Документ подписан простой электронной подписью
Информация о владельце:
ФИО: Гриб Владислав Валерьевич
Должность: Ректор
Дата подписания: 23.09.2026 14:25:42
Уникальный программный ключ:
637517d24e103c3db032acf37e839d98ec1c5bb2f5eb89c29abfcd7f43985447

Образовательное частное учреждение высшего образования
«МОСКОВСКИЙ УНИВЕРСИТЕТ ИМЕНИ А.С. ГРИБОЕДОВА»
(ИМПЭ им. А.С. Грибоедова)

МЕЖДУНАРОДНЫЙ ИНСТИТУТ ИНОСТРАННЫХ ЯЗЫКОВ И МЕЖКУЛЬТУРНЫХ КОММУНИКАЦИЙ

УТВЕРЖДАЮ
Директор Международного
института иностранных языков
и межкультурных
коммуникаций
Д.В. Парамонова
«23» декабря 2025 г.

ФОНД ОЦЕНОЧНЫХ СРЕДСТВ

КВАНТИТАТИВНАЯ ЛИНГВИСТИКА

Направление подготовки
45.04.02 Лингвистика
(уровень магистратуры)

профилю:
«Международная коммуникация в сфере бизнеса и права»

Форма обучения: очная, заочная

Москва

Фонд оценочных средств для дисциплины «Квантитативная лингвистика». Направление подготовки 45.04.02 Лингвистика, направленность (профиль): «Международная коммуникация в сфере бизнеса и права» / Т.А. Борисовская – М.: ИМПЭ им. А.С. Грибоедова.

Фонд оценочных средств является неотъемлемой частью рабочей программы дисциплины.

Разработчик: Т.А. Борисовская

Заведующий кафедрой: _____/ Д. В. Парамонова

1. ФОНД ОЦЕНОЧНЫХ СРЕДСТВ ДЛЯ ТЕКУЩЕГО КОНТРОЛЯ УСПЕВАЕМОСТИ, ПРОМЕЖУТОЧНОЙ АТТЕСТАЦИИ ОБУЧАЮЩИХСЯ ПО ДИСЦИПЛИНЕ

Настоящий Фонд оценочных средств (ФОС) по дисциплине «Квантитативная лингвистика» является неотъемлемым приложением к рабочей программе дисциплины (РПД) «Квантитативная лингвистика». На данный ФОС распространяются все реквизиты утверждения, представленные в РПД по данной дисциплине.

2. ПЕРЕЧЕНЬ ОЦЕНОЧНЫХ СРЕДСТВ

Для определения качества освоения обучающимися учебного материала по дисциплине используются следующие оценочные средства:

№ п/п	Оценочное средство	Краткая характеристика оценочного средства	Представление оценочного средства в ФОС
1	Тестирование	Вид контроля, позволяющий оценить изученный теоретический материал.	Вопросы для проведения тестирования
2	Практические задания	Вид контроля, позволяющий оценить умение обучающегося применять осваиваемую компетенцию в практических ситуациях и при решении производственных задач	Задания к практическому занятию
3	Контрольная работа	Вид контроля, позволяющий определить результат освоения компетенций по дисциплине в рамках рассматриваемой темы, оцениваемый с помощью соответствующих индикаторов достижения компетенций	Задания контрольной работы
4	Самостоятельная работа	Вид контроля, позволяющий оценить проработку теоретического материала, изучение рекомендуемой литературы, выполнение практико-ориентированных заданий (заполнение таблиц, проведение сравнительного анализа, составление схем и др.), решение практических задач, создание презентаций, написание рефератов, подборку нормативного и иного материала и выполнение других заданий	Задания самостоятельной работы
5	Курсовая работа	Вид контроля, позволяющий выявить степень владения базовыми знаниями, умениями и навыками, необходимыми для обучения, и определить уровень владения новым материалом	Индивидуальные задания (темы) для курсовой работы
6	Зачет/Зачет с оценкой/Экзамен	Вид контроля, позволяющий выявить степень овладения знаниями, умениями и навыками, необходимыми для дальнейшего освоения образовательной программы подготовки	Вопросы для подготовки к зачету/зачету с оценкой/экзамену

3. ПАСПОРТ ФОНДА ОЦЕНОЧНЫХ СРЕДСТВ ДИСЦИПЛИНЕ

3.1. Сопроводительная информация.

Разработчик	Т.А. Борисовская
Кафедра	Лингвистики и переводоведения
Наименование дисциплины	Квантитативная лингвистика
Факультет / институт	Международный институт иностранных языков и межкультурных коммуникаций
Направление подготовки / специальность	45.04.02 Лингвистика
Количество вопросов в оценочных заданиях (диапазон)	20-45
Общее время тестирования (мин)	90 мин
Общее количество вопросов/заданий в ФОС	*42*
Размещенность на сайте Университета примерного перечня вопросов, заданий ФОС – для подготовки обучающихся к прохождению оценки (да / нет)	да

3.2. Характеристика оцениваемых компетенций.

Код компетенции	Формулировка компетенции	Индикаторы достижения компетенции
ОПК-6	Способен применять современные технологии при осуществлении сбора, обработки и интерпретации данных эмпирического исследования; составлять и оформлять научную документацию;	**ИОПК-6.1. Знает** современные технологии сбора, обработки и интерпретации полученных экспериментальных данных. **ИОПК-6.2. Умеет** применять современные технологии сбора, обработки и интерпретации полученных экспериментальных данных. **ИОПК-6.3. Владеет** навыками сбора, обработки и интерпретации полученных экспериментальных данных.

4. СОДЕРЖАНИЕ ОЦЕНОЧНЫХ СРЕДСТВ ТЕКУЩЕГО КОНТРОЛЯ

4.1. ТИПОВЫЕ ТЕСТОВЫЕ ЗАДАНИЯ.

Тесты содержат набор вопросов, в полном объеме охватывающие изученный теоретический материал по указанной теме (индикаторы ЗНАТЬ). Выполнение тестов позволяет определить результат освоения компетенций по дисциплине в рамках рассматриваемой темы, оцениваемый с помощью соответствующих индикаторов достижения компетенций. Индивидуальный тестовый сеанс для каждого обучающегося формируется по специальному алгоритму, обеспечивающему заданную тематическую структуру и пропорциональное наличие вопросов разного типа и сложности.

При формировании тестов необходимо использовать задания следующих типов:

Тип задания 1. Задания закрытого типа на установление соответствия.

Тип задания 2. Задания закрытого типа на установление последовательности.

Тип задания 3. Задания комбинированного типа, предполагающие выбор одного правильного ответа из предложенных с последующим объяснением своего выбора.

Тип задания 4. Задания комбинированного типа, предполагающие выбор нескольких ответов из предложенных с последующим объяснением своего выбора.

Тип задания 5. Задания открытого типа с развернутым ответом.

<table>
<tr><th>Тема</th><th>Код и формулировка компетенции</th><th>Индикаторы достижения компетенции</th><th>Тест по теме</th></tr>
<tr>
<td>Тема 1. Современная квантитативная лингвистика как наука. Новые информационные технологии в лингвистических исследованиях и образовательном процессе.</td>
<td>ОПК-6. Способен применять современные технологии при осуществлении сбора, обработки и интерпретации данных эмпирического исследования; составлять и оформлять научную документацию;</td>
<td>ИОПК-6.1. Знает современные технологии сбора, обработки и интерпретации полученных экспериментальных данных.</td>
<td>
Задания закрытого типа на установление соответствия.
Соотнесите каждый метод или технологию сбора данных (1–4) с соответствующим описанием (А–Г). Запишите выбранные буквы рядом с цифрами.
<table>
<tr><td></td><td>Метод / технология</td><td></td><td>Описание</td></tr>
<tr><td>1</td><td>Корпусная лингвистика</td><td>А</td><td>Сбор данных о движении глаз испытуемого при чтении или восприятии речи для изучения когнитивных процессов и языковой обработки.</td></tr>
<tr><td>2</td><td>Онлайн-анкетирование (Google Forms, Яндекс.Взгляд)</td><td>Б</td><td>Автоматизированный сбор и разметка больших массивов текстов с последующим квантитативным анализом частотности грамматических и лексических единиц.</td></tr>
<tr><td>3</td><td>Айтрекинг (eye-tracking)</td><td>В</td><td>Привлечение добровольцев через интернет-платформы (например, Amazon Mechanical Turk) для выполнения микро-заданий по разметке, оценке или сбору языкового материала.</td></tr>
<tr><td>4</td><td>Краудсорсинг</td><td>Г</td><td>Автоматизированный сбор структурированных ответов респондентов с возможностью экспорта данных в форматы, пригодные для статистической обработки.</td></tr>
</table>
Ответ:
<table>
<tr><td>1</td><td>2</td><td>3</td><td>4</td></tr>
<tr><td></td><td></td><td></td><td></td></tr>
</table>
Ответ:
<table>
<tr><td>1</td><td>2</td><td>3</td><td>4</td></tr>
</table>
</td>
</tr>
</table>

<table>
<tr><th>Тема</th><th>Код и формулировка компетенции</th><th>Индикаторы достижения компетенции</th><th>Тест по теме</th></tr>
<tr><td></td><td></td><td></td><td>
<table>
<tr><td>Б</td><td>Г</td><td>А</td><td>В</td></tr>
</table>

Задания закрытого типа на установление соответствия.
Соотнесите каждый программный инструмент (1–4) с его основным функциональным назначением (А–Г). Запишите выбранные буквы рядом с цифрами.
<table>
<tr><td></td><td>Инструмент</td><td></td><td>Назначение</td></tr>
<tr><td>1</td><td>AntConc</td><td>А</td><td>Среда статистического программирования (включая пакеты tidyverse, ggplot2), позволяющая выполнять многомерный анализ (регрессию, кластеризацию) и строить сложные визуализации.</td></tr>
<tr><td>2</td><td>SPSS</td><td>Б</td><td>Построение конкордансов, вычисление ключевых слов (keyness), частотных списков и коллокаций на основе корпусных данных.</td></tr>
<tr><td>3</td><td>R</td><td>В</td><td>Программный пакет для статистического анализа с графическим интерфейсом, широко используемый в социальных науках</td></tr>
<tr><td>4</td><td>TXM</td><td>Г</td><td>Платформа для корпусного анализа с инструментами статистической обработки и визуализации (факторный анализ, иерархическая классификация, кластеризация текстов).</td></tr>
</table>

Ответ:
<table>
<tr><td>1</td><td>2</td><td>3</td><td>4</td></tr>
<tr><td></td><td></td><td></td><td></td></tr>
</table>

Ответ:
<table>
<tr><td>1</td><td>2</td><td>3</td><td>4</td></tr>
<tr><td>Б</td><td>В</td><td>А</td><td>Г</td></tr>
</table>
</td></tr>
</table>

<table>
<tr><th>Тема</th><th>Код и формулировка компетенции</th><th>Индикаторы достижения компетенции</th><th>Тест по теме</th></tr>
<tr><td></td><td></td><td></td><td></td></tr>
<tr><td>Тема 1. Современная квантитативная лингвистика как наука. Новые информационные технологии в лингвистических исследованиях и образовательном процессе.</td><td>ОПК-6. Способен применять современные технологии при осуществлении сбора, обработки и интерпретации данных эмпирического исследования; составлять и оформлять научную документацию;</td><td>ИОПК-6.1. Знает современные технологии сбора, обработки и интерпретации полученных экспериментальных данных.</td><td>Задания закрытого типа на установление последовательности
Расположите перечисленные этапы проведения экспериментального квантитативно-лингвистического исследования с использованием автоматизированного сбора данных в логическом порядке от начала к концу исследования. Запишите цифры в правильной последовательности.

1. Статистическая обработка полученных данных и проверка гипотез (например, в R или SPSS).
2. Визуализация результатов и их лингвистическая интерпретация.
3. Сбор экспериментальных данных с помощью специализированного программного обеспечения (например, PsychoPy, OpenSesame).
4. Формулировка исследовательского вопроса и операционализация переменных.
5. Разработка дизайна эксперимента, создание стимульного материала.
6. Экспорт сырых данных в формат, пригодный для анализа (CSV, JSON).

Ответ:
<table><tr><td></td><td></td><td></td><td></td><td></td><td></td></tr></table>Ответ:
<table><tr><td>4</td><td>5</td><td>3</td><td>6</td><td>1</td><td>2</td></tr></table>
Задания закрытого типа на установление последовательности
Расположите технологические шаги обработки текстовых корпусных данных с использованием методов квантитативной лингвистики в правильном порядке. Запишите цифры в нужной последовательности.

1. Токенизация и лемматизация текстов с помощью автоматических процессоров</td></tr>
</table>

<table>
<tr><th>Тема</th><th>Код и формулировка компетенции</th><th>Индикаторы достижения компетенции</th><th>Тест по теме</th></tr>
<tr><td></td><td></td><td></td><td>(например, TreeTagger, SpaCy).
2. Вычисление статистических показателей (частотность, коллокации, меры разнообразия лексики).
3. Сбор и формирование корпуса текстов (скачивание, очистка от шума, приведение к единому формату).
4. Построение конкордансов и кластерный анализ для выявления паттернов.
5. Разметка метаданных (жанр, автор, дата) для последующего сравнения групп текстов.
6. Интерпретация результатов с опорой на лингвистическую теорию.

