

Министерство науки и высшего образования Российской Федерации
Федеральное государственное бюджетное образовательное
учреждение высшего образования
«Государственный институт русского языка им. А.С. Пушкина»
Филологический факультет
Кафедра общего и русского языкознания

РАБОЧАЯ ПРОГРАММА ДИСЦИПЛИНЫ
ТЕХНОЛОГИИ КОРПУСНОЙ ЛИНГВИСТИКИ

Направление подготовки 45.03.02 Лингвистика

Направленность (профиль) Отечественная и зарубежная лингвистика

Квалификация выпускника бакалавр

Форма обучения очная

Москва

Разработчик (и):

Лунькова Е. С., д. филол. н, доц., профессор кафедры ОиРЯ ФГБОУ ВО
«Государственный институт русского языка им. А. С. Пушкина»

1. Цель освоения дисциплины

Целью освоения дисциплины является освоение основных понятий корпусной лингвистики и формирование представлений об инструментах поиска в различных языковых корпусах в соответствии с поставленными учебными и исследовательскими задачами.

2. Место дисциплины в структуре ОПОП

В соответствии с учебным планом по направлению подготовки 45.03.01 - Филология, дисциплина Б1.В.ДВ.01.02 «Технологии корпусной лингвистики» относится к дисциплинам по выбору.

3. Планируемые результаты обучения по дисциплине, соотнесенные с планируемыми результатами освоения ОПОП

В процессе изучения дисциплины формируются следующие компетенции:

Код и наименование компетенции	Код и наименование индикатора достижения компетенции (ИДК)	Результаты обучения по дисциплине		
		Знать	уметь	владеть
ОПК-1. Способен использовать в профессиональной деятельности, в том числе педагогической, представление об истории, современном состоянии и перспективах развития филологии в целом и ее конкретной области с учетом направленности (профиля) образовательной программы	ОПК-1.3. Корректно интерпретирует различные явления филологии	<i>Предмет и задачи корпусной лингвистики, основные виды языковых корпусов; общие этапы и модули обработки текстов на естественных языках; международные стандарты корпусной лингвистики; основные технологии корпусной лингвистики.</i>	<i>Характеризовать этапы развития, современное состояние и перспективы развития корпусной лингвистики; строить корпуса специальных текстов.</i>	<i>Навыками вариативного поиска данных в языковых корпусах для формирования учебных знаний в различных предметных областях</i>
	ОПК-1.5 Имеет практический опыт работы с языковым и литературным материалом, научным наследием ученых-филологов, в том числе, в педагогической деятельности	<i>Основные положения прочитанных работ по подкорпусам национального русского языка и корпусам специальных текстов, а также сервисам корпусных менеджеров.</i>	<i>Корректно представлять параметры поисковых запросов в языковых корпусах специальных текстов и подбирать текстовый материал для решения соответствующих учебных задач.</i>	<i>Навыками создания различных учебных заданий для студентов-филологов по проблематике корпусной лингвистики</i>
ОПК-2. Способен использовать в	ОПК-2.1. Знает основные положения и	<i>Понятия "язык для специальных целей" и "специальный текст";</i>	<i>Характеризовать основные виды лингвистической</i>	<i>Навыками применения теоретических</i>

профессиональной деятельности, в том числе педагогической, основные положения и концепции в области общего языкознания, теории и истории основного изучаемого языка (языков), теории коммуникации	концепции в области общего языкознания, теории и истории основного изучаемого языка (языков), теории коммуникации, лингвистической терминологии, применяет их в профессиональной, в том числе педагогической деятельности	<i>основные понятия корпусной лингвистики (токенизация, лемматизация, таги, разметка, подкорпус, ирт, поиск точных форм, поиск коллокаций, ключ и коллокат; лексико-грамматический поиск);</i>	<i>разметки подкорпусов НКРЯ и корпусов специальных текстов.</i>	<i>знаний об устройстве языкового корпуса и специального языкового подкорпуса для использования основных алгоритмов поиска</i>
	ОПК-2.3 Осуществляет перевод и (или) интерпретацию текстов различных типов	<i>Основные модули обработки текстов на естественном языке: графематический анализ (сегментация и токенизация), морфологический и синтаксический анализ, семантический и прагматический анализ.</i>	<i>Применять основные понятия корпусной лингвистики к оценке языкового материала; верифицировать извлеченный из корпуса языковой и текстовый материал.</i>	<i>Навыками лингвистического анализа полученных из языкового корпуса данных в прикладном аспекте</i>

