COCA, COHA, Learners' Corpus, Scottish Corpus, НКРЯ. Проекты Linguae и Reverso Context.
4. Корпусоподобные интерфейсы. Принципы внутренней организации корпусов. Интерфейс корпусной базы данных. Разметка. Виды разметок: частеречная разметка, грамматический парсинг, просодическая аннотация, анафорическая аннотация, семантическая разметка и под. Лемма, лемматизация. Корпусный менеджер. Конкордансер. Программы для работы с корпусными базами данных.
5. Алгоритмы решения типовых задач в области теории перевода и методики обучения иностранным языкам. Коллоstrukции. Коллоstrukционный анализ А. Штефановича. Взаимная информация (MI). Частотность. Поиск по регистру. Методика диахронического изучения языковых трендов. Инструменты межкорпусного сопоставления.
6. Раздел корпусной базы данных KWIC, метод лексико-семантических полей, поиск синонимов и антонимов в корпусных базах данных. Параметры ранжирования результатов в корпусных базах данных.

Примерная тематика проектов:

1. Сравните пару синонимов *кинуть* и *бросить*: проверьте влияние на их распределение не менее чем двух факторов. Приведите количественные данные и запросы, по которым они были получены, примените статистические критерии для проверки значимости различий.
2. Постройте частотный словарь словоформ для заданного текста.
3. Выполните лемматизацию и автоматическую частеречную разметку предложенного текста.
4. Создайте собственный корпус по заданной тематике.

Примерные вопросы к зачету (5 семестр)

1. Предмет корпусной лингвистики
2. Полевая лингвистика и корпусная лингвистика
3. Связь корпусной лингвистики с этнолингвистикой
4. История лингвистических корпусов: от картотеки к корпусу.
5. Классификация (типология) корпусов.

6. Корпусная лингвистика: современное состояние.
7. Корпусная лингвистика в России.
8. Обзор существующих корпусов различных типов.
9. Корпус как поисковая система.
10. Корпусоподобные интерфейсы между лингвистом и поисковыми системами Интернета.
11. Лингвистические исследования, базирующиеся на корпусах.
12. Проблемы репрезентативности корпусов.
13. Проблемы хронологии в общезыковых корпусах.
14. Отбор текстов для корпусов.
15. Графематический анализ.
16. Понятие разметки.
17. Типы разметки.
18. Морфологическая разметка.
19. Синтаксические корпуса (treebanks).
20. Семантическая разметка.
21. Технология создания корпусов. Стадии работы.
22. Понятие корпусоида.
23. Автоматическая морфоразметка.
24. Автоматический синтаксический анализ (parsing).
25. Языковые средства представления размеченных текстов (языки SGML, XML).
26. Международные стандарты (TEI, EAGLES, CDIF, XCES).
27. Способы использования корпусов в лингвистических исследованиях.
28. Исследование способов использования корпусов в лексикографии.
29. Изучение средств обработки корпусных данных, представленных на языке XML.
30. Обзор существующих корпусов различных типов.
31. Классификация (типология) корпусов по различным основаниям.
32. Типы корпусов по задачам. Типы корпусов по формальным признакам.

5.4. Методические материалы, определяющие процедуры оценивания знаний, умений, навыков и (или) опыта деятельности, характеризующих этапы формирования компетенций

В рамках освоения дисциплины предусмотрены следующие формы текущего контроля: доклад, практическое задание, проект, устный опрос

Общее количество баллов по дисциплине – 100 баллов.

Максимальное количество баллов, которое может набрать обучающийся в течение семестра за текущий контроль, равняется 80 баллам.

Промежуточная аттестация проводится устно по вопросам в форме зачета. На зачете обучающийся должен дать продемонстрировать знание теоретических аспектов изучаемого курса. Максимальное число баллов, которые выставляются обучающемуся по итогам зачета - 20 баллов.

Требования к зачету

На зачете дается оценка по следующим критериям:

- понимание и степень усвоения теоретического материала курса в соответствии с учебной программой;
- владение понятийным аппаратом и терминологией данной дисциплины;
- умение связать теорию с практическим применением и владение практическими навыками корпусного анализа и использования корпусного подхода и корпусной лингвистики в объеме программы.