Ответ:
<table><tr><td></td><td></td><td></td><td></td><td></td><td></td></tr></table>Ответ:
<table><tr><td>3</td><td>5</td><td>1</td><td>2</td><td>4</td><td>6</td></tr></table></td></tr>
<tr><td>Тема 1. Современная квантитативная лингвистика как наука. Новые информационные технологии в лингвистических исследованиях и образовательном процессе.</td><td>ОПК-6. Способен применять современные технологии при осуществлении сбора, обработки и интерпретации данных эмпирического исследования; составлять и оформлять научную</td><td>ИОПК-6.1. Знает современные технологии сбора, обработки и интерпретации полученных экспериментальных данных.</td><td>Задания комбинированного типа, предполагающие выбор одного правильного ответа из предложенных с последующим объяснением своего выбора.
Исследователь хочет изучить, как носители языка воспринимают редкие лексические единицы в потоке речи в реальном времени. Необходимо зафиксировать не только правильность понимания, но и микропроцессы: латентное время реакции, фиксации взгляда на ключевых словах, зрачковую реакцию. Какая технология наиболее адекватна для этих целей?
Варианты ответов:
1. Корпусный анализ с частотным словарём
2. Онлайн-анкетирование в Google Forms
3. Айтрекинг (eye-tracking) с синхронной регистрацией времени реакции
4. Краудсорсинг на платформе Toloka</td></tr>
</table>

Тема	Код и формулировка компетенции	Индикаторы достижения компетенции	Тест по теме
	документацию;		**Ответ: 3** **Обоснование:** Айтрекинг позволяет регистрировать движение глаз, размер зрачка и фиксации на конкретных стимулах в реальном времени, а при синхронизации с кнопочным устройством — и время реакции. Это даёт количественные данные о когнитивных процессах при восприятии редких слов. ***Задания комбинированного типа, предполагающие выбор одного правильного ответа из предложенных с последующим объяснением своего выбора.*** Лингвист провёл эксперимент с использованием психолингвистической базы данных и получил таблицу с числовыми значениями. Какой порядок действий наиболее соответствует современным стандартам обработки и интерпретации данных в квантитативной лингвистике? **Варианты ответов:** 1. Сразу построить красивые графики → подобрать статистический критерий → выполнить расчёты → сформулировать выводы 2. Выполнить описательную статистику (средние, стандартные отклонения) → проверить нормальность распределения → выбрать параметрический или непараметрический критерий → вычислить значимость → визуализировать → интерпретировать с опорой на теорию 3. Ввести данные в SPSS → нажать «Анализ» → выбрать t-критерий → записать p-value → сделать вывод без визуализации 4. Сразу применить многомерное шкалирование → получить кластеры → описать их содержательно → опубликовать результаты **Ответ: 2** **Обоснование:** Современный стандарт обработки количественных лингвистических данных

Тема	Код и формулировка компетенции	Индикаторы достижения компетенции	Тест по теме
			требует сначала исследовать распределение (описательная статистика, проверка нормальности), затем обоснованно выбрать статистический критерий, вычислить значимость, после чего визуализировать результаты и дать лингвистическую интерпретацию.
Тема 1. Современная квантитативная лингвистика как наука. Новые информационные технологии в лингвистических исследованиях и образовательном процессе.	**ОПК-6**. Способен применять современные технологии при осуществлении сбора, обработки и интерпретации данных эмпирического исследования; составлять и оформлять научную документацию;	**ИОПК-6.1.** Знает современные технологии сбора, обработки и интерпретации полученных экспериментальных данных.	***Задания комбинированного типа, предполагающие выбор нескольких ответов из предложенных с последующим объяснением своего выбора.*** Какие из перечисленных технологий и инструментов относятся к современным методам сбора первичных экспериментальных данных в квантитативной лингвистике? Выберите все верные варианты и объясните, почему остальные не подходят. **Варианты ответов:** 1. Айтрекинг (eye-tracking) 2. Корпусный менеджер AntConc 3. Программа для онлайн-экспериментов PsychoPy 4. Статистический пакет SPSS 5. Платформа краудсорсинга (Yandex.Toloka, Amazon Mechanical Turk) **Ответ: 1, 3, 5** **Обоснование:** 1. Айтрекинг (1) — технология сбора данных о движении глаз, времени фиксации, зрачковых реакциях в реальном времени. 2. PsychoPy (3) — программная среда для создания и проведения психолингвистических экспериментов с регистрацией реакций (точность, время ответа). 3. Краудсорсинг (5) — позволяет собирать размеченные данные, оценки естественности, переводы и т.д. от множества участников через интернет.

Тема	Код и формулировка компетенции	Индикаторы достижения компетенции	Тест по теме
			Задания комбинированного типа, предполагающие выбор нескольких ответов из предложенных с последующим объяснением своего выбора. Какие этапы и подходы являются обязательными для корректной интерпретации количественных лингвистических данных при использовании современных информационных технологий? Выберите все правильные позиции и кратко обоснуйте, почему без них интерпретация будет неполной или ошибочной. **Варианты ответов:** А) Построение описательной статистики (среднее, медиана, стандартное отклонение) Б) Визуализация данных (гистограммы, ящики с усами, графики разброса) В) Проверка статистических допущений (нормальность, гомогенность дисперсий) Г) Автоматический выбор любого критерия по умолчанию (например, t-теста) без проверки условий Д) Интерпретация значимости результатов в контексте лингвистической теории **Ответ: А, Б, В, Д** **Обоснование:** А (описательная статистика) — даёт первичное понимание центральной тенденции и разброса, без неё невозможно оценить характер данных. Б (визуализация) — позволяет обнаружить выбросы, асимметрию, мультимодальность, которые не видны в числах. В (проверка допущений) — необходима для выбора корректного статистического критерия (параметрический / непараметрический), иначе интерпретация p-value может быть ложной. Д (лингвистическая интерпретация) — превращает статистические паттерны в осмысленный вывод о языке; без неё анализ остаётся чисто математическим.

<table>
<tr><th>Тема</th><th>Код и формулировка компетенции</th><th>Индикаторы достижения компетенции</th><th>Тест по теме</th></tr>
<tr><td>Тема 2. Корпусная лингвистика как самостоятельное лингвистическое направление. Характеристики корпуса (требования, репрезентативность и пропорциональность</td><td>ОПК-6. Способен применять современные технологии при осуществлении сбора, обработки и интерпретации данных эмпирического исследования; составлять и оформлять научную документацию;</td><td>ИОПК-6.1. Знает современные технологии сбора, обработки и интерпретации полученных экспериментальных данных.</td><td>
Задания закрытого типа на установление соответствия.
Соотнесите каждый принцип / требование к лингвистическому корпусу (1–4) с его правильным определением или описанием (А–Г).
<table>
<tr><td></td><td>Принцип / требование</td><td></td><td>Описание</td></tr>
<tr><td>1</td><td>Репрезентативность</td><td>А</td><td>Соблюдение заданных пропорций между разными жанрами, стилями, периодами или группами авторов в соответствии с целевой генеральной совокупностью текстов.</td></tr>
<tr><td>2</td><td>Пропорциональность</td><td>Б</td><td>Возможность автоматического поиска и извлечения примеров с помощью программных средств (конкордансеров).</td></tr>
<tr><td>3</td><td>Аннотирование</td><td>В</td><td>Способность корпуса отражать все основные характеристики языка или подъязыка, для которого он создавался, в тех же пропорциях, в которых они встречаются в реальности.</td></tr>
<tr><td>4</td><td>Машинная читаемость</td><td>Г</td><td>Добавление к текстам метаинформации (морфологической, синтаксической, семантической) для возможности квантитативного анализа.</td></tr>
</table>
Ответ:
<table>
<tr><td>1</td><td>2</td><td>3</td><td>4</td></tr>
<tr><td></td><td></td><td></td><td></td></tr>
</table>
Ответ:
<table>
<tr><td>1</td><td>2</td><td>3</td><td>4</td></tr>
<tr><td>В</td><td>А</td><td>Б</td><td>Г</td></tr>
</table>
</td></tr>
</table>

<table>
<tr><th>Тема</th><th>Код и формулировка компетенции</th><th>Индикаторы достижения компетенции</th><th>Тест по теме</th></tr>
<tr><td></td><td></td><td></td><td>
Задания закрытого типа на установление соответствия.
Соотнесите каждый этап построения лингвистического корпуса (1–4) с технологией или программным инструментом, который наиболее часто применяется на этом этапе (А–Г).

<table>
<tr><td></td><td>Этап работы с корпусом</td><td></td><td>Технология / инструмент</td></tr>
<tr><td>1</td><td>Сбор текстовых материалов в больших объёмах</td><td>А</td><td>Веб-краулеры (парсеры) для автоматического скачивания текстов из интернета.</td></tr>
<tr><td>2</td><td>Приведение текстов к единому формату и очистка от шума (HTML-тегов, мусорных символов)</td><td>Б</td><td>Программы лемматизации и морфологической разметки (например, MyStem, TreeTagger).</td></tr>
<tr><td>3</td><td>Лингвистическая разметка (токенизация, лемматизация, частеречная аннотация)</td><td>В</td><td>Скрипты на Python с использованием библиотек BeautifulSoup, re (регулярные выражения).</td></tr>
<tr><td>4</td><td>Оценка репрезентативности и пропорциональности корпуса</td><td>Г</td><td>Статистические методы (хи-квадрат, измерение разнообразия лексики, сравнение частот с эталонным корпусом).</td></tr>
</table>

Ответ:
<table>
<tr><td>1</td><td>2</td><td>3</td><td>4</td></tr>
<tr><td></td><td></td><td></td><td></td></tr>
</table>

Ответ:
<table>
<tr><td>1</td><td>2</td><td>3</td><td>4</td></tr>
<tr><td>А</td><td>В</td><td>Б</td><td>Г</td></tr>
</table>
</td></tr>
</table>

<table>
<tr><th>Тема</th><th>Код и формулировка компетенции</th><th>Индикаторы достижения компетенции</th><th>Тест по теме</th></tr>
<tr><td></td><td></td><td></td><td></td></tr>
<tr><td>Тема 2. Корпусная лингвистика как самостоятельное лингвистическое направление. Характеристики корпуса (требования, репрезентативность и пропорциональность</td><td>ОПК-6. Способен применять современные технологии при осуществлении сбора, обработки и интерпретации данных эмпирического исследования; составлять и оформлять научную документацию;</td><td>ИОПК-6.1. Знает современные технологии сбора, обработки и интерпретации полученных экспериментальных данных.</td><td>Задания закрытого типа на установление последовательности
Расположите перечисленные этапы создания сбалансированного и репрезентативного корпуса текстов в порядке от начала к концу. Запишите цифры в правильной последовательности.

Шаги (в случайном порядке):
1. Оценка репрезентативности и пропорциональности с помощью статистических методов (сравнение частот с эталонным корпусом).
2. Определение целевой генеральной совокупности текстов (жанры, стили, периоды, авторы).
3. Очистка и нормализация собранных текстов (удаление шума, приведение к единому формату).
4. Расчёт необходимых объёмов текстов для каждой страты с учётом пропорций.
5. Сбор текстов из доступных источников (веб-краулеры, оцифрованные архивы).
6. Лингвистическая разметка (токенизация, лемматизация, частеречная аннотация).
Ответ:
<table><tr><td></td><td></td><td></td><td></td><td></td><td></td></tr></table>
Ответ:
<table><tr><td>2</td><td>4</td><td>5</td><td>3</td><td>6</td><td>1</td></tr></table>
Задания закрытого типа на установление последовательности
Расположите шаги в логическом порядке последовательность обработки и интерпретации корпусных данных с использованием современных технологий</td></tr>
</table>

<table>
<tr><th>Тема</th><th>Код и формулировка компетенции</th><th>Индикаторы достижения компетенции</th><th>Тест по теме</th></tr>
<tr><td></td><td></td><td></td><td>от сырых данных к получению интерпретируемого результата. Запишите цифры в правильной последовательности.

Шаги (в случайном порядке):
1. Построение конкордансов и частотных списков с помощью корпусного менеджера (AntConc, TXM).
2. Формулировка исследовательского вопроса (например, различие в употреблении грамматической конструкции).
3. Визуализация результатов (гистограммы, тепловые карты) и их лингвистическая интерпретация.
4. Статистическая проверка значимости различий (хи-квадрат, log-likelihood, точный критерий Фишера).
5. Выбор корпуса или создание подкорпуса, отвечающего критериям репрезентативности и пропорциональности.
6. Экспорт данных из корпуса в формат для статистической обработки (CSV, Excel).

Ответ:
<table><tr><td></td><td></td><td></td><td></td><td></td><td></td></tr></table>

Ответ:
<table><tr><td>2</td><td>5</td><td>1</td><td>6</td><td>4</td><td>3</td></tr></table></td></tr>
<tr><td>Тема 2. Корпусная лингвистика как самостоятельное лингвистическое направление.</td><td>ОПК-6. Способен применять современные технологии при осуществлении</td><td>ИОПК-6.1. Знает современные технологии сбора, обработки и интерпретации</td><td>Задания комбинированного типа, предполагающие выбор одного правильного ответа из предложенных с последующим объяснением своего выбора.
Исследователь создаёт корпус текстов деловой документации начала XXI века. Он собрал 10 000 документов, из которых 9 500 — это приказы, а 500 — договоры и отчёты. Однако в реальной практике того периода договоры и</td></tr>
</table>

Тема	Код и формулировка компетенции	Индикаторы достижения компетенции	Тест по теме
Характеристики корпуса (требования, репрезентативность и пропорциональность	сбора, обработки и интерпретации данных эмпирического исследования; составлять и оформлять научную документацию;	полученных экспериментальных данных.	отчёты составляли 40% всех документов, а приказы — 60%. Какое требование к корпусу нарушено? **Варианты ответов:** А) Репрезентативность Б) Пропорциональность (сбалансированность) В) Машинная читаемость Г) Аннотированность **Ответ: Б** **Обоснование:** Пропорциональность (сбалансированность) означает, что доли разных жанров или типов текстов в корпусе должны соответствовать их долям в генеральной совокупности. В условии договоры и отчёты в реальности составляли 40%, а в корпусе — лишь 5% (500 из 10 000). Это прямое нарушение пропорциональности. ***Задания комбинированного типа, предполагающие выбор одного правильного ответа из предложенных с последующим объяснением своего выбора.*** Лингвист хочет провести квантитативный анализ употребления частицы «бы» в русских художественных текстах XIX века. Какой современный технологический маршрут наиболее адекватен для сбора, обработки и интерпретации данных с гарантией репрезентативности? **Варианты ответов:** А) Найти в интернете 10 случайных романов, вручную выписать все примеры с «бы» и подсчитать их частоту. Б) Скачать из открытых источников готовый частотный словарь русского языка XIX века и взять из него данные по частице «бы».

