

Документ подписан простой электронной подписью  
Информация о владельце:  
ФИО: Романчук Иван Сергеевич  
Должность: Ректор  
Дата подписания: 30.01.2025 11:28:45  
Уникальный программный ключ:  
6319edc2b582ffdacea443f01d5779368d0957ac34f5cd074d81181530452479

Приложение к рабочей  
программе дисциплины

## МЕТОДИЧЕСКИЕ РЕКОМЕНДАЦИИ ПО ОРГАНИЗАЦИИ САМОСТОЯТЕЛЬНОЙ РАБОТЫ ОБУЧАЮЩИХСЯ

Наименование дисциплины	<i>Корпусная лингвистика для переводчиков</i>
Направление подготовки	<i>45.04.02 Лингвистика</i>
Направленность (профиль)	<i>Прикладная лингвистика</i>
Форма обучения	<i>очная</i>
Разработчик(и)	<i>Ковязина М.А., доцент кафедры прикладной и теоретической лингвистики</i>

1. Темы дисциплины для самостоятельного освоения обучающимися  
Отсутствуют

2. План самостоятельной работы

№ п/п	Учебные встречи	Виды самостоятельной работы	Форма отчетности/ контроля	Количество баллов	Рекомендуемый бюджет времени на выполнение (ак.ч.)*
1	2	3	4	5	6
1	УВ №2-3, Практические занятия 1-2, «Введение в корпусную лингвистику»	<p><b>1.</b> Прочитайте Главу 1 в учебном пособии А. А. Шунейко «Корпусная лингвистика» (2024) на образовательной платформе Юрайт (<a href="https://urait.ru/bcode/543746">https://urait.ru/bcode/543746</a>).</p> <p><b>1.1.</b> Ознакомьтесь с разделом 1.1. «Поиск имени» и ответьте на вопросы:</p> <p>а) На чём базируется методология корпусной лингвистики?</p> <p>б) С какими науками корпусная лингвистика тесно связана?</p> <p>в) Как на различных этапах своего развития называлась корпусная лингвистика?</p> <p>г) Каковы основные задачи корпусной лингвистики?</p> <p><b>1.2.</b> Ознакомьтесь с разделом 1.2. «Два типа учёных и познавательных моделей» и ответьте на вопросы:</p> <p>а) Можно ли корпусную лингвистику считать самостоятельным методом исследования?</p> <p>б) Какие теоретические и прикладные вопросы позволяет решать корпусная лингвистика?</p> <p>в) Какие существуют направления в корпусной лингвистике?</p> <p><b>1.3.</b> Ознакомьтесь с разделом 1.3. «Корпус и его окрестности» и ответьте на вопросы:</p> <p>а) Что послужило основой корпусной лингвистики в плане теории?</p> <p>б) Что и как стало основой корпусной лингвистики в плане технологии?</p> <p>в) Когда был дан самый мощный импульс развитию корпусной лингвистики?</p> <p>г) Какие научные направления и поиски</p>	Ключевые слова, ответы на поставленные вопросы	2	8

		сформировали корпусную лингвистику как направление научного знания?			
2	УВ №4, Практическое занятие 3, «Понятие корпуса. Типы корпусов»	<p>1. Посмотрите видео академика РАН В. А. Плуныяна о корпусной лингвистике и ответьте на вопросы:</p> <ol style="list-style-type: none"> <li>1) Что такое корпус?</li> <li>2) Что лингвист делает с корпусом?</li> <li>3) Корпусы какого объема появились в современной лингвистике?</li> <li>4) Почему лингвисты полюбили корпусы?</li> <li>5) Когда появились первые корпусы?</li> <li>6) Корпусы какого объема появились в 90-е гг.?</li> <li>7) В каких неанглоязычных странах в 90-е гг. интенсивно развивалась корпусная лингвистика?</li> <li>8) Какой проект предвосхитил создание Национального корпуса русского языка (НКРЯ)?</li> <li>9) Когда НКРЯ был открыт для свободного доступа?</li> <li>10) Кто составляет целевую аудиторию НКРЯ?</li> <li>11) Каков объем НКРЯ? Какие тексты входят в НКРЯ?</li> <li>12) Существует ли корпус устного русского языка?</li> </ol> <p>2. Посмотрите видео разработчика корпусной платформы <i>English-corpora.org</i> Марка Дэвиса об основных инструментах и возможностях, предоставляемых <i>Корпусом современного американского английского языка (Corpus of Contemporary American English)</i> и другими корпусами на данном сайте: “English-Corpora.org: introduction”. Подготовьте конспект о следующих аспектах:</p> <ul style="list-style-type: none"> <li>– элементы интерфейса платформы;</li> <li>– целевая аудитория платформы;</li> <li>– основные виды запросов;</li> <li>– инструменты Корпуса современного американского английского языка;</li> </ul> <p>способы извлечения статистических данных на базе корпуса.</p>	Ключевые тезисы, ответы на вопросы	4	8
			Текст конспекта		

