

УТВЕРЖДАЮ
 Декан факультета

 (подпись) Шматко А. Д.
 «___» _____ 20__
 ФИО

РАБОЧАЯ ПРОГРАММА ДИСЦИПЛИНЫ ТЕХНОЛОГИИ КОРПУСНОЙ ЛИНГВИСТИКИ

Направление/специальность подготовки	45.05.01 Перевод и переводоведение
Специализация/профиль/программа подготовки	Специальный перевод
Уровень высшего образования	Специалитет
Форма обучения	Очная
Факультет	Р Международного промышленного менеджмента и коммуникации
Выпускающая кафедра	Р7 ТЕОРЕТИЧЕСКАЯ И ПРИКЛАДНАЯ ЛИНГВИСТИКА
Кафедра-разработчик рабочей программы	Р7 ТЕОРЕТИЧЕСКАЯ И ПРИКЛАДНАЯ ЛИНГВИСТИКА

КУРС	СЕМЕСТР	ОБЩАЯ ТРУДОЁМКОСТЬ (ЗАЧЕТНЫХ ЕДИНИЦ)	ЧАСЫ (по наличию видов занятий)									ВИД ПРОМЕЖУТОЧНОГО КОНТРОЛЯ
			ОБЩАЯ ТРУДОЁМКОСТЬ	АУДИТОРНЫЕ ЗАНЯТИЯ				САМОСТОЯТЕЛЬНАЯ РАБОТА				
				ВСЕГО	ЛЕКЦИИ	ЛАБОРАТОРНЫЙ ПРАКТИКУМ	ПРАКТИЧЕСКИЕ ЗАНЯТИЯ	ВСЕГО	КУРСОВОЙ ПРОЕКТ	КУРСОВАЯ РАБОТА	ДРУГИЕ ВИДЫ САМОСТ. РАБОТЫ	
5	9	3	108	34	17	0	17	74	0	0	74	ЭКЗ.

ЛИСТ СОГЛАСОВАНИЯ

РАБОЧАЯ ПРОГРАММА СОСТАВЛЕНА В СООТВЕТСТВИИ С ТРЕБОВАНИЯМИ ФЕДЕРАЛЬНОГО ГОСУДАРСТВЕННОГО ОБРАЗОВАТЕЛЬНОГО СТАНДАРТА ВЫСШЕГО ОБРАЗОВАНИЯ (ФГОС ВО)

45.05.01 Перевод и переводоведение

год набора группы: 2024

Программу составил:

Кафедра Р7 ТЕОРЕТИЧЕСКАЯ И ПРИКЛАДНАЯ ЛИНГВИСТИКА
Мамаев Иван Дмитриевич, старший преподаватель

Программа рассмотрена
на заседании кафедры-разработчика
рабочей программы **Р7 ТЕОРЕТИЧЕСКАЯ И ПРИКЛАДНАЯ ЛИНГВИСТИКА**

Заведующий кафедрой Невзорова Г.Д., к.ф.н., доц.

Программа рассмотрена
на заседании выпускающей кафедры

Р7 ТЕОРЕТИЧЕСКАЯ И ПРИКЛАДНАЯ ЛИНГВИСТИКА

Заведующий кафедрой Невзорова Г.Д., к.ф.н., доц.

РАБОЧАЯ ПРОГРАММА ДИСЦИПЛИНЫ ТЕХНОЛОГИИ КОРПУСНОЙ ЛИНГВИСТИКИ

Разделы рабочей программы

1. ЦЕЛИ ОСВОЕНИЯ ДИСЦИПЛИНЫ
2. МЕСТО ДИСЦИПЛИНЫ В СТРУКТУРЕ ООП ВО
3. СТРУКТУРА И СОДЕРЖАНИЕ ДИСЦИПЛИНЫ
4. ФОРМЫ КОНТРОЛЯ ОСВОЕНИЯ ДИСЦИПЛИНЫ
5. УЧЕБНО-МЕТОДИЧЕСКОЕ И ИНФОРМАЦИОННОЕ ОБЕСПЕЧЕНИЕ ДИСЦИПЛИНЫ
6. МАТЕРИАЛЬНО-ТЕХНИЧЕСКОЕ ОБЕСПЕЧЕНИЕ ДИСЦИПЛИНЫ

Приложения к рабочей программе дисциплины

- Приложение 1. Аннотация рабочей программы
- Приложение 2. Технологии и формы обучения
- Приложение 3. Фонды оценочных средств