Тема	Код и формулировка компетенции	Индикаторы достижения компетенции	Тест по теме
			В) Выбрать сбалансированный по жанрам и периодам корпус (например, Национальный корпус русского языка), сформировать подкорпус художественных текстов 1801–1900 гг., использовать корпусного менеджера для получения частот и коллокаций, затем интерпретировать результаты с учётом статистической значимости. Г) Создать собственный корпус из текстов трёх любимых писателей (Пушкин, Гоголь, Толстой), обработать его в AntConc и представить результаты без проверки репрезентативности. **Ответ: В** **Обоснование:** Вариант В соответствует современным стандартам: использование уже готового репрезентативного и пропорционального корпуса (НКРЯ), автоматическая обработка через корпусного менеджера, возможность статистического анализа.
Тема 2. Корпусная лингвистика как самостоятельное лингвистическое направление. Характеристики корпуса (требования, репрезентативность и пропорциональность	**ОПК-6**. Способен применять современные технологии при осуществлении сбора, обработки и интерпретации данных эмпирического исследования; составлять и оформлять научную документацию;	**ИОПК-6.1.** Знает современные технологии сбора, обработки и интерпретации полученных экспериментальных данных.	***Задания комбинированного типа, предполагающие выбор нескольких ответов из предложенных с последующим объяснением своего выбора.*** Выберите все характеристики, обязательные для репрезентативного корпуса. А) Пропорциональность жанров, периодов и типов текстов Б) Только письменные тексты В) Сбалансированная выборка по стратам Г) Объём не менее 1 млн слов Д) Наличие метаданных (автор, дата, жанр) **Ответ: А, В, Д** **Обоснование:** Б — неверно (корпуса могут включать устную и мультимодальную речь). Г — неверно (репрезентативность не зависит от фиксированного объёма, а определяется целью исследования).

<table>
<tr><th>Тема</th><th>Код и формулировка компетенции</th><th>Индикаторы достижения компетенции</th><th>Тест по теме</th></tr>
<tr><td></td><td></td><td></td><td>Задания комбинированного типа, предполагающие выбор нескольких ответов из предложенных с последующим объяснением своего выбора.
Выберите современные технологии / методы, которые обеспечивают пропорциональность при создании корпуса.
А) Стратифицированная выборка по заранее заданным долям
Б) Случайный отбор текстов без учёта жанра
В) Веб-краулеры с фильтрацией по жанровым меткам
Г) Пост-хок статистическое взвешивание частот
Д) Исключительно ручной отбор одним экспертом

Ответ: А, В, Г
Обоснование: Б — нарушает пропорциональность. Д — субъективен и не обеспечивает воспроизводимости пропорций.</td></tr>
<tr><td>Тема 3. Создания корпусов и их технология. Субкорпуса и принципы их создания Корпусный менеджер. Платформа доступа к корпусу.</td><td>ОПК-6. Способен применять современные технологии при осуществлении сбора, обработки и интерпретации данных эмпирического исследования; составлять и оформлять научную</td><td>ИОПК-6.1. Знает современные технологии сбора, обработки и интерпретации полученных экспериментальных данных.</td><td>Задания закрытого типа на установление соответствия.
Установите соответствие между термином (1–4) и его определением (А–Г).
<table>
<tr><td></td><td>Термин</td><td></td><td>Определение</td></tr>
<tr><td>1</td><td>Субкорпус</td><td>А</td><td>Веб-интерфейс для поиска, фильтрации и визуализации корпусных данных</td></tr>
<tr><td>2</td><td>Корпусный менеджер</td><td>Б</td><td>Часть корпуса, выделенная по одному или нескольким метапризнакам (жанр, автор, время)</td></tr>
<tr><td>3</td><td>Платформа доступа</td><td>В</td><td>Программа для построения конкордансов, частотных списков и коллокаций (AntConc, Sketch Engine)</td></tr>
<tr><td>4</td><td>Стратификация</td><td>Г</td><td>Разбиение корпуса на непересекающиеся слои для обеспечения сбалансированности</td></tr>
</table></td></tr>
</table>

<table>
<tr><th>Тема</th><th>Код и формулировка компетенции</th><th>Индикаторы достижения компетенции</th><th>Тест по теме</th></tr>
<tr><td></td><td>документацию;</td><td></td><td>
Ответ:
<table>
<tr><td>1</td><td>2</td><td>3</td><td>4</td></tr>
<tr><td></td><td></td><td></td><td></td></tr>
</table>
Ответ:
<table>
<tr><td>1</td><td>2</td><td>3</td><td>4</td></tr>
<tr><td>Б</td><td>В</td><td>А</td><td>Г</td></tr>
</table>

Задания закрытого типа на установление соответствия.
Установите соответствие между технологическим этапом создания субкорпуса (1–4) и используемым методом/инструментом (А–Г).
<table>
<tr><td></td><td>Технологический этап создания субкорпуса</td><td></td><td>Метод/инструмент</td></tr>
<tr><td>1</td><td>Фильтрация текстов по дате</td><td>А</td><td>Регулярные выражения по полю <date> в метаданных</td></tr>
<tr><td>2</td><td>Извлечение контекстов для указанной лексемы</td><td>Б</td><td>Функция KWIC (Key Word in Context) в корпусном менеджере</td></tr>
<tr><td>3</td><td>Оценка репрезентативности субкорпуса</td><td>В</td><td>Сравнение частот с полным корпусом через log-likelihood</td></tr>
<tr><td>4</td><td>Добавление морфологической разметки</td><td>Г</td><td>Автоматический тэггер (TreeTagger, MyStem)</td></tr>
</table>
Ответ:
<table>
<tr><td>1</td><td>2</td><td>3</td><td>4</td></tr>
<tr><td></td><td></td><td></td><td></td></tr>
</table>
Ответ:
<table>
<tr><td>1</td><td>2</td><td>3</td><td>4</td></tr>
</table>
</td></tr>
</table>

<table>
<tr><th>Тема</th><th>Код и формулировка компетенции</th><th>Индикаторы достижения компетенции</th><th>Тест по теме</th></tr>
<tr><td></td><td></td><td></td><td>
<table>
<tr><td>А</td><td>Б</td><td>В</td><td>Г</td></tr>
</table>
</td></tr>
<tr><td>Тема 3. Создания корпусов и их технология. Субкорпуса и принципы их создания Корпусный менеджер. Платформа доступа к корпусу.</td><td>ОПК-6. Способен применять современные технологии при осуществлении сбора, обработки и интерпретации данных эмпирического исследования; составлять и оформлять научную документацию;</td><td>ИОПК-6.1. Знает современные технологии сбора, обработки и интерпретации полученных экспериментальных данных.</td><td>Задания закрытого типа на установление последовательности
Установите правильную последовательность создания субкорпуса из большого корпуса с помощью корпусного менеджера

Шаги (в случайном порядке):
1. Экспорт полученных данных (конкордансов, частотных списков) для дальнейшего статистического анализа.
2. Загрузка основного корпуса в интерфейс корпусного менеджера (например, Sketch Engine, AntConc).
3. Интерпретация полученных количественных результатов с опорой на лингвистическую задачу.
4. Формулировка критериев для субкорпуса (жанр, хронологический период, автор).
5. Применение фильтров по метаданным для формирования субкорпуса.
6. Запуск квантитативного запроса (частотность, коллокации) внутри созданного субкорпуса.
Ответ:
<table>
<tr><td></td><td></td><td></td><td></td><td></td><td></td></tr>
</table>

Ответ:
<table>
<tr><td>4</td><td>2</td><td>5</td><td>6</td><td>1</td><td>3</td></tr>
</table>

Задания закрытого типа на установление последовательности
Установите последовательность работы с платформой доступа к корпусу для получения и интерпретации экспериментальных данных</td></tr>
</table>

<table>
<tr><th>Тема</th><th>Код и формулировка компетенции</th><th>Индикаторы достижения компетенции</th><th>Тест по теме</th></tr>
<tr><td></td><td></td><td></td><td>Шаги (в случайном порядке):
1. Визуализация частотного распределения в виде графиков (гистограмм, линейчатых диаграмм) на платформе.
2. Выбор целевого корпуса или субкорпуса в веб-интерфейсе платформы.
3. Формулировка поискового запроса (лексема, грамматическая конструкция).
4. Выгрузка результатов в формате CSV или Excel для статистической проверки (хи-квадрат, log-likelihood).
5. Получение конкорданса и частотных списков с указанием нормализованных частот (на млн слов).
6. Лингвистическая интерпретация значимых различий с учётом репрезентативности субкорпуса.
Ответ:
<table><tr><td></td><td></td><td></td><td></td><td></td><td></td></tr></table>
Ответ:
<table><tr><td>3</td><td>2</td><td>5</td><td>1</td><td>4</td><td>6</td></tr></table></td></tr>
<tr><td>Тема 3. Создания корпусов и их технология. Субкорпуса и принципы их создания Корпусный менеджер. Платформа доступа к корпусу.</td><td>ОПК-6. Способен применять современные технологии при осуществлении сбора, обработки и интерпретации данных эмпирического исследования;</td><td>ИОПК-6.1. Знает современные технологии сбора, обработки и интерпретации полученных экспериментальных данных.</td><td>Задания комбинированного типа, предполагающие выбор одного правильного ответа из предложенных с последующим объяснением своего выбора.
Исследователь хочет сравнить употребление глаголов в научных текстах 1990-х и 2010-х годов. Что ему нужно сделать в первую очередь?

А) Загрузить весь корпус в Excel
Б) Создать два субкорпуса по метапризнаку «время»
В) Построить конкорданс для каждого глагола вручную
Г) Установить платформу доступа на локальный сервер</td></tr>
</table>

Тема	Код и формулировка компетенции	Индикаторы достижения компетенции	Тест по теме
	составлять и оформлять научную документацию;		**Ответ: Б** **Обоснование:** Без создания субкорпусов (выборки по хронологическому признаку) невозможно корректно сравнить периоды. Остальные действия либо преждевременны, либо неверны. ***Задания комбинированного типа, предполагающие выбор одного правильного ответа из предложенных с последующим объяснением своего выбора.*** Какой инструмент позволяет получить частотный список лемм в субкорпусе с нормализацией на 1 млн слов? А) Текстовый редактор Б) Корпусный менеджер (AntConc, Sketch Engine) В) Система компьютерной вёрстки Г) Грамматический словарь **Ответ: Б** **Обоснование:** Корпусный менеджер специально предназначен для квантитативной обработки (частоты, коллокации, конкордансы) внутри корпуса или субкорпуса с автоматической нормализацией. Остальные варианты не дают таких возможностей.
Тема 3. Создания корпусов и их технология. Субкорпуса и принципы их создания Корпусный	**ОПК-6**. Способен применять современные технологии при осуществлении сбора, обработки и интерпретации данных	**ИОПК-6.1.** Знает современные технологии сбора, обработки и интерпретации полученных экспериментальных данных.	***Задания комбинированного типа, предполагающие выбор нескольких ответов из предложенных с последующим объяснением своего выбора.*** Выберите все корректные действия при создании субкорпуса с помощью корпусного менеджера. А) Фильтрация по метаданным (жанр, дата, автор) Б) Ручное удаление текстов из основного корпуса В) Использование регулярных выражений для отбора по полю <year> Г) Случайная выборка без учёта пропорций

Тема	Код и формулировка компетенции	Индикаторы достижения компетенции	Тест по теме
менеджер. Платформа доступа к корпусу.	эмпирического исследования; составлять и оформлять научную документацию;		Д) Сохранение субкорпуса как отдельной коллекции **Ответ: А, В, Д** **Обоснование:** А и В — штатные способы фильтрации в корпусных менеджерах; Д — функция сохранения для повторного использования. Б — неавтоматизировано, противоречит технологии. Г — нарушает репрезентативность. ***Задания комбинированного типа, предполагающие выбор нескольких ответов из предложенных с последующим объяснением своего выбора.*** Какие возможности платформы доступа к корпусу (Sketch Engine, НКРЯ) необходимы для интерпретации экспериментальных данных? А) Построение конкордансов (KWIC) Б) Вычисление коллокационной силы (MI, logDice) В) Редактирование исходных текстов корпуса Г) Экспорт частотных списков в CSV Д) Автоматическая запись устной речи **Ответ: А, Б, Г** **Обоснование:** А, Б, Г — стандартные инструменты квантитативной обработки и выгрузки данных для интерпретации. В — не предусмотрено (корпус неизменяем). Д — относится к сбору, а не к платформе доступа.