4. Объем, структура и содержание дисциплины

4.1. Трудоемкость дисциплины

Вид учебной работы	очная форма обучения
	7 семестр
Общая трудоемкость по учебному плану (з.е.)	3
Количество часов по учебному плану (час.)	108
Контактная работа (всего)	36
В том числе:	
Лекции	24
Практические занятия	12
Лабораторные работы	
Самостоятельная работа (всего)	36
Курсовая работа (для ОПОП бакалавриата)	Не предусмотрена
Форма промежуточной аттестации – зачет	36

4.2. Структура и содержание дисциплины

Наименование и краткое содержание разделов дисциплины	Количество часов по учебному плану	Контактная работа (часы), из них		Самостоятельная работа	Форма промежуточной аттестации	Формы текущего контроля успеваемости
		Лекции	Практические занятия			
	Очная форма	Очная форма	Очная форма	Очная форма	Очная форма	Очная форма
1. Корпусная лингвистика как научная дисциплина. Национальный корпус русского языка. Корпусы специальных текстов. Корпусная лингвистика и ее понятийный аппарат. История корпусной лингвистики. Виды языковых корпусов. Варьирование объема и баланса корпуса. НКРЯ.	10	4	-	6		Т
2. Создание специальных корпусов текстов. Разметка корпусов, стандартизация в корпусной лингвистике. Настройка пользовательских подкорпусов в НКРЯ для исследовательских целей: газетный, акцентологический, устный, диалектный, мультимедийный, параллельный, поэтический подкорпуса. Панхронический корпус. Поиск по грамматической и метатекстовой разметке в НКРЯ. Функции "График", "Статистика", "Частотность" в НКРЯ. Специальные корпуса текстов на русском языке.	18	4	4	10		Д, УО
3. Сервис корпусного менеджера. Формирование поисковых запросов разного типа в НКРЯ: семантическая, грамматическая и метатекстовая разметка. Оптимизация параметров морфосинтаксической разметки Программа AntConc: поиск текстовой информации на русском и иностранных языках, функция REGEX: вариативные запросы поиска. Сервис NoSketchEngine.	44	16	8	20		Д, Т, Э
Форма промежуточной аттестации – зачет	36				36	
Итого	108	24	12	36	36	

Примечание:

Условные обозначения форм контроля:

УО – устный опрос; Т – тест; Д – доклад с презентацией; Э – эссе

5. Образовательные технологии

При реализации дисциплины «Основы корпусной лингвистики» применяются следующие образовательные технологии:

- ☐ лекции;
- ☐ практические занятия;
- ☐ доклады;
- ☐ эссе;
- ☐ письменные домашние работы по извлечению из корпуса, анализу и интерпретации данных;
- ☐ тестирование;

Самостоятельная работа охватывает проработку обучающимися отдельных вопросов теоретического курса, выполнение домашних работ, включающих

- ☐ проблемное обучение (эссе);
- ☐ подготовку докладов и презентаций к ним;
- ☐ выполнение письменных домашних работ;
- ☐ работа с учебной и научной литературой.

6. Учебно-методическое обеспечение самостоятельной работы обучающихся

6.1. Список учебной литературы

Основная литература:

1. *Шунейко, А. А.* Корпусная лингвистика: учебник для вузов / А. А. Шунейко. — Москва: Издательство Юрайт, 2026. — 222 с. — (Высшее образование). — ISBN 978-5-534-13603-6. — Текст: электронный // Образовательная платформа Юрайт [сайт]. — URL: <https://urait.ru/bcode/588309> (дата обращения: 04.07.2026).