1. ЦЕЛИ ОСВОЕНИЯ ДИСЦИПЛИНЫ

Целью освоения дисциплины является формирование следующих компетенций:

ОПК-4 — способен работать с электронными словарями, различными источниками информации, осуществлять поиск, хранение, обработку и анализ информации, представлять ее в требуемом формате с использованием информационных, компьютерных и сетевых технологий

Формированию компетенций служит достижение следующих результатов образования:

ОПК-4

знания:

сущность процедуры разметки и основные стандарты корпусной разметки;
понятийная база и терминологический минимум корпусной лингвистики;
основные типы корпусов, требования, предъявляемые к корпусам;

умения:

работа с программными средствами и информационными ресурсами корпусной лингвистики;

навыки:

работа с корпусными базами данных в целях лингвистического анализа;
осуществление поиска и исследований на базе корпусов.

2. МЕСТО ДИСЦИПЛИНЫ В СТРУКТУРЕ ООП ВО

Дисциплина **ТЕХНОЛОГИИ КОРПУСНОЙ ЛИНГВИСТИКИ** является дисциплиной **обязательной части блока 1** программы подготовки по направлению *45.05.01 Перевод и переводоведение*.

Содержание дисциплины является логическим продолжением дисциплин: **ОБЩАЯ МОРФОЛОГИЯ, ЛЕКСИКОЛОГИЯ**.

Содержание дисциплины является основой для освоения дисциплин: **НАУЧНО-ИССЛЕДОВАТЕЛЬСКАЯ РАБОТА**.

Предварительные компетенции, сформированные у обучающегося до начала изучения дисциплины:

- ОПК-1 — способен применять знания иностранных языков и знания о закономерностях функционирования языков перевода, а также использовать систему лингвистических знаний при осуществлении профессиональной деятельности

3. СТРУКТУРА И СОДЕРЖАНИЕ ДИСЦИПЛИНЫ

Общая трудоемкость дисциплины составляет 3 з.е., 108 ч.

3.1. Содержание (дидактика) дисциплины

КУРС	СЕМЕСТР	Наименование разделов и дидактических единиц	ВСЕГО	Аудиторные занятия в контактной форме			Самостоятельная работа студентов	Формируемая компетенция, %
				ВСЕГО	Лекции	Практические занятия		ОПК-4
5	9	Раздел 1. Введение в корпусную лингвистику. Классификация корпусов. Языковой материал в лингвистическом исследовании. Общее представление о корпусах и корпусной лингвистике. Основные понятия. Основные требования, предъявляемые к корпусам. Компьютерная поддержка корпусов, программы для работы с корпусами. Принципы классификации корпусов. Типы корпусов. Наиболее известные корпусы.	26	6	3	3	20	25
5	9	Раздел 2. История развития корпусной лингвистики. Национальные корпусы текстов. Зарождение и развитие корпусной лингвистики. Опыт создания корпусов в США и в Европе. Национальные корпусы России, Западной Европы и США. Характеристика корпуса на примере Национального корпуса русского языка.	26	6	3	3	20	25
5	9	Раздел 3. Стандарты и типы разметки в корпусах. Аннотирование корпусов. Принципы разметки. Виды корпусной разметки. Морфологическая разметка корпуса. Синтаксическая разметка корпуса. Семантическая разметка корпуса.	32	14	7	7	18	25
5	9	Раздел 4. Сферы практического использования корпусов. Корпусные методы исследований. Сферы применения корпусов. Особенности применения корпусов текстов. Применение корпусов в гуманитарных исследованиях. Лингвистические исследования, базирующиеся на корпусах. Проблемы и перспективы развития корпусной лингвистики.	24	8	4	4	16	25
Всего за 9 семестр			108	34	17	17	74	100
Всего по дисциплине			108	34	17	17	74	100