Критерии оценивания тестового задания:

Оценка	Критерии оценивания
отлично	от 90 до 100 % правильно выполненных заданий
хорошо	от 70 до 89 % правильно выполненных заданий
удовлетворительно	от 50 до 69 % правильно выполненных заданий
неудовлетворительно	менее 50 % правильно выполненных заданий

4.2 ТИПОВЫЕ ПРАКТИЧЕСКИЕ ЗАДАНИЯ

Практические задания должны отражать умение обучающегося применять осваиваемую компетенцию в практических ситуациях и при решении производственных задач (индикаторы УМЕТЬ).

Тема	Код и формулировка компетенции	Индикаторы достижения компетенции	Тест по теме
Тема 1. Современная квантитативная лингвистика как наука. Новые информационные технологии в лингвистических исследованиях и образовательном процессе.	ОПК-6. Способен применять современные технологии при осуществлении сбора, обработки и интерпретации данных эмпирического исследования; составлять и оформлять научную документацию;	ИОПК-6.2. Умеет применять современные технологии сбора, обработки и интерпретации полученных экспериментальных данных.	***Задания комбинированного типа, предполагающие выбор одного правильного ответа из предложенных с последующим объяснением своего выбора.*** Лингвист получил массив ответов испытуемых (время реакции, точность). Какой инструмент оптимален для проверки статистической значимости различий между двумя группами? А) Текстовый редактор Б) Пакет SPSS или R (t-критерий) В) Программа для записи экрана Г) Онлайн-анкета без экспорта **Ответ: Б** **Обоснование**: SPSS и R позволяют применить параметрические критерии (t-тест) для интерпретации экспериментальных данных. Остальные варианты не предназначены для статистической обработки. ***Задания комбинированного типа, предполагающие выбор одного правильного ответа из предложенных с последующим объяснением своего выбора.*** В эксперименте по восприятию редких слов нужно зафиксировать микродвижения глаз. Какую технологию следует применить? А) Корпусный менеджер Б) Айтрекинг (eye-tracker) В) Опрос в Google Forms

			Г) Частотный словарь **Ответ: Б** **Обоснование**: Айтрекинг собирает данные о фиксациях, саккадах и зрачках в реальном времени. Корпусные и опросные методы не дают таких физиологических показателей.
Тема 1. Современная квантитативная лингвистика как наука. Новые информационные технологии в лингвистических исследованиях и образовательном процессе.	ОПК-6. Способен применять современные технологии при осуществлении сбора, обработки и интерпретации данных эмпирического исследования; составлять и оформлять научную документацию;	ИОПК-6.3. Владеет навыками сбора, обработки и интерпретации полученных экспериментальных данных.	***Задания комбинированного типа, предполагающие выбор нескольких ответов из предложенных с последующим объяснением своего выбора.*** Выберите все технологии, которые используются для сбора экспериментальных данных в квантитативной лингвистике. А) Айтрекинг Б) Корпусный менеджер (AntConc) В) Платформа онлайн-экспериментов (PsychoPy, Pavlovia) Г) Пакет статистического анализа R Д) Краудсорсинговые платформы (Toloka, Prolific) **Ответ: А, В, Д** **Обоснование:** Айтрекинг (А) фиксирует физиологические реакции, PsychoPy (В) собирает поведенческие данные (время, точность), краудсорсинг (Д) — массовые оценки и разметку. AntConc (Б) и R (Г) — инструменты обработки, а не сбора. ***Задания комбинированного типа, предполагающие выбор нескольких ответов из предложенных с последующим объяснением своего выбора.*** Какие действия входят в корректную обработку и интерпретацию количественных лингвистических данных после проведения эксперимента? А) Проверка нормальности распределения Б) Визуализация данных (ящик с усами, гистограмма)

			В) Автоматический выбор t-критерия без проверки условий Г) Расчёт описательной статистики (среднее, медиана, σ) Д) Интерпретация p-value в отрыве от лингвистической теории **Ответ: А, Б, Г** **Обоснование:** Проверка нормальности (А), визуализация (Б) и описательная статистика (Г) — обязательные этапы современной обработки. В — ошибочен (нужна проверка условий). Д — неверен (статистика должна интерпретироваться лингвистически).
Тема 1. Современная квантитативная лингвистика как наука. Новые информационные технологии в лингвистических исследованиях и образовательном процессе.	ОПК-6. Способен применять современные технологии при осуществлении сбора, обработки и интерпретации данных эмпирического исследования; составлять и оформлять научную документацию;	ИОПК-6.2. Умеет применять современные технологии сбора, обработки и интерпретации полученных экспериментальных данных.	***Задания открытого типа с развернутым ответом.*** Исследователь проводит эксперимент: испытуемые читают предложения с редкими словами, а компьютер фиксирует точки остановки взгляда и движения зрачка. Какая технология сбора данных здесь используется? Ответ_____ **Ответ: Айтрекинг** ***Задания открытого типа с развернутым ответом.*** Лингвист получил числовые данные о частоте употребления конструкции в разных жанрах и хочет выполнить кластерный анализ и построить графики. Какой бесплатный язык программирования и среда статистических вычислений ему следует применить? Ответ:_____ **Ответ: R**
Тема 2.	ОПК-6. Способен	ИОПК-6.2. Умеет	***Задания комбинированного типа, предполагающие выбор одного***

Корпусная лингвистика как самостоятельное лингвистическое направление. Характеристики корпуса (требования, репрезентативность и пропорциональность	применять современные технологии при осуществлении сбора, обработки и интерпретации данных эмпирического исследования; составлять и оформлять научную документацию;	применять современные технологии сбора, обработки и интерпретации полученных экспериментальных данных.	***правильного ответа из предложенных с последующим объяснением своего выбора.*** Корпус содержит 90% газетных текстов и 10% научных, хотя в реальной речи их соотношение 50/50. Какое требование нарушено? А) Аннотированность Б) Пропорциональность В) Машинная читаемость Г) Динамичность **Ответ: Б** **Обоснование:** Нарушена пропорциональность (сбалансированность) – доли жанров не соответствуют генеральной совокупности. ***Задания комбинированного типа, предполагающие выбор одного правильного ответа из предложенных с последующим объяснением своего выбора.*** Какой инструмент позволяет автоматически получить коллокации слова «свобода» в субкорпусе публицистики 2000-х годов? А) Текстовый редактор Б) Корпусный менеджер (AntConc, Sketch Engine) В) Система распознавания речи Г) Грамматический словарь **Ответ: Б** **Обоснование**: Корпусный менеджер предназначен для построения конкордансов, частотных списков и коллокаций с учётом метаданных субкорпуса.
Тема 2. Корпусная лингвистика как самостоятельное	ОПК-6. Способен применять современные	ИОПК-6.3. Владеет навыками сбора, обработки и	***Задания комбинированного типа, предполагающие выбор нескольких ответов из предложенных с последующим объяснением своего выбора.*** Выберите условия, необходимые для репрезентативности корпуса.

лингвистическое направление. Характеристики корпуса (требования, репрезентативность и пропорциональность	технологии при осуществлении сбора, обработки и интерпретации данных эмпирического исследования; составлять и оформлять научную документацию;	интерпретации полученных экспериментальных данных.	А) Соблюдение пропорций жанров, периодов, типов речи Б) Только письменные тексты В) Объём более 10 млн слов Г) Наличие метаданных для стратификации Д) Случайная выборка внутри каждой страты **Ответ: А, Г, Д** **Обоснование:** А (пропорции), Г (метаданные), Д (случайная выборка) обеспечивают репрезентативность. Б и В – необязательны. ***Задания комбинированного типа, предполагающие выбор нескольких ответов из предложенных с последующим объяснением своего выбора.*** Какие современные технологии позволяют оценить пропорциональность корпуса? А) Сравнение частот с эталонным корпусом (хи-квадрат) Б) Построение конкордансов для одного слова В) Статистический анализ распределения метапризнаков Г) Визуализация страт через диаграммы Д) Ручной подсчёт документов **Ответы: А, В, Г** **Обоснование:** А (статсравнение), В (анализ распределения), Г (визуализация) – автоматизированные методы. Б – не оценивает пропорции. Д – несовременная технология.
Тема 2. Корпусная лингвистика как самостоятельное лингвистическое направление.	ОПК-6. Способен применять современные технологии при осуществлении сбора,	ИОПК-6.2. Умеет применять современные технологии сбора, обработки и	***Задания открытого типа с развернутым ответом.*** Вы формируете корпус русских художественных текстов XX века: включаете 80% прозы и 20% поэзии, хотя в реальной литературе этого периода соотношение 50% на 50%. Какое требование к корпусу нарушено?

Характеристики корпуса (требования, репрезентативность и пропорциональность	обработки и интерпретации данных эмпирического исследования; составлять и оформлять научную документацию;	интерпретации полученных экспериментальных данных.	Ответ_____ ***Ответ: пропорциональность*** ***Задания открытого типа с развернутым ответом.*** Вам нужно автоматически получить список коллокатов слова «исследование» с указанием силы связи (MI score) в подкорпусе научных статей. Какой тип программы для этого необходим? Ответ_____ **Ответ: конкордансер**
Тема 3. Создания корпусов и их технология. Субкорпуса и принципы их создания Корпусный менеджер. Платформа доступа к корпусу.	ОПК-6. Способен применять современные технологии при осуществлении сбора, обработки и интерпретации данных эмпирического исследования; составлять и оформлять научную документацию;	ИОПК-6.2. Умеет применять современные технологии сбора, обработки и интерпретации полученных экспериментальных данных.	***Задания комбинированного типа, предполагающие выбор одного правильного ответа из предложенных с последующим объяснением своего выбора.*** Вы создаёте субкорпус для анализа эволюции термина «искусственный интеллект» в научных статьях 1970–2020 гг. Какой принцип создания субкорпуса вы примените в первую очередь? А) Фильтрацию по полю «год публикации» в метаданных Б) Случайную выборку 1000 статей без учёта даты В) Построение конкорданса по всем словам Г) Ручной отбор статей по названиям **Ответ: А** **Обоснование:** Для анализа изменений во времени необходимо сначала выделить субкорпус по хронологическому признаку через фильтрацию метаданных, что позволяет далее сравнивать разные периоды. ***Задания комбинированного типа, предполагающие выбор одного правильного ответа из предложенных с последующим объяснением своего выбора.*** Вы завершили обработку данных в корпусном менеджере и должны

			представить результаты в виде научной статьи. Что необходимо включить в раздел «Методология»? А) Только список использованной литературы Б) Название корпуса, версию, критерии формирования субкорпуса, название корпусного менеджера и статистические методы В) Все полученные конкордансы целиком Г) Графики без указания источника данных **Ответ: Б** **Обоснование**: Научная документация требует указания корпуса, субкорпуса, инструмента (корпусного менеджера) и методов анализа для воспроизводимости исследования.
Тема 3. Создания корпусов и их технология. Субкорпуса и принципы их создания Корпусный менеджер. Платформа доступа к корпусу.	ОПК-6. Способен применять современные технологии при осуществлении сбора, обработки и интерпретации данных эмпирического исследования; составлять и оформлять научную документацию;	ИОПК-6.3. Владеет навыками сбора, обработки и интерпретации полученных экспериментальных данных.	***Задания комбинированного типа, предполагающие выбор нескольких ответов из предложенных с последующим объяснением своего выбора.*** Вы готовите корпусное исследование диахронических изменений в научном стиле. Какие действия необходимо выполнить при создании субкорпуса для эмпирического исследования? А) Определить метапризнаки (жанр, время, автор) Б) Использовать корпусный менеджер для фильтрации В) Включить все тексты основного корпуса без изменений Г) Проверить репрезентативность субкорпуса Д) Оформить описание субкорпуса в научной документации **Ответ: А, Б, Г, Д** **Обоснование:** В (включение всех текстов) противоречит созданию субкорпуса. Остальные – обязательные шаги. ***Задания комбинированного типа, предполагающие выбор нескольких ответов из предложенных с последующим объяснением своего выбора.***

			Вы завершили анализ и готовите научную статью по результатам корпусного исследования. Какие элементы должны быть указаны в разделе «Материалы и методы» научной статьи по корпусному исследованию? А) Название и версия корпуса Б) Логин и пароль от платформы доступа В) Критерии формирования субкорпуса Г) Название корпусного менеджера и версия Д) Статистические критерии для интерпретации **Ответы: А, В, Г, Д** **Объяснение:** Б (логин и пароль) не относится к научной документации. Остальные – необходимы для воспроизводимости.
Тема 3. Создания корпусов и их технология. Субкорпуса и принципы их создания Корпусный менеджер. Платформа доступа к корпусу.	ОПК-6. Способен применять современные технологии при осуществлении сбора, обработки и интерпретации данных эмпирического исследования; составлять и оформлять научную документацию;	ИОПК-6.2. Умеет применять современные технологии сбора, обработки и интерпретации полученных экспериментальных данных.	***Задания открытого типа с развернутым ответом.*** Вы создаёте субкорпус для исследования эволюции юридической терминологии. Вы отобрали тексты законов только за 2020 год, хотя реальные изменения происходили с 1990 по 2020 год. Какой принцип создания субкорпуса нарушен? Ответ:_____ **Ответ: репрезентативность** ***Задания открытого типа с развернутым ответом.*** Вы провели анализ в корпусном менеджере и теперь оформляете раздел «Материалы и методы» для научной статьи. Как называется документ, в котором фиксируются все параметры корпуса, субкорпуса, инструментов и статистических критериев для обеспечения воспроизводимости исследования? Ответ:_____ **Ответ: протокол**

Критерии оценивания практических занятий:

Оценка	Критерии оценивания
отлично	Выставляется, если обучающийся умеет увязывать теорию с практикой (решает задачи, формулирует выводы, умеет пояснить полученные результаты), владеет понятийным аппаратом, полно и глубоко овладел материалом по заданной теме, обосновывает свои суждения и даёт правильные ответы на вопросы преподавателя
хорошо	Выставляется, если обучающийся умеет увязывать теорию с практикой (решает задачи и формулирует выводы, умеет пояснить полученные результаты), владеет понятийным аппаратом, полно и глубоко овладел материалом по заданной теме, но содержание ответов имеют некоторые неточности и требуют уточнения и комментария со стороны преподавателя
удовлетворительно	Выставляется, если обучающийся знает и понимает материал по заданной теме, но изложение неполное, непоследовательное, допускаются неточности в определении понятий, студент не может обосновать свои ответы на уточняющие вопросы преподавателя
неудовлетворительно	Выставляется, если обучающийся допускает ошибки в определении понятий, искажающие их смысл, беспорядочно и неуверенно излагает материал. Делает ошибки в ответах на уточняющие вопросы преподавателя

4.3. ТИПОВЫЕ ЗАДАНИЯ ДЛЯ КОНТРОЛЬНЫХ РАБОТ

Контрольные работы содержат несколько практических заданий по индивидуальным вариантам, в полном объеме охватывающих изученный материал по указанной теме (индикаторы УМЕТЬ). Выполнение контрольных работ позволяет определить результат освоения компетенций по дисциплине в рамках рассматриваемой темы, оцениваемый с помощью соответствующих индикаторов достижения компетенций.