Дополнительная литература:

1. Захаров В. И., Богданова С. Ю. Корпусная лингвистика: учебник. 3-е изд., перераб. СПб.: Изд-во С.-Петерб.ун-та, 2020. - 234с. ISBN: 978-5-288-05997-1
2. Моделирование в корпусной лингвистике: специализированные корпуса русского языка / В.П. Захаров, И.В. Азарова, О.А. Митрофанова, А.М. Попов, М.В. Хохлова; отв.ред. В.П. Захаров. — СПб.: Изд-во С.-Петерб.ун-та, 2019. - 208с. — ISBN: 978-5-288-05902-5

6.2. Словари и справочники:

1. Энциклопедия «Русский язык»/Институт русского языка им. В.В. Виноградова РАН, Гл. ред. А.М. Молдован. — 3-е изд., переработанное и дополненное. — М.: АСТ-ПРЕСС ШКОЛА, 2020. — 904 с. — ISBN: 978-5-907126-47-3

6.3. Периодические издания:

1. Вопросы языкознания: Научный журнал. - <https://vja.ruslang.ru/ru/archive> (Архив журнала с открытым доступом в Интернет).

6.4. Ресурсы информационно-телекоммуникационной сети «Интернет»

Ресурсы сети Интернет, необходимые для изучения дисциплины:

Наименование	Ссылка на ресурс	Доступность (свободный доступ / ограниченный доступ)
1	2	3
Информационно-справочные системы		
Библиографические базы данных по общественным наукам ИНИОН	http://inion.ru/	Свободный доступ
Методология: словарь системы	http://www.anovikov.ru/dict/met_sl.htm	Свободный доступ

основных понятий. Сайт академика РАО Новикова А.М.		
Электронно-библиотечные системы		
ЭБС Юрайт	https://urait.ru	После регистрации с компьютера Института – доступ с любого компьютера, подключенного к сети Интернет.
Профессиональные базы данных		
Справочно-информационный портал ГРАМОТА.РУ – русский язык для всех	http://gramota.ru/	Свободный доступ
«ИНФРА–М»	http://znanium.com	Свободный доступ
BOOK.ru	https://www.book.ru	После регистрации с компьютера Института – доступ с любого компьютера, подключенного к сети Интернет
Научная электронная библиотека РИНЦ	https://elibrary.ru/	Свободный доступ
Научная электронная библиотека «КиберЛенинка»	https://cyberleninka.ru/	Свободный доступ

6.5. Содержание самостоятельной работы

РАЗДЕЛ 1. 1. Корпусная лингвистика как научная дисциплина. Национальный корпус русского языка. Корпусы специальных текстов.

Вопросы для самостоятельной работы:

1. Каковы предмет и объект корпусной лингвистики?
2. Расскажите об основных этапах развития корпусной лингвистики.
3. Каковы взаимоотношения корпусной лингвистики с другими лингвистическими дисциплинами?
4. Охарактеризуйте НКРЯ как корпус текстов национального языка.
5. Опишите создание специальных корпусов как многокритериальную теоретическую задачу.

Задания для самостоятельной работы:

1. Определите пределы варьирования объема и баланса избранного для анализа корпуса специальных текстов, его тематические и жанровые характеристики. Опишите структуру и текстовые единицы избранного корпуса.

РАЗДЕЛ 2 Создание специальных корпусов текстов.

Вопросы для самостоятельной работы:

1. Опишите варьирование основных параметров применительно к различным типам специальных корпусов текстов.
2. Охарактеризуйте корпус терминологических текстов.
3. Приведите примеры возможных поисковых запросов для выбранного корпуса.

Задания для самостоятельной работы

1. Домашняя работа №1 «Анализ результатов поискового запроса в корпусе терминологических текстов»
2. Домашняя работа №2 «Специфика использования параллельных корпусов русского языка (на примере поисковых запросов в одном из специальных параллельных подкорпусов)».

РАЗДЕЛ 3. Сервис корпусного менеджера.

Вопросы для самостоятельной работы:

1. Перечислите и охарактеризуйте форматы лингвистической разметки в языковых корпусах.
2. Опишите систему грамматических классов и подклассов слов в русском языке применительно к задачам корпусной лингвистики.
3. Охарактеризуйте возможности синтаксической аннотации в корпусах текстов.
4. Охарактеризуйте понятие n-граммы, объясните их отличие от коллокаций.
5. Что такое ключ и коллокат в поиске коллокаций? Как осуществляется поиск опосредованных коллокаций в НКРЯ?
6. Охарактеризуйте формальную модель выделения терминов и терминологических сочетаний из корпусов.
7. Опишите автоматическое извлечение ключевых слов и словосочетаний в корпусах специальных текстов.