Обязательным условием допуска студентов к зачету является выполнение всех предусмотренных учебной программой видов текущего контроля. На зачете осуществляется комплексная проверка компетенций, знаний, навыков, умений студентов.

Один теоретический вопрос на проверку владения материалом курса корпусные исследования в переводе

Шкала оценивания зачета

Критерии оценивания	Баллы
- обучающийся глубоко и всесторонне усвоил программный материал; – уверенно, логично, последовательно и грамотно его излагает; – опираясь на знания основной и дополнительной литературы, тесно привязывает усвоенные научные положения с практической деятельностью направления; – умело обосновывает и аргументирует выдвигаемые им идеи; – делает выводы и обобщения; – свободно владеет системой специализированных понятий.	20-15
– обучающийся твердо усвоил программный материал, грамотно и по существу излагает его, опираясь на знания основной литературы; – не допускает существенных неточностей; – увязывает усвоенные знания с практической деятельностью направления; – аргументирует научные положения; – делает выводы и обобщения; – владеет системой специализированных понятий	14-10
- обучающийся усвоил только основной программный материал, по существу излагает его, опираясь на знания только основной литературы; – допускает несущественные ошибки и неточности; – испытывает затруднения в практическом применении знаний направления; – слабо аргументирует научные положения; – затрудняется в формулировании выводов и обобщений; – частично владеет системой специализированных понятий.	9-5
– обучающийся не усвоил значительной части программного материала; – допускает существенные ошибки и неточности при рассмотрении проблем в конкретном направлении; – испытывает трудности в практическом применении знаний; – не может аргументировать научные положения; – не формулирует выводов и обобщений.	4-0

Итоговая шкала оценивания дисциплины

Итоговая оценка складывается из оценки за выполнения всех предусмотренных в

программе дисциплины форм отчетности в рамках текущего контроля, а также оценки на промежуточной аттестации.

Баллы, полученные по текущему контролю и промежуточной аттестации	Оценка в традиционной системе
81-100	зачтено
61-80	зачтено
41-60	зачтено
0-40	не зачтено

6. УЧЕБНО-МЕТОДИЧЕСКОЕ И РЕСУРСНОЕ ОБЕСПЕЧЕНИЕ ДИСЦИПЛИНЫ

6.1. Основная литература

1. Захаров, В. П. Корпусная лингвистика : учебник / В. П. Захаров, С. Ю. Богданова. - 3-е изд., перераб. - Санкт-Петербург : СПбГУ, 2020. - 234 с. - ISBN 978-5-288-05997-1. - Текст : электронный. - URL: <https://znanium.com/catalog/product/1244746> (дата обращения: 09.06.2024). – Режим доступа: по подписке.
2. Моделирование в корпусной лингвистике: специализированные корпуса русского языка : монография / В. П. Захаров, И. В. Азаров, О. А. Митрофанова [и др.]. - СПб : Изд-во С.-Петерб. ун-та, 2019. - 208 с. - ISBN 978-5-288-05902-5. - Текст : электронный. - URL: <https://znanium.com/catalog/product/1080953> (дата обращения: 09.06.2024). – Режим доступа: по подписке.
3. Черняховская, Л. А. Перевод и смысловая структура : монография / Л. А. Черняховская. - 2-е изд, доп. - Москва : Флинта, 2021. - 280 с. - ISBN 978-5-9765-4637-0. - Текст : электронный. - URL: <https://znanium.com/catalog/product/1852018> (дата обращения: 09.06.2024). – Режим доступа: по подписке.

6.2. Дополнительная литература

1. Грудева, Е.В. Корпусная лингвистика : учеб. пособие / Е.В. Грудева. - 3-е изд., стер. - Москва : ФЛИНТА, 2017. - 165 с. - ISBN 978-5-9765-1497-3. - Текст : электронный. - URL: <https://znanium.com/catalog/product/1032488> (дата обращения: 09.06.2024). – Режим доступа: по подписке.
2. Гацура, Н. И. Перевод и лингвистический анализ текста : учебно-методическое пособие / Н.И. Гацура. - Омск : Издательство Омского государственного университета, 2018. - 148 с. - ISBN 978-5-7779-2263-2. - Текст : электронный. - URL: <https://znanium.ru/catalog/product/2144433> (дата обращения: 09.06.2024). – Режим доступа: по подписке.