Тема	**Код и формулировка компетенции**	**Индикаторы достижения компетенции**	**Типовой вариант контрольной работы**	
			Вариант 1	**Вариант 2**
Тема 1. Современная квантитативная лингвистика как наука. Новые информационные технологии в лингвистических исследованиях и образовательном процессе. Тема 2. Корпусная лингвистика как самостоятельное лингвистическое направление. Характеристики корпуса (требования, репрезентативность и пропорциональность Тема 3. Создания	ОПК-6. Способен применять современные технологии при осуществлении сбора, обработки и интерпретации данных эмпирического исследования; составлять и оформлять научную документацию;	ИОПК-6.2. Умеет применять современные технологии сбора, обработки и интерпретации полученных экспериментальных данных.	***Задание 1.*** Рассчитать индекс лексического разнообразия (TTR = Types/Tokens) для текста из 200 слов, содержащего 85 уникальных словоформ. **Ответ**: TTR = 85/200 = 0,425 (или 42,5%). Вывод: умеренное лексическое разнообразие. ***Задание 2.*** Составить мини-алгоритм проверки пропорциональности корпуса при сборе данных из соцсетей. **Ответ:** Определить параметры выборки	***Задание 1.*** Предложить алгоритм обработки корпуса текстов для выявления частотности неологизмов с использованием современных ИТ-инструментов. **Ответ:** 1. Загрузка корпуса в Python (NLTK/Spacy) → 2. Лемматизация и фильтрация по словарю → 3. Выделение слов, отсутствующих в базовых лексиконах → 4. Частотный анализ → 5. Визуализация (matplotlib/WordCloud). ***Задание 2.*** Оценить репрезентативность корпуса из 1 млн слов, если 90% текстов — художественная литература, а цель исследования — анализ современной разговорной речи.

корпусов и их технология. Субкорпуса и принципы их создания Корпусный менеджер. Платформа доступа к корпусу.			(платформа, возраст, регион) → Задать квоты по каждому параметру → Парсинг данных (Python: BeautifulSoup/API) → Контроль пропорций в реальном времени → Статистическая проверка (χ^2-тест). ***Задание 3***. Сформировать субкорпус для исследования молодежного сленга из национального корпуса (2015–2024 гг.). **Ответ:** Фильтры: возраст авторов 14–25 лет, жанр: соцсети/блоги, дата: 2015–2024 → Экспорт метаданных → Проверка объёма (мин. 100 тыс. слов) → Сохранение как субкорпуса в формате XML/TEI.	**Ответ:** Корпус нерепрезентативен: дисбаланс жанров искажает результаты. Для изучения разговорной речи нужна пропорция: ≥40% устной/неформальной речи, сбалансированная по возрасту, региону, полу говорящих. ***Задание 3.*** Извлечь частотный список коллокаций для слова «краш» через корпусный менеджер. **Ответ:** Загрузить субкорпус в платформу (Sketch Engine) → Запрос: word="краш" & pos =".*" → Функция «Collocations» (MI-score ≥ 3, window 5L/5R) → Экспорт результатов (.csv) → Интерпретация: визуализация в WordCloud.

Критерии оценивания контрольной работы:

Оценка	Критерии оценивания
отлично	Выставляется, если обучающийся умеет увязывать теорию с практикой (решает задачи, формулирует выводы, умеет пояснить полученные результаты), владеет понятийным аппаратом, полно и глубоко овладел материалом по заданной теме, обосновывает свои суждения и даёт

	правильные ответы на вопросы преподавателя
хорошо	Выставляется, если обучающийся умеет увязывать теорию с практикой (решает задачи и формулирует выводы, умеет пояснить полученные результаты), владеет понятийным аппаратом, полно и глубоко овладел материалом по заданной теме, но содержание ответов имеют некоторые неточности и требуют уточнения и комментария со стороны преподавателя.
удовлетворительно	Выставляется, если обучающийся знает и понимает материал по заданной теме, но изложение неполное, непоследовательное, допускаются неточности в определении понятий, студент не может обосновать свои ответы на уточняющие вопросы преподавателя
неудовлетворительно	Выставляется, если обучающийся допускает ошибки в определении понятий, искажающие их смысл, беспорядочно и неуверенно излагает материал. Делает ошибки в ответах на уточняющие вопросы преподавателя

4.4. ТИПОВЫЕ ЗАДАНИЯ ДЛЯ САМОСТОЯТЕЛЬНОЙ РАБОТЫ

Самостоятельная работа (индикаторы ЗНАТЬ, УМЕТЬ – на выбор) включает в себя проработку теоретического материала, изучение рекомендуемой литературы, выполнение практико-ориентированных заданий (заполнение таблиц, проведение сравнительного анализа, составление схем и др.), решение практических задач, создание презентаций, написание рефератов, подборка нормативного и иного материала и выполнение других заданий.

Тема	**Код и формулировка компетенции**	**Индикаторы достижения компетенции**	**Типовое задание для самостоятельной работы**
Тема 1. Современная квантитативная лингвистика как наука. Новые информационные технологии в лингвистических исследованиях и образовательном процессе.	ОПК-6. Способен применять современные технологии при осуществлении сбора, обработки и интерпретации данных эмпирического исследования; составлять и оформлять научную документацию;	ИОПК-6.1. Знает современные технологии сбора, обработки и интерпретации полученных экспериментальных данных. ИОПК-6.2. Умеет применять современные технологии сбора,	***Задание. Подготовить рефераты и презентации по теме: «Современная квантитативная лингвистика как наука. Новые информационные технологии в лингвистических исследованиях и образовательном процессе».*** 1. Методы математической статистики в анализе языковых данных: от частотных словарей к машинному обучению. 2. Корпусные технологии в обучении иностранным языкам: инструментарий и педагогическая эффективность. 3. Автоматизация сбора и обработки экспериментальных лингвистических данных: алгоритмы и программные среды. 4. Визуализация квантитативных результатов в лингвистике: современные инструменты и методики интерпретации. 5. Применение нейросетевых моделей в количественном анализе текста: возможности и ограничения.

<table>
<tr><th>Тема</th><th>Код и формулировка компетенции</th><th>Индикаторы достижения компетенции</th><th>Типовое задание для самостоятельной работы</th></tr>
<tr><td></td><td></td><td>обработки и интерпретации полученных экспериментальных данных.</td><td>6. Обеспечение репрезентативности лингвистических корпусов: цифровые методы балансировки данных.
7. Языки программирования Python и R в квантитативной лингвистике: сравнительный обзор применения.
8. Цифровые платформы для создания учебных субкорпусов в лингвистическом образовании.
9. Статистическое моделирование языковой динамики на основе больших данных: методы верификации результатов.
10. Интеграция квантитативных методов и ИТ-инструментов в образовательный процесс: от теории к практике.</td></tr>
<tr><td>Тема 2. Корпусная лингвистика как самостоятельное лингвистическое направление. Характеристики корпуса (требования, репрезентативность и пропорциональность</td><td>ОПК-6. Способен применять современные технологии при осуществлении сбора, обработки и интерпретации данных эмпирического исследования; составлять и оформлять научную документацию;</td><td>ИОПК-6.1. Знает современные технологии сбора, обработки и интерпретации полученных экспериментальных данных.

ИОПК-6.2. Умеет применять современные технологии сбора, обработки и интерпретации полученных экспериментальных данных.</td><td>Задание на сравнительный анализ: Корпусная лингвистика. Характеристики корпуса
Задание: Сравнить два корпуса по критериям репрезентативности и пропорциональности:
<table>
<tr><td>Параметр</td><td>Корпус А (НКРЯ)</td><td>Корпус Б (соцсети, парсинг)</td></tr>
<tr><td>Объём</td><td>200 млн слов</td><td>5 млн слов</td></tr>
<tr><td>Жанры</td><td>12 типов, сбалансировано</td><td>1 жанр (посты)</td></tr>
<tr><td>Метаданные</td><td>Пол, возраст, регион, дата</td><td>Только дата</td></tr>
<tr><td>Обработка</td><td>Ручная разметка + автопроверка</td><td>Только токенизация</td></tr>
</table>
Вопрос: Какой корпус репрезентативнее для исследования общеязыковой лексики и почему? Какие ИТ-инструменты улучшат Корпус Б?
Ответ:
Репрезентативнее Корпус А: сбалансирован по жанрам, имеет полные метаданные, верифицированную разметку.
Улучшение Корпуса Б:</td></tr>
</table>

Тема	Код и формулировка компетенции	Индикаторы достижения компетенции	Типовое задание для самостоятельной работы
		ИОПК-6.3. Владеет навыками сбора, обработки и интерпретации полученных экспериментальных данных.	Добавить квоты по жанрам/демографии (pandas) → Расширить метаданные (API соцсетей) → Автоматическая лемматизация (Spacy/MyStem) → Валидация пропорций (χ^2-тест в R/Python).
Тема 3. Создания корпусов и их технология. Субкорпуса и принципы их создания Корпусный менеджер. Платформа доступа к корпусу.	ОПК-6. Способен применять современные технологии при осуществлении сбора, обработки и интерпретации данных эмпирического исследования; составлять и оформлять научную документацию;	ИОПК-6.1. Знает современные технологии сбора, обработки и интерпретации полученных экспериментальных данных. ИОПК-6.2. Умеет применять современные технологии сбора, обработки и интерпретации полученных экспериментальных данных. ИОПК-6.3. Владеет навыками сбора, обработки и интерпретации	***Практическое задание: Создание корпусов, субкорпуса, корпусный менеджер*** Создать субкорпус «Научные статьи по лингвистике (2020–2025)» из открытого корпуса и извлечь топ-10 ключевых терминов. **Алгоритм выполнения:** 1. Выбрать платформу (Sketch Engine / NoSketch Engine / AntConc). 2. Задать фильтры: жанр = «научная статья», тематика = «лингвистика», дата = 2020–2025. 3. Сформировать субкорпус → экспортировать метаданные. 4. Запустить частотный анализ с фильтрацией стоп-слов. 5. Построить список ключевых слов (KeyWords, MI-score ≥ 3). 6. Визуализировать результаты (WordCloud / matplotlib). **Ответ:** • Субкорпус сохранён в формате TEI/XML, объём ~150 тыс. слов. • Топ-10 терминов: корпус, дискурс, лексема, семантика, прагматика, грамматика, синтаксис, морфема, деривация, коллокация. • Интерпретация: ядро терминосферы соответствует современной квантитативной лингвистике.

Тема	Код и формулировка компетенции	Индикаторы достижения компетенции	Типовое задание для самостоятельной работы
		полученных экспериментальных данных.	

Критерии оценивания самостоятельной работы:

Оценка	Критерии оценивания
отлично	выставляется если работа носит научно-исследовательский характер, проанализирован и сделан сравнительный анализ нескольких литературных источников, приведены примеры
хорошо	выставляется если проанализирован и сделан сравнительный анализ нескольких литературных источников, приведены примеры
удовлетворительно	выставляется если проведен сравнительный анализ научно-методической литературы, приведены примеры
неудовлетворительно	выставляется если работа прошла проверку на антиплагиат и соответствует требованиям оформления

6 МАТЕРИАЛЫ ДЛЯ ДИАГНОСТИЧЕСКОЙ РАБОТЫ

Задания для диагностической работы должны обеспечивать оценку полностью или частично сформированных компетенций. Каждое задание должно быть привязано к тому или иному индикатору сформированности компетенций.

При формировании заданий для диагностической работы необходимо использовать тестовые задания следующих типов:

Тип задания 1. Задания закрытого типа на установление соответствия.

Тип задания 2. Задания закрытого типа на установление последовательности.

Тип задания 3. Задания комбинированного типа, предполагающие выбор одного правильного ответа из предложенных с последующим объяснением своего выбора.

Тип задания 4. Задания комбинированного типа, предполагающие выбор нескольких ответов из предложенных с последующим объяснением своего выбора.

Тип задания 5. Задания открытого типа с развернутым ответом.