Задания для самостоятельной работы

Домашняя работа №3. Подберите примеры 50 фрагментов терминологических текстов для избранного слова-термина с IPM не ниже 10. Сравните результаты с аналогичными данными в основном подкорпусе НКРЯ.

Домашняя работа №4. Проанализируйте тексты в избранном корпусе НКРЯ с использованием статистических методов и данных

сравнительного корпуса.

Домашняя работа №5. Работа в системе NoSketchEngine (*AntConc): проектирование и технологический процесс создания корпуса.

Дополнительное задание (на дополнительные баллы)

1. Создайте и загрузите в систему NoSketchEngine (*AntConc) пользовательский подкорпус, проведите в нем анализ выбранных слов и сочетаний: а) слово 1 (не менее 100 контекстов); б) слово 2; в)..., по 1 баллу за каждое проанализированное и описанное слово.

7. Оценочные средства для проведения текущего контроля успеваемости и промежуточной аттестации обучающихся

Оценка качества учебных достижений обучающихся по дисциплине осуществляется в виде текущего контроля успеваемости и промежуточной аттестации.

7.1. Текущий контроль успеваемости

Текущий контроль успеваемости учитывает следующие виды текущей аттестации: выполнение обучающимся всех видов контрольных мероприятий, предусмотренных рабочей программой дисциплины, и самостоятельную работу обучающегося. Все формы текущего контроля и активность участия студентов в практических и (или) лабораторных занятиях, критерии их оценивания представлены в рабочей программе дисциплины, технологической карте и фонде оценочных материалов (Приложения 1-2).

7.1.1. Тематика и требования к структуре и содержанию докладов

Темы докладов

1. Принципы построения и разметки (аннотации) лингвистических корпусов. Сравнение морфологической, синтаксической и семантической разметки на примере Национального корпуса русского языка (НКРЯ).

2. Корпусный менеджер Sketch Engine в практике лексикографа: автоматическое составление словарей частотности, коллокаций и тезаурусов.

3. Сравнительный анализ параллельных корпусов для перевода: использование ресурса OPUS или русско-английского подкорпуса НКРЯ для выявления переводческих соответствий и безэквивалентной лексики.

4. Динамика функционирования русской лексики по данным газетного корпуса: неологизмы 2015–2026 годов (например, *удаленка*, *ковид-диссидент*, *нейросетка*) и уходящие слова (*пейджер*, *факс*).

5. Автоматическая стилистическая атрибуция текстов с помощью машинного обучения: определение авторства анонимных произведений XIX века методами стилометрии.

6. Методы извлечения терминологии из узкоспециализированных текстовых коллекций (на материале медицинских или юридических текстов). Применение статистических мер MI-score и T-score.

7. Корпусные методы исследования грамматической синонимии: употребление конструкций «несмотря на» / «невзирая на» / «вопреки» в современном русском языке.

8. Создание пользовательских учебных корпусов: методика отбора текстов и разработка аннотации для преподавания русского языка как иностранного (РКИ).

9. Диахронический анализ поэтического синтаксиса: эволюция ритмико-синтаксических структур от поэзии А.С. Пушкина до И.А. Бродского (можно использовать других авторов) на основе поэтического подкорпуса НКРЯ.

10. Извлечение мнений и сентимент-анализ русскоязычных текстов: проблемы автоматической классификации тональности (сарказм, ирония, контекстуальная полярность) на данных социальных сетей.

11. Корпус устной речи vs корпус письменных текстов: фонетические редукции, хезитации, самокоррекция и специфика диалогической разметки (на примере устного корпуса русского языка).

12. Лингвистический ландшафт города в зеркале корпусных технологий: сбор и количественный анализ микротекстов городской среды (вывесок, объявлений, граффити).

13. Применение дистрибутивной семантики и векторных моделей (Word2Vec, FastText) для анализа семантических полей: выявление скрытых смысловых связей между словами в больших массивах текста.

14. Критерии репрезентативности и сбалансированности корпуса: проблема создания малого специализированного корпуса (от 1 до 5 миллионов словоупотреблений) и экстраполяция его данных на весь язык.