3	УВ №6-8, Практические занятия 4-6, «Аннотирование корпуса»	<p>1. Прочитайте раздел 1.4. «Разметка корпуса в контексте его создания» в учебном пособии А. А. Шунейко «Корпусная лингвистика» (2024) на образовательной платформе Юрайт (<a href="https://urait.ru/bcode/543746">https://urait.ru/bcode/543746</a>). Ответьте на вопросы:</p> <p>а) Что такое разметка или аннотирование текста?</p> <p>б) Для чего осуществляется разметка текста?</p> <p>в) Как осуществляется разметка текста?</p> <p>г) Зависит ли качество корпуса от качества и сложности разметки? Приведите примеры.</p> <p>2. Посмотрите видео профессора Университета Васэда (Токио, Япония) Лоренса Энтони “TagAnt 1.0 - Getting started” о приложении, позволяющем осуществлять частеречную (или морфологическую) разметку текста на английском языке. Зафиксируйте в письменной форме основные операции, необходимые для получения морфологически размеченного текста.</p>	<p>Ответы на вопросы</p> <p>Основные тезисы, конспект</p>	2	8
4	УВ №10-12, Практические занятия 7-9, «Методология работы с корпус-менеджером»	<p>Посмотрите обучающие видео профессора Университета Васэда (Токио, Япония) Лоренса Энтони о корпусном менеджере AntConc:</p> <ol style="list-style-type: none"> <li>1) “AntConc 4 (ver. 4.2) - Getting started”</li> <li>2) “AntConc 4 (ver. 4.2) - Corpus Manager Basics”</li> <li>3) “AntConc 4 (ver. 4.0) - Word list tool basics”</li> <li>4) “AntConc 4 (ver. 4.0) - Keyword list tool basics”</li> <li>5) “AntConc 4 (ver. 4.0) - Collocate tool basics”</li> </ol> <p>Подготовьте конспект об инструментах данного приложения, позволяющих получать данные о наиболее частотных единицах, ключевых словах и сочетаемости.</p>	Основные тезисы, конспект	4	8
5	УВ №14-15, Практические занятия 10-11, «Двуязычные и многоязычные корпуса»	<p>Ознакомьтесь с разделами 1.7. «Когда и как родилась корпусная лингвистика» и 1.9. «Основные корпуса второго и третьего поколения». Ответьте на вопросы:</p> <p>а) Как называется корпус, в котором содержится только один тип</p>	Ключевые слова, ответы на вопросы	2	6

		<p>языковых единиц?</p> <p>б) Как называется корпус, в котором представлено несколько языков?</p> <p>в) Когда и где был разработан первый языковой корпус?</p> <p>г) Какие возможности для лингвистики дало создание первого корпуса?</p> <p>Какой корпус помимо письменной части включает разговорную часть, состоящую из неформальных бесед и разговорного языка?</p>			
6	УВ №17, Практическое занятие 12, «Создание собственного корпуса»	<p>Ознакомьтесь с требованиями к корпусному проекту, который разрабатывается студентами индивидуально или в парах и выносится на защиту на завершающем занятии. Студентам предлагается разработать одноязычный или двуязычный корпус не менее 10 тысяч словоупотреблений и на его базе провести лингвистическое исследование.</p>	Проект языкового корпуса	2	10
7	УВ №18, Практическое занятие 13, «Защита корпусных проектов»	<p>Подготовка к защите проекта корпусного исследования: студентам (индивидуально или в парах) предлагается разработать одноязычный или двуязычный корпус не менее 10 тысяч словоупотреблений и на его базе провести лингвистическое исследование.</p> <p>Проект состоит из следующих обязательных модулей:</p> <ul style="list-style-type: none"> <li>- обоснование корпуса, описание источников материала;</li> <li>- выявление базовых параметров корпуса (соотношение types/tokens, относительная частота и коллокации, составление конкорданса для некоторых слов, ключевые слова);</li> <li>- описание какой-либо лингвистической особенности корпуса, которая вскрылась в процессе выполнения проекта.</li> </ul>	Проект языкового корпуса: текст доклада (слайды презентации) и архивный файл корпуса zip или rar, включающий .docx или .pdf файл с названием корпуса и кратким описанием его параметров, а также собственно файл(ы) текста/текстов в формате *.txt.	20	16
8	Подготовка к занятиям	Повторение материала и изучение дополнительной литературы	Участие в обсуждениях		21
9	Подготовка к экзамену по дисциплине	Повторение изученного материала	Выполнение контрольных заданий		21
					<b>Итого: 106</b>