3.2. Аудиторный практикум

№ п/п	Номер и наименование раздела дисциплины	Тема практического занятия	Объем, ауд. часов
1	Раздел 1. Введение в корпусную лингвистику. Классификация корпусов.	Принципы классификации корпусов. Типы корпусов. Наиболее известные корпусы.	1
2		Языковой материал в лингвистическом исследовании. Общее представление о корпусах и корпусной лингвистике. Основные понятия. Основные требования, предъявляемые к корпусам. Компьютерная поддержка корпусов, программы для работы с корпусами.	2
3	Раздел 2. История развития корпусной лингвистики.	Зарождение и развитие корпусной лингвистики. Опыт создания корпусов в США и в Европе. Национальные корпусы России, Западной Европы и США.	2
4	Национальные корпусы текстов.	Характеристика корпуса на примере Национального корпуса русского языка.	1
5	Раздел 3. Стандарты и типы разметки в корпусах.	Аннотирование корпусов. Принципы разметки. Виды корпусной разметки.	3
6		Морфологическая разметка корпуса. Синтаксическая разметка корпуса. Семантическая разметка корпуса.	4
7	Раздел 4. Сферы практического использования корпусов. Корпусные методы исследований.	Лингвистические исследования, базирующиеся на корпусах.	1
8		Проблемы и перспективы развития корпусной лингвистики.	1
9		Сферы применения корпусов. Особенности применения корпусов текстов. Применение корпусов в гуманитарных исследованиях.	2
Всего за 9 семестр			17

3.3. Самостоятельная работа студента (СРС)

№ п/п	Номер и наименование раздела дисциплины	Содержание учебного задания	Объем, часов
1	Раздел 1. Введение в корпусную лингвистику. Классификация	Изучение темы 1 по конспектам лекций и рекомендованной литературе для подготовки к	20

	корпусов.	практическим занятиям. Подготовка докладов.	
2	Раздел 2. История развития корпусной лингвистики. Национальные корпуса текстов.	Изучение темы 2 по конспектам лекций и рекомендованной литературе для подготовки к практическим занятиям. Подготовка докладов.	20
3	Раздел 3. Стандарты и типы разметки в корпусах.	Изучение темы 3 по конспектам лекций и рекомендованной литературе для подготовки к практическим занятиям. Подготовка докладов.	18
4	Раздел 4. Сферы практического использования корпусов. Корпусные методы исследований.	Изучение темы 4 по конспектам лекций и рекомендованной литературе для подготовки к практическим занятиям. Подготовка докладов.	16
Всего за 9 семестр			74

4. ФОРМЫ КОНТРОЛЯ ОСВОЕНИЯ ДИСЦИПЛИНЫ

СЕМЕСТР	НЕДЕЛИ СЕМЕСТРА																
	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17
9			Докл, Задан			ДР	Докл, Задан			ДР	Докл, Задан					ДР	Вопр. Экз

Условные обозначения:

- ДР – диагностическая работа;
- Докл – доклад;
- Задан – задание;
- Вопр. Экз – вопросы к экзамену.

Текущий контроль успеваемости студентов проводится в дискретные временные интервалы в следующих формах:

- диагностическая работа;
- доклад;
- задание;
- вопросы к экзамену.

Промежуточная аттестация проводится в формах:

- экзамен.

5. УЧЕБНО-МЕТОДИЧЕСКОЕ И ИНФОРМАЦИОННОЕ ОБЕСПЕЧЕНИЕ ДИСЦИПЛИНЫ

5.1. Основная литература по дисциплине:

1. А. А. Шунейко. . Корпусная лингвистика. Москва: Юрайт, 2022, эл. рес.
2. А. В. Гребенщикова. . Основы квантитативной лингвистики и новых информационных технологий. Москва: Флинта, 2018, эл. рес.
3. Е. В. Булдакова. . Компьютерная лексикография. СПб.БГТУ "ВОЕНМЕХ" им. Д. Ф. Устинова, 2009, эл. рес.

5.2. Дополнительная литература по дисциплине:

1. А. Н. Баранов. . Введение в прикладную лингвистику. М.: Эдиториал УРСС, 2003, 2 экз.

5.3. Периодические издания:

не требуются.

5.4. Перечень ресурсов информационно-телекоммуникационной сети "Интернет", необходимых для освоения дисциплины, электронные библиотечные системы:

1. <https://e.lanbook.com/> — ЭБС Лань;
2. <https://urait.ru/> — Главная – Образовательная платформа Юрайт. Для вузов и ссузов.;
3. http://library.voenmeh.ru/jirbis2/index.php?option=com_irbis&view=irbis&Itemid=474 — Фундаментальная библиотека БГТУ «ВОЕНМЕХ» им. Д.Ф. Устинова.

Современные профессиональные базы данных:

1. <https://rusneb.ru> – Национальная электронная библиотека (НЭБ);
2. <https://cyberleninka.ru/> - Научная электронная библиотека «Киберленинка»;
<http://www.rfbr.ru/rffi/ru/library> - Полнотекстовая электронная библиотека Российского фонда фундаментальных исследований.