15. Эволюция значений многозначных слов (полисемия): диахроническое исследование семантических переходов глаголов движения (*идти*, *вести*) на исторических срезах НКРЯ.

Требования к структуре и содержанию доклада:

- 1) регламент выступления - 10 минут; наличие текста доклада и компьютерной презентации (презентация и текст доклада не должны дублировать друг друга);
- 2) использование самостоятельно подобранного текстового материала и его анализа в аспекте, определяемом темой доклада;
- 3) наличие учебного задания для студентов группы по теме доклада.

Критерии оценки доклада

Критерии	Показатели
Наличие необходимых элементов	Есть письменный текст доклада, компьютерная презентация, не дублирующие друг друга, студент выступил с докладом
Соблюдение требований к структуре и содержанию доклада	В докладе используется самостоятельно подобранный текстовый материал, есть его анализ и учебное задание

7.1.2. Тематика и требования к структуре и содержанию эссе

Примерная тематика эссе:

Языковые корпуса: сбалансированность и репрезентативность.

Поиск в корпусе: семантика, морфология, синтаксис.

Метатекстовая разметка в языковых корпусах.

Корпуса устной и письменной речи.

Диалектные корпуса русского языка.

Интернет-корпуса славянских языков.

Интернет-корпуса германских языков.

Эссе по книге Захарова В. И., Богдановой С. Ю. Корпусная лингвистика (СПб., 2020).

Требования к структуре и содержанию эссе:

- 1) титульный лист;
- 2) текст эссе (не менее 10 с. формата А4, 14 шрифт через 1,5 интервала);

- 3) наличие самостоятельно подобранных иллюстраций
- 4) наличие анализа прочитанного и выводов.

Критерии оценки эссе

Критерии	Показатели
Корректность оформления	Наличие титульного листа, грамотность, соблюдение требований к оформлению, соблюдение объема
Полноценность содержания	Наличие анализа прочитанного, собственной позиции, ее аргументации и выводов.
Иллюстративный материал	Наличие самостоятельно выбранного текстового материала

7.1.3. Тематика и требования к структуре и содержанию рефератов

Не предусмотрены.

7.1.4. Требования к выполнению контрольной работы

Не предусмотрено.

7.1.5. Требования к выполнению тестов

Примерные вопросы к тесту № 1:

1. Продолжите высказывание, выбрав наиболее точное определение понятия.

1. Какое из высказываний является определением прикладной лингвистики?

- a) область языкознания, направленная на объективное установление состояния отдельного языка, его истории и закономерностей;
- b) область языкознания, связанная с использованием компьютерных инструментов — программ, технологий организации и обработки данных — для моделирования функционирования языка в тех или иных условиях;
- c) область языкознания, связанная с разработкой методов решения практических задач использования языка;
- d) область языкознания, связанная с применением компьютерных моделей языка в лингвистике и в смежных с ней дисциплинах.

2. К направлениям компьютерной лингвистики не относится

- a) компьютерная лексикография;
- b) компьютерно-опосредованная коммуникация;
- c) системы обработки естественного языка;
- d) машинный перевод.

3. Информатика — это

- a) наука об управлении, связи и переработке информации;
- b) наука о накоплении, обработке и передаче информации с помощью ЭВМ;
- c) наука о накоплении, обработке и передаче информации о строении языка с помощью ЭВМ;
- d) наука об использовании компьютерных инструментов для моделирования

функционирования языка в тех или иных условиях.

4. Разное количество информации в одном и том же сообщении для разных людей зависит не от...

- a) накопленных ими знаний;
- b) уровня понимания сообщения;
- c) их интереса к сообщению;
- d) их уровня владения компьютерной техникой.

Следствие третьей информационной революции состоит в том, что...

- a) информация становится общедоступной;
- b) информацию можно автоматически обрабатывать и передавать с большой скоростью;
- c) информацию можно легко найти с помощью инструментов поиска и совместно производить;
- d) информация может накапливаться.

Для современного человека преобладающей является...

- a) звуковая информация;
- b) визуальная (символьная) информация;
- c) вкусовая и тактильная информация;
- d) визуальная (образная) информация.