Все типы заданий должны быть представлены не менее одного раза.

<table>
<tr><th>№ п/п</th><th>Тема занятия</th><th>Код компетенции</th><th>Индикатор</th><th>Тип задания</th><th colspan="2">Задание</th></tr>
<tr><th colspan="5"></th><th>Вариант 1</th><th>Вариант 2</th></tr>
<tr><td>1.</td><td>Тема 1. Современная квантитативная лингвистика как наука. Новые информационные технологии в лингвистических исследованиях и образовательном процессе.</td><td>ОПК-6. Способен применять современные технологии при осуществлении сбора, обработки и интерпретации данных эмпирического исследования; составлять и оформлять научную документацию;</td><td>Тема 1. Современная квантитативная лингвистика как наука. Новые информационные технологии в лингвистических исследованиях и образовательном</td><td>1</td><td>Соотнесите каждый метод или технологию сбора данных (1–4) с соответствующим описанием (А–Г). Запишите выбранные буквы рядом с цифрами.
<table>
<tr><th></th><th>Метод / технология</th><th></th><th>Описание</th></tr>
<tr><td>1</td><td>Корпусная лингвистика</td><td>А</td><td>Сбор данных о движении глаз испытуемого при чтении или восприятии речи для изучения когнитивных процессов и языковой</td></tr>
</table></td><td>Соотнесите каждый программный инструмент (1–4) с его основным функциональным назначением (А–Г). Запишите выбранные буквы рядом с цифрами.
<table>
<tr><th></th><th>Инструмент</th><th></th><th>Назначение</th></tr>
<tr><td>1</td><td>AntConc</td><td>А</td><td>Среда статистического программирования (включая пакеты tidyverse, ggplot2), позволяющая выполнять многомерный анализ (регрессию, кластеризацию) и строить сложные</td></tr>
</table></td></tr>
</table>

			процессе.					обработки.				визуализации.
					2	Онлайн-анкетирование (Google Forms, Яндекс.Взгляд)	Б	Автоматизированный сбор и разметка больших массивов текстов с последующим квантитативным анализом частотности грамматических и лексических единиц.	2	SPSS	Б	Построение конкордансов, вычисление ключевых слов (keyness), частотных списков и коллокаций на основе корпусных данных.
					3	Айтрекинг (eye-tracking)	В	Привлечение добровольцев через интернет-платформы (например, Amazon Mechanical Turk) для выполнения микро-заданий по разметке, оценке или сбору языкового материала.	3	R	В	Программный пакет для статистического анализа с графическим интерфейсом, широко используемый в социальных науках
					4	Краудсорсинг	Г	Автоматизированный сбор структурированных ответов респондентов с возможностью экспорта данных в форматы, пригодные для	4	TXM	Г	Платформа для корпусного анализа с инструментами статистической обработки и визуализации (факторный анализ, иерархическая классификация, кластеризация текстов).

Ответ:

1	2	3	4

Ответ:

<table>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>
<table>
<tr><td></td><td></td><td></td><td>статистической обработки.</td></tr>
</table>
Ответ:
<table>
<tr><td>1</td><td>2</td><td>3</td><td>4</td></tr>
<tr><td></td><td></td><td></td><td></td></tr>
</table>
Ответ:
<table>
<tr><td>1</td><td>2</td><td>3</td><td>4</td></tr>
<tr><td>Б</td><td>Г</td><td>А</td><td>В</td></tr>
</table>
</td>
<td>
<table>
<tr><td>1</td><td>2</td><td>3</td><td>4</td></tr>
<tr><td>Б</td><td>В</td><td>А</td><td>Г</td></tr>
</table>
</td>
</tr>
<tr>
<td>2.</td>
<td>Тема 1. Современная квантитативная лингвистика как наука. Новые информационные технологии в лингвистических исследованиях и образовательном процессе.</td>
<td>ОПК-6. Способен применять современные технологии при осуществлении сбора, обработки и интерпретации данных эмпирического исследования; составлять и оформлять научную документацию;</td>
<td>ИОПК-6.1. Знает современные технологии сбора, обработки и интерпретации полученных экспериментальных данных.</td>
<td>2</td>
<td>Расположите перечисленные этапы проведения экспериментального квантитативно-лингвистического исследования с использованием автоматизированного сбора данных в логическом порядке от начала к концу исследования. Запишите цифры в правильной последовательности.

1. Статистическая обработка полученных данных и проверка гипотез (например, в R или SPSS).
2. Визуализация результатов и их лингвистическая интерпретация.
3. Сбор экспериментальных данных с помощью специализированного программного обеспечения (например, PsychoPy, OpenSesame).
4. Формулировка исследовательского вопроса и операционализация переменных.
5. Разработка дизайна эксперимента, создание стимульного материала.</td>
<td>Расположите технологические шаги обработки текстовых корпусных данных с использованием методов квантитативной лингвистики в правильном порядке. Запишите цифры в нужной последовательности.

1. Токенизация и лемматизация текстов с помощью автоматических процессоров (например, TreeTagger, SpaCy).
2. Вычисление статистических показателей (частотность, коллокации, меры разнообразия лексики).
3. Сбор и формирование корпуса текстов (скачивание, очистка от шума, приведение к единому формату).
4. Построение конкордансов и кластерный анализ для выявления паттернов.
5. Разметка метаданных (жанр, автор,</td>
</tr>
</table>

					6. Экспорт сырых данных в формат, пригодный для анализа (CSV, JSON). Ответ: | | | | | | | | **Ответ:** | **4** | **5** | **3** | **6** | **1** | **2** |	дата) для последующего сравнения групп текстов. 6. Интерпретация результатов с опорой на лингвистическую теорию. Ответ: | | | | | | | | **Ответ:** | **3** | **5** | **1** | **2** | **4** | **6** |
3.	Тема 1. Современная квантитативная лингвистика как наука. Новые информационные технологии в лингвистических исследованиях и образовательном процессе.	**ОПК-6**. Способен применять современные технологии при осуществлении сбора, обработки и интерпретации данных эмпирического исследования; составлять и оформлять научную документацию;	**ИОПК-6.1.** Знает современные технологии сбора, обработки и интерпретации полученных экспериментальных данных.	*3*	Исследователь хочет изучить, как носители языка воспринимают редкие лексические единицы в потоке речи в реальном времени. Необходимо зафиксировать не только правильность понимания, но и микропроцессы: латентное время реакции, фиксации взгляда на ключевых словах, зрачковую реакцию. Какая технология наиболее адекватна для этих целей? **Варианты ответов:** 1. Корпусный анализ с частотным словарём 2. Онлайн-анкетирование в Google Forms 3. Айтрекинг (eye-tracking) с синхронной регистрацией времени реакции 4. Краудсорсинг на платформе Toloka **Ответ: 3**	Лингвист провёл эксперимент с использованием психолингвистической базы данных и получил таблицу с числовыми значениями. Какой порядок действий наиболее соответствует современным стандартам обработки и интерпретации данных в квантитативной лингвистике? **Варианты ответов:** 1. Сразу построить красивые графики → подобрать статистический критерий → выполнить расчёты → сформулировать выводы 2. Выполнить описательную статистику (средние, стандартные отклонения) → проверить нормальность распределения → выбрать параметрический или непараметрический критерий →

					Обоснование: Айтрекинг позволяет регистрировать движение глаз, размер зрачка и фиксации на конкретных стимулах в реальном времени, а при синхронизации с кнопочным устройством — и время реакции. Это даёт количественные данные о когнитивных процессах при восприятии редких слов.	вычислить значимость → визуализировать → интерпретировать с опорой на теорию 3. Ввести данные в SPSS → нажать «Анализ» → выбрать t-критерий → записать p-value → сделать вывод без визуализации 4. Сразу применить многомерное шкалирование → получить кластеры → описать их содержательно → опубликовать результаты **Ответ: 2** **Обоснование:** Современный стандарт обработки количественных лингвистических данных требует сначала исследовать распределение (описательная статистика, проверка нормальности), затем обоснованно выбрать статистический критерий, вычислить значимость, после чего визуализировать результаты и дать лингвистическую интерпретацию.
4.	Тема 1. Современная квантитативная лингвистика как наука. Новые информационные технологии	**ОПК-6**. Способен применять современные технологии при осуществлении сбора, обработки и	**ИОПК-6.1.** Знает современные технологии сбора, обработки и интерпретации	*4*	Какие из перечисленных технологий и инструментов относятся к современным методам сбора первичных экспериментальных данных в квантитативной лингвистике? Выберите все верные варианты и объясните, почему остальные не подходят.	Какие этапы и подходы являются обязательными для корректной интерпретации количественных лингвистических данных при использовании современных информационных технологий? Выберите все правильные позиции и кратко обоснуйте, почему без них

	в лингвистических исследованиях и образовательном процессе.	интерпретации данных эмпирического исследования; составлять и оформлять научную документацию;	полученных экспериментальных данных.		**Варианты ответов:** 1. Айтрекинг (eye-tracking) 2. Корпусный менеджер AntConc 3. Программа для онлайн-экспериментов PsychoPy 4. Статистический пакет SPSS 5. Платформа краудсорсинга (Yandex.Toloka, Amazon Mechanical Turk) Ответ: 1, 3, 5 Обоснование: 1. Айтрекинг (1) — технология сбора данных о движении глаз, времени фиксации, зрачковых реакциях в реальном времени. 2. PsychoPy (3) — программная среда для создания и проведения психолингвистических экспериментов с регистрацией реакций (точность, время ответа). 3. Краудсорсинг (5) — позволяет собирать размеченные данные, оценки естественности, переводы и т.д. от множества участников через интернет.	интерпретация будет неполной или ошибочной. **Варианты ответов:** А) Построение описательной статистики (среднее, медиана, стандартное отклонение) Б) Визуализация данных (гистограммы, ящики с усами, графики разброса) В) Проверка статистических допущений (нормальность, гомогенность дисперсий) Г) Автоматический выбор любого критерия по умолчанию (например, t-теста) без проверки условий Д) Интерпретация значимости результатов в контексте лингвистической теории **Ответ: А, Б, В, Д** **Обоснование:** А (описательная статистика) — даёт первичное понимание центральной тенденции и разброса, без неё невозможно оценить характер данных. Б (визуализация) — позволяет обнаружить выбросы, асимметрию, мультимодальность, которые не видны в числах. В (проверка допущений) — необходима для выбора корректного статистического критерия

<table>
<tr><td></td><td></td><td></td><td></td><td></td><td></td><td>(параметрический / непараметрический), иначе интерпретация p-value может быть ложной.
Д (лингвистическая интерпретация) — превращает статистические паттерны в осмысленный вывод о языке; без неё анализ остаётся чисто математическим.</td></tr>
<tr><td>5.</td><td>Тема 2. Корпусная лингвистика как самостоятельное лингвистическое направление. Характеристики корпуса (требования, репрезентативность и пропорциональность</td><td>ОПК-6. Способен применять современные технологии при осуществлении сбора, обработки и интерпретации данных эмпирического исследования; составлять и оформлять научную документацию;</td><td>ИОПК-6.3. Владеет навыками сбора, обработки и интерпретации полученных экспериментальных данных..</td><td>1</td><td>Соотнесите каждый принцип / требование к лингвистическому корпусу (1–4) с его правильным определением или описанием (А–Г).
<table><tr><td></td><td>Принцип / требование</td><td></td><td>Описание</td></tr><tr><td>1</td><td>Репрезентативность</td><td>А</td><td>Соблюдение заданных пропорций между разными жанрами, стилями, периодами или группами авторов в соответствии с целевой генеральной совокупностью текстов.</td></tr></table></td><td>Соотнесите каждый этап построения лингвистического корпуса (1–4) с технологией или программным инструментом, который наиболее часто применяется на этом этапе (А–Г).
<table><tr><td></td><td>Этап работы с корпусом</td><td></td><td>Технология / инструмент</td></tr><tr><td>1</td><td>Сбор текстовых материалов в в больших объёмах</td><td>А</td><td>Веб-краулеры (парсеры) для автоматического скачивания текстов из интернета.</td></tr><tr><td>2</td><td>Приведение текстов к единому</td><td>Б</td><td>Программы лемматизации и</td></tr></table></td></tr>
</table>

					2	Пропорциональность	Б	Возможность автоматического поиска и извлечения примеров с помощью программных средств (конкордансеров).		формату и очистка от шума (HTML-тегов, мусорных символов)		морфологической разметки (например, MyStem, TreeTagger).
					3	Аннотирование	В	Способность корпуса отражать все основные характеристики языка или подъязыка, для которого он создавался, в тех же пропорциях, в которых они встречаются в реальности.	3	Лингвистическая разметка (токенизация, лемматизация, частеречная аннотация)	В	Скрипты на Python с использованием библиотек BeautifulSoup, re (регулярные выражения).
					4	Машинная читаемость	Г	Добавление к текстам метаинформации (морфологической, синтаксической, семантической) для возможности квантитативного анализа.	4	Оценка репрезентативности и пропорциональности корпуса	Г	Статистические методы (хи-квадрат, измерение разнообразия лексики, сравнение частот с эталонным корпусом).
					Ответ:				Ответ:			
									1	2	3	4
									Ответ:			

<table>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>
<table>
<tr><td>1</td><td>2</td><td>3</td><td>4</td></tr>
<tr><td></td><td></td><td></td><td></td></tr>
</table>
Ответ:
<table>
<tr><td>1</td><td>2</td><td>3</td><td>4</td></tr>
<tr><td>В</td><td>А</td><td>Б</td><td>Г</td></tr>
</table>
</td>
<td>
<table>
<tr><td>1</td><td>2</td><td>3</td><td>4</td></tr>
<tr><td>А</td><td>В</td><td>Б</td><td>Г</td></tr>
</table>
</td>
</tr>
<tr>
<td>6.</td>
<td>Тема 2. Корпусная лингвистика как самостоятельное лингвистическое направление. Характеристики корпуса (требования, репрезентативность и пропорциональность</td>
<td>ОПК-6. Способен применять современные технологии при осуществлении сбора, обработки и интерпретации данных эмпирического исследования; составлять и оформлять научную документацию;</td>
<td>ИОПК-6.1. Знает современные технологии сбора, обработки и интерпретации полученных экспериментальных данных.</td>
<td>2</td>
<td>Расположите перечисленные этапы создания сбалансированного и репрезентативного корпуса текстов в порядке от начала к концу. Запишите цифры в правильной последовательности.