3. Требования и рекомендации по выполнению самостоятельных работ обучающихся, критерии оценивания

**УВ №17-18 Практическое занятие 12 «Создание собственного корпуса» и практическое занятие 13 «Защита корпусных проектов»:**

Студентам предлагается разработать одноязычный или двуязычный корпус не менее 10 тысяч словоупотреблений и на его базе провести лингвистическое исследование.

**Проект состоит из следующих обязательных модулей:**

- обоснование корпуса, описание источников материала;
- выявление базовых параметров корпуса (соотношение *types/tokens*, относительная частота и коллокации, составление конкорданса для некоторых слов, ключевые слова);
- описание какой-либо лингвистической особенности корпуса, которая вскрылась в процессе выполнения проекта.

Материалы проекта сдаются преподавателю в электронном виде: студент формирует архивный файл корпуса *zip* или *rar*, включающий *.docx* или *.pdf* файл с названием корпуса и кратким описанием его параметров, а также собственно файл(ы) текста/текстов в формате *.txt*, включенных в корпус. В названии архивного файла должна быть указана фамилия (или фамилии) исполнителя проекта; файл отправляется в соответствующее задание в онлайн-классе дисциплины (*lxp*).

Проект корпуса защищается на семинарском занятии. В ходе устной защиты студентам предлагается ответить на следующие вопросы:

1. Является ли корпус одноязычным / многоязычным (параллельным)?
2. Что послужило источником / источниками для корпуса?
  - укажите жанр(ы) текстов, включённых в корпус;
  - укажите ключевые слова (*seed words*), использованные при автоматическом формировании корпуса.
3. Цель создания корпуса?
4. Объем корпуса? (количество слов / токенов). Если корпус состоит из нескольких блоков (разные стили, хронологические периоды), можно указать процентное соотношение частей корпуса.
5. Список наиболее частотных единиц корпуса? (*Antconc > Word List*)
6. Список ключевых слов, выделенных на базе корпуса (*Keywords*)?
7. Коэффициент лексического разнообразия – *types/tokens ratio* (для корпуса в целом, либо для его частей).
8. Какая разметка применена в корпусе?

- *частеречная / POS – part of speech tagging, используемый набор меток (tagset)?*
  - *Free CLAWS web tagger <http://ucrel-api.lancaster.ac.uk/claws/free.html>*
  - *TagAnt <https://www.laurenceanthony.net/software/tagant/>*
- *семантическая (USAS, English online semantic tagger <http://ucrel-api.lancaster.ac.uk/usas/tagger.html>)*

**9.** *Результаты лингвистического исследования, проведённого на базе полученного корпуса (исследование сочетаемости, частотности определённых единиц, выделение синонимов, применение CQL при анализе функционирования единиц и пр. (AntConc > Concordance / Clusters & N-Grams / Collocates и пр.).*

**10.** *Каковы перспективы применения полученного корпуса?*

*Студент может получить 20 баллов максимум за корпусный проект, проведённое на его базе исследование и успешную защиту проекта. Баллы распределяются следующим образом:*

- *формулирование цели проекта (1 балл);*
- *использование корпусного менеджера для создания корпуса и проведения анализа (2 балла);*
- *подсчитан объем корпуса, выявлены наиболее частотные единицы (2 балла);*
- *выделены ключевые слова корпуса (1 балл);*
- *кратко описана применённая в корпусе разметка (1 балл);*
- *проведено небольшое лингвистическое исследование на базе корпуса, продемонстрированы его результаты (3 балла);*
- *защитное слово (4 балла max);*
- *наличие иллюстративных материалов, слайдов (3 балла max);*
- *отчетность в электронном виде (3 балла).*

**4.** *Рекомендации по самоподготовке к промежуточной аттестации по дисциплине*  
*Необходимое количество баллов для итоговой аттестации по дисциплине – 61-100.*

*Студенты, набравшие по итогам изучения дисциплины 61 балл и более, имеют право на аттестацию по итогам текущего контроля согласно балльно-рейтинговой шкале оценивания.*

*Студенты, не получившие необходимого количества баллов, сдают дифференцированный зачет, который представляет собой защиту корпусного проекта (требования описаны выше).*