Адекватность информации — это ...

- a) степень соответствия информации объективной реальности окружающего мира;
- b) степень соответствия информации, полученной потребителем, тому, что автор вложил в ее содержание;
- c) достаточность информации для принятия решения;
- d) степень соответствия информации текущему моменту времени.

Машинный синтаксис — это ...

- a) правила строения имен;
- b) правила построения слов в более сложные структуры;
- c) соотношение слова и его значения;
- d) правила перевода письменного символа в устный.

Естественный язык — это ...

- a) знаковая система, используемая человеком с момента рождения;
- b) знаковая система, используемая человеком в непринужденной обстановке;
- c) знаковая система, созданная для естественных наук;
- d) знаковая система, стихийно возникшая и закрепившаяся в обществе.

10. Волапюк — это...

- a) специализированный язык науки;
- b) родной язык одного из малочисленных племен;
- c) неспециализированный искусственный язык;
- d) система символического кодирования.

11. Какие из следующих приложений не являются текстовыми редакторами?

- a) MS Excel;
- b) Corel WordPerfect;

- c) MS Works;
- d) Adobe InCopy.

12. Microsoft Word не включает...

- a) функции настольных издательских систем;
- b) функцию удалённого доступа;
- c) функцию редактирования графических объектов;
- d) шаблоны типовых таблиц.

13. К устройствам ввода данных не относится

- a) сканер;
- b) принтер;
- c) клавиатура;
- d) цифровой фотоаппарат.

14. OCR —это ...

- a) система автоматического распознавания символов;
- b) система переводческой памяти;
- c) система машинного перевода;
- d) функция текстового процессора.

15. Реферат — это...

- a) связный текст, который кратко выражает тему, предмет, цель, методы и результаты исследования;
- b) процесс составления содержания документа (книги, статьи, патента на изобретение и др.);
- c) краткое изложение содержания документа, дающее общее представление о его теме;
- d) краткий текст, выполняющий сигнальную функцию (информирует о том, что есть публикация на определенную тему).

16. Слово, относящееся к основному содержанию текста и повторяющееся в нем несколько раз, в автоматическом реферировании называется ...

- a) лейтмотивом;
- b) термином;
- c) символом;
- d) ключевым словом.

17. Метод автоматического аннотирования, при котором важные слова выделяются в заголовке, подзаголовке, начале и конце текста, называется ...

- a) статистическим;
- b) логико-семантическим;
- c) позиционным;
- d) функциональным.

18. Совокупность специально отобранных текстов, размеченных по различным лингвистическим параметрам и обеспеченных системой поиска, называется...

- a) базой данных;
- b) словарем;
- c) информационным массивом;
- d) корпусом.

19. Разметка бывает ...

- a) морфологической; синтаксической; семантической и просодической;
- b) полнотекстовой и фрагментной;
- c) синхронической и диахронической;
- d) звуковой, письменной, смешанной.

20. УНК —это...

- a) корпус естественного языка, представительный по отношению ко всему языку;
- b) универсальный национальный код;
- c) собрание текстов, которое существует в Интернете;
- d) собрание текстов, размеченных по различным лингвистическим параметрам и обеспеченных системой поиска.

21. Требования к корпусам

- a) полнота, адекватность, актуальность, компьютерная поддержка;
- b) устойчивость, тиражируемость, адаптируемость, оптимальность временных параметров, комфорт пользователя;
- c) репрезентативность, полнота, экономичность, структуризация, компьютерная поддержка;
- d) полнота, экономичность, достоверность, структуризация, компьютерная поддержка.

22. Корпусный менеджер ...

- a) обеспечивает сортировку результатов поиска, статистические подсчеты, составление списков слов на основе корпуса;
- b) это специальная программа поиска по корпусу;
- c) это человек, составляющий корпуса и управляющий ими;
- d) это специальная программа подготовки текстов к их включению в корпус.

23. ПОД —это ...

- a) вид информационно-поисковой системы;
- b) специальная программа поиска по корпусу;
- c) поисковый образ документа;
- d) поисковая оценка данных.

24. Одна из основных проблем компьютерного анализа речи состоит в том, что ...