Шаги (в случайном порядке):
1. Оценка репрезентативности и пропорциональности с помощью статистических методов (сравнение частот с эталонным корпусом).
2. Определение целевой генеральной совокупности текстов (жанры, стили, периоды, авторы).
3. Очистка и нормализация собранных текстов (удаление шума, приведение к единому формату).
4. Расчёт необходимых объёмов текстов для каждой страты с учётом пропорций.
5. Сбор текстов из доступных источников (веб-краулеры, оцифрованные архивы).
6. Лингвистическая разметка (токенизация, лемматизация, частеречная аннотация).</td>
<td>Расположите шаги в логическом порядке последовательность обработки и интерпретации корпусных данных с использованием современных технологий от сырых данных к получению интерпретируемого результата. Запишите цифры в правильной последовательности.

1. Шаги (в случайном порядке):
2. Построение конкордансов и частотных списков с помощью корпусного менеджера (AntConc, TXM).
3. Формулировка исследовательского вопроса (например, различие в употреблении грамматической конструкции).
4. Визуализация результатов (гистограммы, тепловые карты) и их лингвистическая интерпретация.
5. Статистическая проверка значимости различий (хи-квадрат, log-likelihood, точный критерий Фишера).
6. Выбор корпуса или создание</td>
</tr>
</table>

					Ответ: | | | | | | | **Ответ:** **| 2 | 4 | 5 | 3 | 6 | 1 |**	подкорпуса, отвечающего критериям репрезентативности и пропорциональности. 7. Экспорт данных из корпуса в формат для статистической обработки (CSV, Excel). Ответ: | | | | | | | **Ответ:** **| 2 | 5 | 1 | 6 | 4 | 3 |**
7.	Тема 2. Корпусная лингвистика как самостоятельное лингвистическое направление. Характеристики корпуса (требования, репрезентативность и пропорциональность	**ОПК-6**. Способен применять современные технологии при осуществлении сбора, обработки и интерпретации данных эмпирического исследования; составлять и оформлять научную документацию;	**ИОПК-6.3. Владеет** навыками сбора, обработки и интерпретации полученных экспериментальных данных..	*3*	Исследователь создаёт корпус текстов деловой документации начала XXI века. Он собрал 10 000 документов, из которых 9 500 — это приказы, а 500 — договоры и отчёты. Однако в реальной практике того периода договоры и отчёты составляли 40% всех документов, а приказы — 60%. Какое требование к корпусу нарушено? **Варианты ответов:** А) Репрезентативность Б) Пропорциональность (сбалансированность) В) Машинная читаемость Г) Аннотированность **Ответ: Б** **Обоснование:**	Лингвист хочет провести квантитативный анализ употребления частицы «бы» в русских художественных текстах XIX века. Какой современный технологический маршрут наиболее адекватен для сбора, обработки и интерпретации данных с гарантией репрезентативности? **Варианты ответов:** А) Найти в интернете 10 случайных романов, вручную выписать все примеры с «бы» и подсчитать их частоту. Б) Скачать из открытых источников готовый частотный словарь русского языка XIX века и взять из него данные по частице «бы».

					Пропорциональность (сбалансированность) означает, что доли разных жанров или типов текстов в корпусе должны соответствовать их долям в генеральной совокупности. В условии договоры и отчёты в реальности составляли 40%, а в корпусе — лишь 5% (500 из 10 000). Это прямое нарушение пропорциональности.	В) Выбрать сбалансированный по жанрам и периодам корпус (например, Национальный корпус русского языка), сформировать подкорпус художественных текстов 1801–1900 гг., использовать корпусного менеджера для получения частот и коллокаций, затем интерпретировать результаты с учётом статистической значимости. Г) Создать собственный корпус из текстов трёх любимых писателей (Пушкин, Гоголь, Толстой), обработать его в AntConc и представить результаты без проверки репрезентативности. **Ответ: В** **Обоснование:** Вариант В соответствует современным стандартам: использование уже готового репрезентативного и пропорционального корпуса (НКРЯ), автоматическая обработка через корпусного менеджера, возможность статистического анализа.
8.	Тема 2. Корпусная лингвистика как самостоятельн	**ОПК-6**. Способен применять современные технологии при	**ИОПК-6.1.** Знает современны е технологии сбора,	*4*	Выберите все характеристики, обязательные для репрезентативного корпуса. А) Пропорциональность жанров, периодов и типов текстов	Выберите современные технологии / методы, которые обеспечивают пропорциональность при создании корпуса. А) Стратифицированная выборка по

<table>
<tr>
<td></td>
<td>ое лингвистическое направление. Характеристики корпуса (требования, репрезентативность и пропорциональность</td>
<td>осуществлении сбора, обработки и интерпретации данных эмпирического исследования; составлять и оформлять научную документацию;</td>
<td>обработки и интерпретации полученных экспериментальных данных.</td>
<td></td>
<td>Б) Только письменные тексты
В) Сбалансированная выборка по стратам
Г) Объём не менее 1 млн слов
Д) Наличие метаданных (автор, дата, жанр)

Ответ: А, В, Д
Обоснование: Б — неверно (корпуса могут включать устную и мультимодальную речь). Г — неверно (репрезентативность не зависит от фиксированного объёма, а определяется целью исследования).</td>
<td>заранее заданным долям
Б) Случайный отбор текстов без учёта жанра
В) Веб-краулеры с фильтрацией по жанровым меткам
Г) Пост-хок статистическое взвешивание частот
Д) Исключительно ручной отбор одним экспертом
Ответ: А, В, Г
Обоснование: Б — нарушает пропорциональность. Д — субъективен и не обеспечивает воспроизводимости пропорций.</td>
</tr>
<tr>
<td>9.</td>
<td>Тема 3. Создания корпусов и их технология. Субкорпуса и принципы их создания Корпусный менеджер. Платформа доступа к корпусу.</td>
<td>ОПК-6. Способен применять современные технологии при осуществлении сбора, обработки и интерпретации данных эмпирического исследования; составлять и оформлять научную документацию;</td>
<td>ИОПК-6.1. Знает современные технологии сбора, обработки и интерпретации полученных экспериментальных данных.</td>
<td>1</td>
<td>Установите соответствие между термином (1–4) и его определением (А–Г).
<table>
<tr><td></td><td>Термин</td><td></td><td>Определение</td></tr>
<tr><td>1</td><td>Субкорпус</td><td>А</td><td>Веб-интерфейс для поиска, фильтрации и визуализации корпусных данных</td></tr>
<tr><td>2</td><td>Корпусный менеджер</td><td>Б</td><td>Часть корпуса, выделенная по одному или нескольким метапризнакам (жанр, автор,</td></tr>
</table>
</td>
<td>Установите соответствие между технологическим этапом создания субкорпуса (1–4) и используемым методом/инструментом (А–Г).
<table>
<tr><td></td><td>Технологический этап создания субкорпуса</td><td></td><td>Метод/инструмент</td></tr>
<tr><td>1</td><td>Фильтрация текстов по дате</td><td>А</td><td>Регулярные выражения по полю <date> в метаданных</td></tr>
<tr><td>2</td><td>Извлечение</td><td>Б</td><td>Функция KWIC (Key Word in</td></tr>
</table>
</td>
</tr>
</table>

<table>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>
<table>
<tr><td></td><td></td><td></td><td>время)</td></tr>
<tr><td>3</td><td>Платформа доступа</td><td>В</td><td>Программа для построения конкордансов, частотных списков и коллокаций (AntConc, Sketch Engine)</td></tr>
<tr><td>4</td><td>Стратификация</td><td>Г</td><td>Разбиение корпуса на непересекающиеся слои для обеспечения сбалансированности</td></tr>
</table>
Ответ:
<table>
<tr><td>1</td><td>2</td><td>3</td><td>4</td></tr>
<tr><td></td><td></td><td></td><td></td></tr>
</table>
Ответ:
<table>
<tr><td>1</td><td>2</td><td>3</td><td>4</td></tr>
<tr><td>Б</td><td>В</td><td>А</td><td>Г</td></tr>
</table>
</td>
<td>
<table>
<tr><td></td><td>контекстов для указанной лексемы</td><td></td><td>Context) в корпусном менеджере</td></tr>
<tr><td>3</td><td>Оценка репрезентативности субкорпуса</td><td>В</td><td>Сравнение частот с полным корпусом через log-likelihood</td></tr>
<tr><td>4</td><td>Добавление морфологической разметки</td><td>Г</td><td>Автоматический тэггер (TreeTagger, MyStem)</td></tr>
</table>
Ответ:
<table>
<tr><td>1</td><td>2</td><td>3</td><td>4</td></tr>
<tr><td></td><td></td><td></td><td></td></tr>
</table>
Ответ:
<table>
<tr><td>1</td><td>2</td><td>3</td><td>4</td></tr>
<tr><td>А</td><td>Б</td><td>В</td><td>Г</td></tr>
</table>
</td>
</tr>
<tr>
<td>10.</td>
<td>Тема 3. Создания корпусов и их технология. Субкорпуса и принципы их создания</td>
<td>ОПК-6. Способен применять современные технологии при осуществлении сбора,</td>
<td>ИОПК-6.3. Владеет навыками сбора, обработки и интерпретации</td>
<td>2</td>
<td>Установите правильную последовательность создания субкорпуса из большого корпуса с помощью корпусного менеджера

Шаги (в случайном порядке):
1. Экспорт полученных данных</td>
<td>Установите последовательность работы с платформой доступа к корпусу для получения и интерпретации экспериментальных данных

Шаги (в случайном порядке):</td>
</tr>
</table>

<table>
<tr>
<td></td>
<td>Корпусный менеджер. Платформа доступа к корпусу.</td>
<td>обработки и интерпретации данных эмпирического исследования; составлять и оформлять научную документацию;</td>
<td>полученных экспериментальных данных.</td>
<td></td>
<td>(конкордансов, частотных списков) для дальнейшего статистического анализа.
2. Загрузка основного корпуса в интерфейс корпусного менеджера (например, Sketch Engine, AntConc).
3. Интерпретация полученных количественных результатов с опорой на лингвистическую задачу.
4. Формулировка критериев для субкорпуса (жанр, хронологический период, автор).
5. Применение фильтров по метаданным для формирования субкорпуса.
6. Запуск квантитативного запроса (частотность, коллокации) внутри созданного субкорпуса.
Ответ:
<table><tr><td></td><td></td><td></td><td></td><td></td><td></td></tr></table>
Ответ:
<table><tr><td>4</td><td>2</td><td>5</td><td>6</td><td>1</td><td>3</td></tr></table></td>
<td>1. Визуализация частотного распределения в виде графиков (гистограмм, линейчатых диаграмм) на платформе.
2. Выбор целевого корпуса или субкорпуса в веб-интерфейсе платформы.
3. Формулировка поискового запроса (лексема, грамматическая конструкция).
4. Выгрузка результатов в формате CSV или Excel для статистической проверки (хи-квадрат, log-likelihood).
5. Получение конкорданса и частотных списков с указанием нормализованных частот (на млн слов).
6. Лингвистическая интерпретация значимых различий с учётом репрезентативности субкорпуса.
Ответ:
<table><tr><td></td><td></td><td></td><td></td><td></td><td></td></tr></table>
Ответ:
<table><tr><td>3</td><td>2</td><td>5</td><td>1</td><td>4</td><td>6</td></tr></table></td>
</tr>
<tr>
<td>11.</td>
<td>Тема 3. Создания корпусов и их технология. Субкорпуса и принципы их создания Корпусный</td>
<td>ОПК-6. Способен применять современные технологии при осуществлении сбора, обработки и</td>
<td>ИОПК-6.1. Знает современные технологии сбора, обработки и интерпретации</td>
<td>3</td>
<td>Исследователь хочет сравнить употребление глаголов в научных текстах 1990-х и 2010-х годов. Что ему нужно сделать в первую очередь?

А) Загрузить весь корпус в Excel
Б) Создать два субкорпуса по метапризнаку «время»</td>
<td>Какой инструмент позволяет получить частотный список лемм в субкорпусе с нормализацией на 1 млн слов?