Информационные справочные системы:

1. Техэксперт – Информационный портал технического регулирования: Нормы, правила, стандарты РФ;
2. http://library.voenmeh.ru/jirbis2/index.php?option=com_irbis&view=irbis&Itemid=457 - БД ГОСТов собственной генерации БГТУ "ВОЕНМЕХ" им. Д. Ф. Устинова;
3. <http://www.consultant.ru/>- КонсультантПлюс- информационный портал правовой информации.

5.5. Программное обеспечение:

не требуется.

5.6. Информационные технологии:

взаимодействие с обучающимися посредством ЭИОС Moodle БГТУ «ВОЕНМЕХ» им. Д.Ф. Устинова.

6. МАТЕРИАЛЬНО-ТЕХНИЧЕСКОЕ ОБЕСПЕЧЕНИЕ ДИСЦИПЛИНЫ

6.1. Лекционные занятия:

специализированные требования по оборудованию отсутствуют; аудитория с посадочными местами по количеству студентов; доска.

6.2. Практические занятия:

1. Проектор;
2. Интерактивная доска.

6.3. Прочее:

1. рабочее место преподавателя, оснащенное компьютером с доступом в Интернет;
2. рабочие места студентов, оснащенные компьютерами с доступом в Интернет, предназначенные для работы в электронной образовательной среде.

Аннотация рабочей программы

Дисциплина **ТЕХНОЛОГИИ КОРПУСНОЙ ЛИНГВИСТИКИ** является дисциплиной **обязательной части блока 1** программы подготовки по направлению *45.05.01 Перевод и переводоведение*. Дисциплина реализуется на факультете *Р* Международного промышленного менеджмента и коммуникации БГТУ "ВОЕНМЕХ" им. Д.Ф. Устинова кафедрой *Р7 ТЕОРЕТИЧЕСКАЯ И ПРИКЛАДНАЯ ЛИНГВИСТИКА*.

Дисциплина нацелена на формирование *компетенций*:

ОПК-4 способен работать с электронными словарями, различными источниками информации, осуществлять поиск, хранение, обработку и анализ информации, представлять ее в требуемом формате с использованием информационных, компьютерных и сетевых технологий.

Содержание дисциплины охватывает круг вопросов, связанных с изучением предмета и задач корпусной лингвистики; места корпусной лингвистики в исследовании языка, основных разделов корпусной лингвистики.

Программой дисциплины предусмотрены следующие **виды контроля**:

Текущий контроль успеваемости студентов проводится в дискретные временные интервалы в следующих формах:

- диагностическая работа;
- доклад;
- задание;
- вопросы к экзамену.

Промежуточная аттестация проводится в формах:

- экзамен.

Общая трудоемкость освоения дисциплины составляет **3 з.е., 108 ч.** Программой дисциплины предусмотрены лекционные занятия (**17 ч.**), практические занятия (**17 ч.**), самостоятельная работа студента (**74 ч.**).

ТЕХНОЛОГИИ И ФОРМЫ ОБУЧЕНИЯ

Рекомендации по освоению дисциплины для студента

Трудоемкость освоения дисциплины составляет 108 ч., из них 34 ч. аудиторных занятий, и 74 ч., отведенных на самостоятельную работу студента.

Рекомендации по распределению учебного времени по видам самостоятельной работы и разделам дисциплины приведены в таблице.

Контроль освоения дисциплины производится в соответствии с Положением о текущем, рубежном контроле успеваемости и промежуточной аттестации обучающихся.

Формы контроля и критерии оценивания приведены в приложении 3 к Рабочей программе.