- a) невозможно создать искусственный интеллект;
- b) компьютер не умеет работать со смыслом;
- c) у компьютера нет дополнительных источников информации (ситуация, контекст, прошлый опыт в данной области и т.п.);
- d) разработчики не желают делиться своими профессиональными секретами.

25. Электронный словарь — это ...

- a) введенный в компьютер бумажный словарь, снабженный средствами поиска и отображения информации;
- b) организованное собрание слов с комментариями, в которых описываются особенности структуры и/или функционирования этих слов;

- с) организованное собрание слов с описанием их значения, особенностей употребления, структурных свойств, сочетаемости, соотношения с лексическими системами других языков и т.д.;
- д) словарь в специальном машинном формате, предназначенный для применения на ЭВМ пользователем или компьютерной программой.
26. К зонам словарной статьи не относится
- а) лексический вход (вокабула, лемма);
- б) зона грамматической информации;
- с) зона стилистических помет;
- д) словник.
27. Что включает в себя понятие АСПОТ?
- а) словарь в специальном машинном формате, предназначенный для применения на ЭВМ пользователем;
- б) компьютерные версии хорошо известных словарей (Вебстер, Коллинз, Ожегов...);
- с) словарь в специальном машинном формате, предназначенный для применения на ЭВМ компьютерной программой;
- д) словари, предназначенные для обычного пользователя.
28. Что не относится к понятию термина?
- а) слово (словосочетание) метаязыка науки, а также областей конкретной практической деятельности человека;
- б) понятие задается через свойства, реализуемые в системе;
- с) использование основывается не на интуиции, а на четких определениях;
- д) сопоставляется, как правило, несколько значений.
29. Что не относится к процессу и понятию машинного перевода?
- а) междисциплинарность;
- б) использование машинных средств;
- с) принципиальное сходство этапов понимания и синтеза текста;
- д) учет языковых и экстралингвистических знаний.
30. Типовая парадигма лексемы в автоматическом морфологическом анализе — это ...
- а) последовательность букв от начала словоформы, общая для всех словоформ;
- б) элементы, описывающие формоизменение конкретной лексемы;
- с) совокупность наборов машинных окончаний;
- д) совпадение основ разных слов.
31. Требования к системам МП включают ...
- а) устойчивость, тиражируемость, адаптируемость, оптимальность временных параметров, комфорт пользователя;
- б) полнота, адекватность, актуальность, достоверность;
- с) репрезентативность, полнота, экономичность, адекватность, компьютерная поддержка;
- д) репрезентативность, полнота, экономичность, структуризация, компьютерная поддержка.
32. Аббревиатура CALL относится к ...

- а) науке об использовании компьютерных инструментов для моделирования функционирования языка в тех или иных условиях;
- б) обучению иностранному языку;
- с) обучению языку с помощью компьютера;
- д) использованию компьютеров в обучении.

7.1.6. Требования к подготовке проекта

Не предусмотрен.

7.1.7. Требования к подготовке курсовой работы

Не предусмотрена.

7.2. Промежуточная аттестация

Промежуточная аттестация по дисциплине обеспечивает оценку степени полноты и качества освоения компетенций в соответствии с планируемыми результатами обучения по дисциплине.

7.2.1. Оценивание результатов освоения образовательной программы в части дисциплин (модулей), практик

Результаты освоения дисциплины оцениваются по шкале «зачтено» - «не зачтено (зачет)».

Текущий контроль, промежуточная аттестация осуществляются в соответствии с локальным нормативным актом Института.

Оценочные материалы для проведения промежуточной аттестации обучающихся приведены в Приложении 2 к рабочей программе дисциплины.

7.2.2. Примерные вопросы для подготовки к зачету:

1. История корпусной лингвистики.
2. Виды языковых корпусов.
3. Токенизация и лемматизация в языковом корпусе.
4. Национальный корпус русского языка: история создания, принципы организации, виды подкорпусов. Сбалансированность и репрезентативность как главные характеристики корпусов национальных языков.
5. Корпусы специальных текстов.
6. Сервис корпусного менеджера.
7. Форматы лингвистической разметки.
8. Программные средства морфологической разметки.
9. Сервис NoSketchEngine.
10. Программа AntConc, функции RegEx.