А) Текстовый редактор
Б) Корпусный менеджер (AntConc, Sketch Engine)
В) Система компьютерной вёрстки</td>
</tr>
</table>

	менеджер. Платформа доступа к корпусу.	интерпретации данных эмпирического исследования; составлять и оформлять научную документацию;	полученных экспериментальных данных.		В) Построить конкорданс для каждого глагола вручную Г) Установить платформу доступа на локальный сервер **Ответ: Б** **Обоснование**: Без создания субкорпусов (выборки по хронологическому признаку) невозможно корректно сравнить периоды. Остальные действия либо преждевременны, либо неверны.	Г) Грамматический словарь **Ответ: Б** **Обоснование:** Корпусный менеджер специально предназначен для квантитативной обработки (частоты, коллокации, конкордансы) внутри корпуса или субкорпуса с автоматической нормализацией. Остальные варианты не дают таких возможностей.
12.	Тема 3. Создания корпусов и их технология. Субкорпуса и принципы их создания Корпусный менеджер. Платформа доступа к корпусу.	**ОПК-6**. Способен применять современные технологии при осуществлении сбора, обработки и интерпретации данных эмпирического исследования; составлять и оформлять научную документацию;	**ИОПК-6.1.** Знает современные технологии сбора, обработки и интерпретации полученных экспериментальных данных.	*4*	Выберите все корректные действия при создании субкорпуса с помощью корпусного менеджера. А) Фильтрация по метаданным (жанр, дата, автор) Б) Ручное удаление текстов из основного корпуса В) Использование регулярных выражений для отбора по полю <year> Г) Случайная выборка без учёта пропорций Д) Сохранение субкорпуса как отдельной коллекции **Ответ: А, В, Д** **Обоснование:** А и В — штатные способы фильтрации в корпусных менеджерах; Д — функция сохранения для повторного использования. Б — неавтоматизировано, противоречит технологии. Г — нарушает	Какие возможности платформы доступа к корпусу (Sketch Engine, НКРЯ) необходимы для интерпретации экспериментальных данных? А) Построение конкордансов (KWIC) Б) Вычисление коллокационной силы (MI, logDice) В) Редактирование исходных текстов корпуса Г) Экспорт частотных списков в CSV Д) Автоматическая запись устной речи **Ответ: А, Б, Г** **Обоснование:** А, Б, Г — стандартные инструменты квантитативной обработки и выгрузки данных для интерпретации. В — не предусмотрено (корпус неизменяем). Д — относится к сбору, а не к платформе доступа.

					репрезентативность	
13.	Тема 1. Современная квантитативная лингвистика как наука. Новые информационные технологии в лингвистических исследованиях и образовательном процессе.	ОПК-6. Способен применять современные технологии при осуществлении сбора, обработки и интерпретации данных эмпирического исследования; составлять и оформлять научную документацию;	ИОПК-6.2. Умеет применять современные технологии сбора, обработки и интерпретации полученных экспериментальных данных.	3	Лингвист получил массив ответов испытуемых (время реакции, точность). Какой инструмент оптимален для проверки статистической значимости различий между двумя группами? А) Текстовый редактор Б) Пакет SPSS или R (t-критерий) В) Программа для записи экрана Г) Онлайн-анкета без экспорта **Ответ: Б** **Обоснование**: SPSS и R позволяют применить параметрические критерии (t-тест) для интерпретации экспериментальных данных. Остальные варианты не предназначены для статистической обработки.	В эксперименте по восприятию редких слов нужно зафиксировать микродвижения глаз. Какую технологию следует применить? А) Корпусный менеджер Б) Айтрекинг (eye-tracker) В) Опрос в Google Forms Г) Частотный словарь **Ответ: Б** **Обоснование**: Айтрекинг собирает данные о фиксациях, саккадах и зрачках в реальном времени. Корпусные и опросные методы не дают таких физиологических показателей.
14.	Тема 1. Современная квантитативная лингвистика как наука. Новые информационные технологии в лингвистических	ОПК-6. Способен применять современные технологии при осуществлении сбора, обработки и интерпретации данных эмпирического	ИОПК-6.2. Умеет применять современные технологии сбора, обработки и интерпретации полученных эксперимент	4	Выберите все технологии, которые используются для сбора экспериментальных данных в квантитативной лингвистике. А) Айтрекинг Б) Корпусный менеджер (AntConc) В) Платформа онлайн-экспериментов (PsychoPy, Pavlovia) Г) Пакет статистического анализа R Д) Краудсорсинговые платформы (Toloka, Prolific)	Какие действия входят в корректную обработку и интерпретацию количественных лингвистических данных после проведения эксперимента? А) Проверка нормальности распределения Б) Визуализация данных (ящик с усами, гистограмма) В) Автоматический выбор t-критерия без проверки условий

	исследованиях и образовательном процессе.	исследования; составлять и оформлять научную документацию;	альных данных.		**Ответ: А, В, Д** **Обоснование:** Айтрекинг (А) фиксирует физиологические реакции, PsychoPy (В) собирает поведенческие данные (время, точность), краудсорсинг (Д) — массовые оценки и разметку. AntConc (Б) и R (Г) — инструменты обработки, а не сбора.	Г) Расчёт описательной статистики (среднее, медиана, σ) Д) Интерпретация p-value в отрыве от лингвистической теории **Ответ: А, Б, Г** **Обоснование**: Проверка нормальности (А), визуализация (Б) и описательная статистика (Г) — обязательные этапы современной обработки. В — ошибочен (нужна проверка условий). Д — неверен (статистика должна интерпретироваться лингвистически).
15.	Тема 1. Современная квантитативная лингвистика как наука. Новые информационные технологии в лингвистических исследованиях и образовательном процессе.	ОПК-6. Способен применять современные технологии при осуществлении сбора, обработки и интерпретации данных эмпирического исследования; составлять и оформлять научную документацию;	ИОПК-6.2. Умеет применять современные технологии сбора, обработки и интерпретации полученных экспериментальных данных.	*5*	Исследователь проводит эксперимент: испытуемые читают предложения с редкими словами, а компьютер фиксирует точки остановки взгляда и движения зрачка. Какая технология сбора данных здесь используется? Ответ запишите словом в именительном падеже Ответ_____ **Ответ: Айтрекинг**	Лингвист получил числовые данные о частоте употребления конструкции в разных жанрах и хочет выполнить кластерный анализ и построить графики. Какой бесплатный язык программирования и среда статистических вычислений ему следует применить? Ответ: _____ **Ответ: R**
16.	Тема 2. Корпусная	ОПК-6. Способен	ИОПК-6.2. Умеет	*3*	Корпус содержит 90% газетных текстов и 10% научных, хотя в реальной речи их	Какой инструмент позволяет автоматически получить коллокации

	лингвистика как самостоятельное лингвистическое направление. Характеристики корпуса (требования, репрезентативность и пропорциональность	применять современные технологии при осуществлении сбора, обработки и интерпретации данных эмпирического исследования; составлять и оформлять научную документацию;	применять современные технологии сбора, обработки и интерпретации полученных экспериментальных данных.		соотношение 50/50. Какое требование нарушено? А) Аннотированность Б) Пропорциональность В) Машинная читаемость Г) Динамичность **Ответ: Б** **Обоснование:** Нарушена пропорциональность (сбалансированность) – доли жанров не соответствуют генеральной совокупности.	слова «свобода» в субкорпусе публицистики 2000-х годов? А) Текстовый редактор Б) Корпусный менеджер (AntConc, Sketch Engine) В) Система распознавания речи Г) Грамматический словарь **Ответ: Б** **Обоснование:** Корпусный менеджер предназначен для построения конкордансов, частотных списков и коллокаций с учётом метаданных субкорпуса.
17.	Тема 2. Корпусная лингвистика как самостоятельное лингвистическое направление. Характеристики корпуса (требования, репрезентативность и пропорциональность	ОПК-6. Способен применять современные технологии при осуществлении сбора, обработки и интерпретации данных эмпирического исследования; составлять и оформлять научную документацию;	**ИОПК-6.3. Владеет** навыками сбора, обработки и интерпретации полученных экспериментальных данных.	*4*	Выберите условия, необходимые для репрезентативности корпуса. А) Соблюдение пропорций жанров, периодов, типов речи Б) Только письменные тексты В) Объём более 10 млн слов Г) Наличие метаданных для стратификации Д) Случайная выборка внутри каждой страты **Ответ: А, Г, Д** **Обоснование:** А (пропорции), Г (метаданные), Д (случайная выборка) обеспечивают репрезентативность. Б и В – необязательны.	Какие современные технологии позволяют оценить пропорциональность корпуса? А) Сравнение частот с эталонным корпусом (хи-квадрат) Б) Построение конкордансов для одного слова В) Статистический анализ распределения метапризнаков Г) Визуализация страт через диаграммы Д) Ручной подсчёт документов **Ответы: А, В, Г** **Обоснование:** А (статсравнение), В (анализ распределения), Г (визуализация) – автоматизированные методы. Б – не оценивает пропорции.

						Д – несовременная технология.
18.	Тема 2. Корпусная лингвистика как самостоятельное лингвистическое направление. Характеристики корпуса (требования, репрезентативность и пропорциональность	ОПК-6. Способен применять современные технологии при осуществлении сбора, обработки и интерпретации данных эмпирического исследования; составлять и оформлять научную документацию;	**ИОПК-6.3. Владеет** навыками сбора, обработки и интерпретации полученных экспериментальных данных..	*5*	Вы формируете корпус русских художественных текстов XX века: включаете 80% прозы и 20% поэзии, хотя в реальной литературе этого периода соотношение 50% на 50%. Какое требование к корпусу нарушено? Ответ_____ **Ответ: пропорциональность**	Вам нужно автоматически получить список коллокатов слова «исследование» с указанием силы связи (MI score) в подкорпусе научных статей. Какой тип программы для этого необходим? Ответ_____ **Ответ: конкордансер / корпусный менеджер**
19.	Тема 3.	ОПК-6.	ИОПК-6.2.	*3*	Вы создаёте субкорпус для анализа	Вы завершили обработку данных в

	Создания корпусов и их технология. Субкорпуса и принципы их создания Корпусный менеджер. Платформа доступа к корпусу.	Способен применять современные технологии при осуществлении сбора, обработки и интерпретации данных эмпирического исследования; составлять и оформлять научную документацию;	Умеет применять современные технологии сбора, обработки и интерпретации полученных экспериментальных данных.		эволюции термина «искусственный интеллект» в научных статьях 1970–2020 гг. Какой принцип создания субкорпуса вы примените в первую очередь? А) Фильтрацию по полю «год публикации» в метаданных Б) Случайную выборку 1000 статей без учёта даты В) Построение конкорданса по всем словам Г) Ручной отбор статей по названиям **Ответ: А** **Обоснование:** Для анализа изменений во времени необходимо сначала выделить субкорпус по хронологическому признаку через фильтрацию метаданных, что позволяет далее сравнивать разные периоды.	корпусном менеджере и должны представить результаты в виде научной статьи. Что необходимо включить в раздел «Методология»? А) Только список использованной литературы Б) Название корпуса, версию, критерии формирования субкорпуса, название корпусного менеджера и статистические методы В) Все полученные конкордансы целиком Г) Графики без указания источника данных **Ответ: Б** **Обоснование**: Научная документация требует указания корпуса, субкорпуса, инструмента (корпусного менеджера) и методов анализа для воспроизводимости исследования.
20.	Тема 3. Создания корпусов и их технология. Субкорпуса и принципы их создания Корпусный менеджер. Платформа	ОПК-6. Способен применять современные технологии при осуществлении сбора, обработки и интерпретации данных	ИОПК-6.2. Умеет применять современные технологии сбора, обработки и интерпретации полученных	*4*	Вы готовите корпусное исследование диахронических изменений в научном стиле. Какие действия необходимо выполнить при создании субкорпуса для эмпирического исследования? А) Определить метапризнаки (жанр, время, автор) Б) Использовать корпусный менеджер для фильтрации В) Включить все тексты основного	Вы завершили анализ и готовите научную статью по результатам корпусного исследования. Какие элементы должны быть указаны в разделе «Материалы и методы» научной статьи по корпусному исследованию? А) Название и версия корпуса Б) Логин и пароль от платформы доступа

	доступа к корпусу.	эмпирического исследования; составлять и оформлять научную документацию;	экспериментальных данных.		корпуса без изменений Г) Проверить репрезентативность субкорпуса Д) Оформить описание субкорпуса в научной документации **Ответ: А, Б, Г, Д** **Обоснование:** В (включение всех текстов) противоречит созданию субкорпуса. Остальные – обязательные шаги.	В) Критерии формирования субкорпуса Г) Название корпусного менеджера и версия Д) Статистические критерии для интерпретации **Ответы: А, В, Г, Д** **Объяснение:** Б (логин и пароль) не относится к научной документации. Остальные – необходимы для воспроизводимости.
21.	Тема 3. Создания корпусов и их технология. Субкорпуса и принципы их создания Корпусный менеджер. Платформа доступа к корпусу.	ОПК-6. Способен применять современные технологии при осуществлении сбора, обработки и интерпретации данных эмпирического исследования; составлять и оформлять научную документацию;	**ИОПК-6.3. Владеет** навыками сбора, обработки и интерпретации полученных экспериментальных данных.	*5*	Вы создаёте субкорпус для исследования эволюции юридической терминологии. Вы отобрали тексты законов только за 2020 год, хотя реальные изменения происходили с 1990 по 2020 год. Какой принцип создания субкорпуса нарушен? Ответ: _____ **Ответ: репрезентативность**	Вы провели анализ в корпусном менеджере и теперь оформляете раздел «Материалы и методы» для научной статьи. Как называется документ, в котором фиксируются все параметры корпуса, субкорпуса, инструментов и статистических критериев для обеспечения воспроизводимости исследования? Ответ: _____ **Ответ: протокол /протокол исследования**

Критерии оценивания тестового задания:

Оценка	Критерии оценивания
отлично	от 90 до 100 % правильно выполненных заданий

хорошо	от 70 до 89 % правильно выполненных заданий
удовлетворительно	от 50 до 69 % правильно выполненных заданий
неудовлетворительно	менее 50 % правильно выполненных заданий