Наименование работы	Рекомендуемая литература	Трудоемкость, час.
Раздел 1. Введение в корпусную лингвистику. Классификация корпусов.		
Изучение темы 1 по конспектам лекций и рекомендованной литературе для подготовки к практическим занятиям. Подготовка докладов.	А. Н. Баранов. . Введение в прикладную лингвистику: М.: Эдиториал УРСС, 2003 (3) А. В. Гребенщикова. . Основы квантитативной лингвистики и новых информационных технологий: Москва: Флинта, 2018 (5)	20
Итого по разделу 1		20
Раздел 2. История развития корпусной лингвистики. Национальные корпуса текстов.		
Изучение темы 2 по конспектам лекций и рекомендованной литературе для подготовки к практическим занятиям. Подготовка докладов.	Е. В. Булдакова. . Компьютерная лексикография: СПб.БГТУ "ВОЕНМЕХ" им. Д. Ф. Устинова, 2009 (5) А. В. Гребенщикова. . Основы квантитативной лингвистики и новых информационных технологий: Москва: Флинта, 2018 (5)	20
Итого по разделу 2		20
Раздел 3. Стандарты и типы разметки в корпусах.		
Изучение темы 3 по конспектам лекций и рекомендованной литературе для подготовки к практическим занятиям. Подготовка докладов.	А. А. Шунейко. . Корпусная лингвистика: Москва: Юрайт, 2022 (1) А. В. Гребенщикова. . Основы квантитативной лингвистики и новых информационных технологий: Москва: Флинта, 2018 (5)	18
Итого по разделу 3		18
Раздел 4. Сферы практического использования корпусов. Корпусные методы исследований.		
Изучение темы 4 по конспектам лекций и рекомендованной литературе для подготовки к практическим занятиям. Подготовка докладов.	А. Н. Баранов. . Введение в прикладную лингвистику: М.: Эдиториал УРСС, 2003 (3) А. А. Шунейко. . Корпусная лингвистика: Москва: Юрайт, 2022 (1)	16
Итого по разделу 4		16

ФОНД ОЦЕНОЧНЫХ СРЕДСТВ

Фонд оценочных средств, позволяющие оценить результаты обучения по данной дисциплине, включают в себя:

- диагностическая работа
- доклад;
- задание;
- вопросы к экзамену;
- экзамен.

Критерии оценивания

Диагностическая работа

Диагностическая работа проводится в форме теста в ЭИОС Moodle:

- при правильном ответе менее чем на 60% вопросов - не аттестация;
- при правильном ответе на 60% вопросов и более - аттестация.

Доклад

В семестре предусмотрено выполнение студентами двух докладов.

Оценка «отлично»:

Студент чётко, грамотно и последовательно выстраивает свой ответ; умеет показать связь теории и практики. Раскрывает тему доклада глубоко и всесторонне. Приводит собственные примеры, иллюстрирующие глубокое понимание теоретического материала; исчерпывающе отвечает на вопросы, не допускает неточностей. Доклад чётко структурирован: состоит из введения, основной части и заключения. В заключении подводятся итоги доклада.

Оценка «хорошо»:

Студент грамотно и последовательно излагает материал, умеет показать связь теории и практики, не допускает существенных неточностей в ответах на вопросы, может правильно применять теоретические положения, испытывает незначительные затруднения при самостоятельном обобщении материала. Приводит примеры только из основных источников, не приводит собственных примеров, иллюстрирующих теорию. Раскрывает тему доклада глубоко и всесторонне. Доклад чётко структурирован: состоит из введения, основной части и заключения. В заключении подводятся итоги доклада.

Оценка «удовлетворительно»:

Студент не знает отдельных положений, в ответе допускает неточности, недостаточно правильные формулировки, нарушает последовательность в изложении материала, испытывает затруднения при самостоятельном обобщении материала. Приводит примеры из курса лекций, иллюстрирующие теорию. Доклад не является чётко структурированным.

Оценка «неудовлетворительно»:

Студент не знает значительной части материала, в ответе допускает существенные ошибки, неправильные формулировки, испытывает значительные затруднения при самостоятельном обобщении материала. Не приводит примеров, иллюстрирующих теорию. Доклад не является чётко структурированным.

Задание

Задание направлено на развитие практических навыков работы в корпусными ресурсами (Национальный корпус русского языка, веб-корпуса и пр.): отработка языка корпусных запросов, создание подкорпусов и пр.

Оценка «отлично»:

Студент представил результаты задания, которые корректны и не содержат ошибок.

Оценка «хорошо»:

Студент представил результаты задания, которые корректны и содержат не более 2-х грубых ошибок.

Оценка «удовлетворительно»:

Студент представил результаты задания, которые корректны и содержат не более 4-х грубых ошибок.

Оценка «неудовлетворительно»:

В задании выявлено 5 и более грубых ошибок.

Вопросы к экзамену

1. История корпусной лингвистики.
2. Этапы создания корпуса.
3. Виды лингвистической разметки данных.
4. Метаразметка в корпусных ресурсах.
5. Национальный корпус русского языка.
6. Британский национальный корпус.
7. Генеральный Интернет-корпус русского языка.
8. Семейство корпусов TenTen.
9. Семейство корпусов Aranea.
10. Корпусный менеджер Sketch Engine. Автоматическое создание собственного корпуса.
11. Корпус «Один речевой день». Особенности разметки устных корпусов.
12. Информационная система Google Ngram Viewer.
13. Корпус RusIdiolect. Профилирование автора письменного текста.
14. Исследования славянских рукописей и древнерусских текстов: ресурс СКАТ и корпус «Манускрипт».
15. Практическое использование корпусов в переводоведении.
16. Исследование коллокаций в корпусах текстов. Метрики устойчивости.
17. Стандарт TEI.
18. Корпус драматических текстов DraCor. Особенности разметки и визуализации данных.
19. Язык запросов к корпусу CQL.
20. Веб как корпусный ресурс.