6.3. Ресурсы информационно-телекоммуникационной сети «Интернет»

1. <http://ruscorpora.ru/instruction-main.pdf> (основной подкорпус);
2. <http://ruscorpora.ru/instruction-syntax.html> (синтаксический подкорпус);
3. <http://ruscorpora.ru/instruction-accent.pdf> (акцентологический подкорпус);
4. <http://ruscorpora.ru/instruction-murco.html> (мультимедийный подкорпус);
5. Национальный корпус русского языка (<http://www.ruscorpora.ru>)
6. Корпусы на базе Corpus Technologies (<http://web-corpora.net>)
7. Корпусы Марка Дэвиса: английские, испанский и португальский (<http://corpus.byu.edu>)
8. Корпусы, доступные через интерфейс университета Лидс (<http://corpus.leeds.ac.uk>)
9. Британский национальный корпус <http://www.natcorp.ox.ac.uk/>

10. Корпусы, доступные через интерфейс Kontext (https://kontext.korpus.cz/first_form), в частности:
11. Параллельный корпус Intercorp (<http://www.korpus.cz/intercorp>)
12. Корпусы Araneum
13. Параллельный корпус ParaSol (<http://parasol.unibe.ch>, требуется регистрация)
14. Генеральный интернет-корпус русского языка (<http://int.webcorpora.ru/drake>, требуется регистрация)
15. «Рассказы о сновидениях и другие корпуса звучащей речи» (<http://spokencorpora.ru/>)
16. LinguistList, программное обеспечение: <http://linguistlist.org/sp/software-index.cfm>
17. SIL International, программное обеспечение: <http://software.sil.org/products/>
18. Sketch Engine: <https://the.sketchengine.co.uk/> (требуется регистрация)

Ресурсы сети Интернет

Notepad++ <https://notepad-plus-plus.org/>
MS Office (Word, Excel) или OpenOffice
Google Chrome <https://www.google.ru/chrome/>
Яндекс.Браузер <https://browser.yandex.ru/>
Морфологический анализатор MyStem <https://tech.yandex.ru/mystem/>
Корпусный менеджер AntConc <http://www.laurenceanthony.net/software.html>
Google книги n-грамм анализ: <https://books.google.com/ngrams>
ELAN – <http://tla.mpi.nl/tools/tla-tools/elan/>
FLEX: Language Explorer – <http://fieldworks.sil.org/flex/>

7. МЕТОДИЧЕСКИЕ УКАЗАНИЯ ПО ОСВОЕНИЮ ДИСЦИПЛИНЫ

1. Методические рекомендации по организации самостоятельной работы обучающихся.

8. ИНФОРМАЦИОННЫЕ ТЕХНОЛОГИИ ДЛЯ ОСУЩЕСТВЛЕНИЯ ОБРАЗОВАТЕЛЬНОГО ПРОЦЕССА ПО ДИСЦИПЛИНЕ

Лицензионное программное обеспечение:

Microsoft Windows
Microsoft Office
Kaspersky Endpoint Security

Информационные справочные системы:

Система ГАРАНТ
Система «Консультант Плюс»

Профессиональные базы данных:

fgosvo.ru – Портал Федеральных государственных образовательных стандартов высшего образования

pravo.gov.ru - Официальный интернет-портал правовой информации

www.edu.ru – Федеральный портал Российское образование

Свободно распространяемое программное обеспечение, в том числе отечественного производства

ОМС Плеер (для воспроизведения Электронных Учебных Модулей)
7-zip
Google Chrome

9. МАТЕРИАЛЬНО-ТЕХНИЧЕСКОЕ ОБЕСПЕЧЕНИЕ ДИСЦИПЛИНЫ

Материально-техническое обеспечение дисциплины включает в себя:

- учебные аудитории для проведения занятий лекционного и семинарского типа, групповых и индивидуальных консультаций, текущего контроля и промежуточной аттестации, укомплектованные учебной мебелью, доской, демонстрационным оборудованием;
- помещения для самостоятельной работы, укомплектованные учебной мебелью, персональными компьютерами с подключением к сети Интернет и обеспечением доступа к электронным библиотекам и в электронную информационно-образовательную среду .