7.2.3. Примерные практические задания к зачету:

1. Составьте подкорпусы в рамках НКРЯ на материалах газетных текстов и сформулируйте в них вариативные поисковые запросы.
2. Постройте графики частотности для пяти слов в современном русском языке, интерпретируйте результаты (по материалам НКРЯ).

3. Рассчитайте ipm для пяти терминологических глаголов в русском языке (по материалам НКРЯ).
4. Выясните, какие существительные, прилагательные, глаголы, наречия и т.д.) наиболее частотны в русском языке.
5. Составьте корпус произведений какого-либо автора на русском или иностранном языках с помощью специальной программы. Выясните, какие слова являются в нем ключевыми.
6. Найдите все существительные (прилагательные, глаголы и др.) в специальном корпусе текстов.
7. Составьте автоматизированную аннотацию заданного текста.

На экзамене можно получить максимальный балл – 15, из них:

- ответ на теоретические вопросы – 8 баллов (по 4 на каждый из 2-х вопросов)
- практическое задание – 7 баллов.

7.2.4. Критерии и шкала результатов обучения по дисциплине

Критерии оценивания результатов обучения по дисциплине при проведении промежуточной аттестации:

Планируемые результаты обучения по дисциплине	Критерии оценивания результатов обучения по дисциплине			
	«не зачтено»	«зачтено»		
Сформированность знаний (знать)	Уровень знаний ниже минимальных требований. Имели место грубые ошибки. Невозможно оценить полноту знаний вследствие отказа обучающегося от ответа.	Минимально допустимый уровень знаний. Допущено много негрубых ошибок.	Уровень знаний в объеме, соответствующем программе подготовки. Допущено несколько незначительных ошибок.	Уровень знаний в объеме, соответствующем программе подготовки, или в объеме, превышающем программу подготовки, продемонстрирован без ошибок и погрешностей.
Сформированность умений (уметь)	Отсутствие минимальных умений. При решении стандартных задач не продемонстрированы основные умения. Имели место грубые ошибки. Невозможно оценить наличие умений вследствие отказа обучающегося от ответа.	Продemonстрированы основные умения. Решены типовые задачи с негрубыми ошибками. Выполнены все задания, но не в полном объеме.	Продemonстрированы все основные умения. Решены все основные задачи, но некоторые с негрубыми ошибками. Выполнены все задания, в полном объеме, но некоторые с недочетами.	Продemonстрированы все основные умения. Решены все основные задачи. Выполнены все задания, в полном объеме без недочетов.
С ф о р м	При решении стандартных задач	Продemonстрированы минимальный	Продemonстрированы базовые навыки.	Продemonстрировано всестороннее

	не продемонстрированы базовые навыки. Имели место грубые ошибки. Невозможно оценить наличие навыков вследствие отказа обучающегося от ответа.	набор навыков. Решены стандартные задачи с некоторыми недочетами.	Решены стандартные задачи, но некоторые с недочетами.	владение навыками. Задачи, в том числе нестандартные, решены без ошибок и недочетов.
--	---	--	---	---

8. Материально-техническое обеспечение дисциплины

Для реализации рабочей программы дисциплины имеются помещения для проведения занятий лекционного и семинарского типа, текущего контроля и промежуточной аттестации, а также помещения для самостоятельной работы. Помещения представляют собой учебные аудитории для проведения учебных занятий, оснащенные оборудованием и техническими средствами обучения, служащими для представления информации большой аудитории.

Помещение для проведения занятий лекционного типа поточная аудитория на 100 посадочных мест укомплектовано: специализированной мебелью посадочные места, столы, доска, компьютер, интерактивная доска,

Помещение для проведения занятий семинарского типа, текущего контроля и промежуточной аттестации укомплектовано: специализированной мебелью, компьютером, интерактивной доской, мультимедиа проектором, звуковой системой.

Помещение для самостоятельной работы оснащено компьютерной техникой с возможностью подключения к сети «Интернет» и обеспечением доступа к электронной информационно-образовательной среде Института.

Институт обеспечен следующим необходимым комплектом лицензионного и свободно распространяемого программного обеспечения, в том числе отечественного производства:

1. Microsoft Windows 10,
2. Microsoft Office 2016 стандартная,
3. Big Blue Button.