Экзамен

Оценка за экзамен выставляется согласно баллам по технологической карте, которые обучающийся набрал за семестр. При несогласии с баллами студент имеет право сдать экзамен, список экзаменационных вопросов представлен в секции «Вопросы к экзамену».

Паспорт фонда оценочных средств

КУРС	СЕМЕСТР	Наименование разделов и дидактических единиц	ВСЕГО	Аудиторные занятия в контактной форме			Самостоятельная работа студентов	Формируемая компетенция, %	НАИМЕНОВАНИЕ ОЦЕНОЧНОГО СРЕДСТВА
				ВСЕГО	Лекции	Практические занятия		ОПК-4	
5	9	Раздел 1. Введение в корпусную лингвистику. Классификация корпусов.	26	6	3	3	20	25	Доклад
5	9	Раздел 2. История развития корпусной лингвистики. Национальные корпуса текстов.	26	6	3	3	20	25	Доклад
5	9	Раздел 3. Стандарты и типы разметки в корпусах.	32	14	7	7	18	25	Доклад, Задание
5	9	Раздел 4. Сферы практического использования корпусов. Корпусные методы исследований.	24	8	4	4	16	25	Доклад, Задание, Вопросы к экзамену
Всего за 9 семестр			108	34	17	17	74	100	
Всего по дисциплине			108	34	17	17	74	100	

Критерии оценивания

ОПК-4

- Вопросы открытого типа:*
- № 1 Первый корпус текстов назывался _____ корпус.
- № 2 POS-разметка в корпусах известна как _____.
- № 3 Синтаксическая разметка в НКРЯ представлена в двух форматах — в формате _____ и в формате Universal Dependencies.
- № 4 ГИКРЯ — _____ интернет-корпус русского языка.
- № 5 _____ корпус НКРЯ посвящен истории русского ударения, играющего ключевую роль в его разметке.
- № 6 В _____ корпусе НКРЯ текст снабжается видео- или аудиозаписью.
- № 7 _____ корпус НКРЯ включает в себя тексты СМИ.
- № 8 Метрика log-likelihood для извлечения коллокаций из корпуса известна в русскоязычной традиции как метрика _____ правдоподобия.
- № 9 CQL расшифровывается как _____ Query Language.
- № 10 Что такое корпусная разметка?
- Вопросы закрытого типа:*
- № 1 Какой корпус в НКРЯ может быть использован в переводческой работе?
- устный
- газетный
- мультимедийный
- параллельный
- № 2 НКРЯ разрабатывался в таких городах, как:
- Сталинград
- Москва
- Санкт-Петербург
- Воронеж
- № 3 Задача снятия омонимии в корпусе известна как:
- тропонимия
- дизамбигуация
- холонимия
- меронимия
- № 4 Для морфологической разметки Британского национального корпуса применялась система:
- CLAWS
- Python
- GitHub
- Request
- № 5 Объём Британского национального корпуса равен:
- 10 млн словоупотреблений
- 100 млн словоупотреблений
- 1 млн словоупотреблений

- 100 млрд словоупотреблений
- № 6 Сколько процентов занимают устные тексты в Британском национальном корпусе?
- 10%
- 30%
- 90%
- 50%
- № 7 Относительная частота словоупотреблений в корпусе известна как:
- likelihood
- dice
- ipm
- tf-idf
- № 8 Сколько процентов занимают письменные тексты в Британском национальном корпусе?
- 10%
- 30%
- 90%
- 50%
- № 9 Установите правильную последовательность создания корпуса:
1. Разметка текстов.
 2. Сбор текстов.
 3. Загрузка корпуса на некоторый ресурс.
 4. Тестирование корпуса.
 5. Начало функционирования корпуса.
- № 10 Соотнесите корпус и год его создания.
1. CoCoCo
 2. База данных русскоязычных коллокаций
 3. Британский национальный корпус
 4. ГИКРЯ
- A. 2015
- Б. 2012
- В. 1994
- Г. 2021