

Документ подписан простой электронной подписью

Информация о владельце:

ФИО: Макаренко Елена Николаевна

Должность: Ректор

Дата подписания: 07.08.2022 Министерство науки и высшего образования Российской Федерации

Уникальный идентификатор: c098bc0c1041cb2a4cf976cf171d6715d99a6ae00adc8e27b555che1e2dbd7c78  
Федеральное государственное бюджетное образовательное учреждение высшего образования «Ростовский государственный экономический университет (РИНХ)»

УТВЕРЖДАЮ

Директор Института магистратуры

 Иванова Е.А.

« 29 » 08 2022 г.

**Рабочая программа дисциплины  
Квантитативные методы в лингвострановедческих исследованиях**

Направление 41.04.01 Зарубежное регионоведение  
магистерская программа 41.04.01.01 "Этнические, социальные и  
лингвокультурологические особенности стран Восточной Азии"

Для набора 2022 года

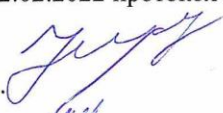
Квалификация  
магистр


КАФЕДРА **Иностранные языки для гуманитарных специальностей****Распределение часов дисциплины по семестрам**


Семестр (<Курс>.<Семестр на курсе>)	2 (1.2)		Итого	
	15 2/6			
Вид занятий	уп	рп	уп	рп
Лекции	10	10	10	10
Практические	10	10	10	10
Итого ауд.	20	20	20	20
Контактная работа	20	20	20	20
Сам. работа	52	52	52	52
Часы на контроль	36	36	36	36
Итого	108	108	108	108

**ОСНОВАНИЕ**

Учебный план утвержден учёным советом вуза от 22.02.2022 протокол № 7.

Программу составил(и): к.ф.н., доц., Казанская Е.В. 

Зав. кафедрой: к.филол.н., доцент Гермашева Т.М. 

Методическим советом направления: к.ф.н., проф., Евсюкова Т.В. 

### 1. ЦЕЛИ ОСВОЕНИЯ ДИСЦИПЛИНЫ

- |     |  |
|-----|--|
| 1.1 | Цели освоения дисциплины: Основной целью дисциплины является формирование лингвоинформационной компетентности студентов. |
|-----|--|

### 2. ТРЕБОВАНИЯ К РЕЗУЛЬТАТАМ ОСВОЕНИЯ ДИСЦИПЛИНЫ

**ОПК-2: Способен осуществлять поиск и применять перспективные информационно-коммуникационные технологии и программные средства для комплексной постановки и решения задач профессиональной деятельности**

#### В результате освоения дисциплины обучающийся должен:

**Знать:**

основные понятия количественной лингвистики, стилистики, атрибуции; принципы построения частотных словарей; принципы работы с экспертными программами в области психолингвистики и корпусными технологиями (соотнесено с индикатором ОПК-2.1);

**Уметь:**

применять новые информационные технологии в процессе изучения иностранного языка и профессиональной переводческой деятельности (соотнесено с индикатором ОПК-2.2);

**Владеть:**

основами применения количественных методов в лингвистической практике в условиях информационной среды (соотнесено с индикатором ОПК-2.3).

### 3. СТРУКТУРА И СОДЕРЖАНИЕ ДИСЦИПЛИНЫ

Код занятия	Наименование разделов и тем /вид занятия/	Семестр / Курс	Часов	Компетенции	Литература
	<b>Раздел 1. «Количественная лингвистика как направление общего языкознания. Понятие и сущность лингвостатистического анализа»</b>				
1.1	Тема «История развития количественной лингвистики и взаимосвязь дисциплины с другими науками» /Лек/	2	2	ОПК-2	Л1.1 Л1.2Л2.2 Л2.1
1.2	Тема «Ключевые понятия количественной лингвистики» /Лек/	2	2	ОПК-2	Л1.1 Л1.2Л2.2 Л2.1
1.3	Тема «Стилистика и атрибуция» /Пр/	2	2	ОПК-2	Л1.1 Л1.2Л2.2 Л2.1
1.4	Тема «Аспекты процесса атрибуции текста, идентификация авторства текстов» /Пр/	2	2	ОПК-2	Л1.1 Л1.2Л2.2 Л2.1
1.5	Тема «Области применения структурно- вероятностной модели языка» /Ср/	2	10	ОПК-2	Л1.1 Л1.2Л2.2 Л2.1
1.6	Тема «Национальные корпуса текстов» /Ср/	2	10	ОПК-2	Л1.1 Л1.2Л2.2 Л2.1
1.7	Тема «Глубинный анализ текстов (Text Mining)» /Ср/	2	10	ОПК-2	Л1.1 Л1.2Л2.2 Л2.1
1.8	Тема «Программное обеспечение для идентификации авторства текстов» /Ср/	2	10	ОПК-2	Л1.1 Л1.2Л2.2 Л2.1
	<b>Раздел 2. «Количественные исследования лексики»</b>				
2.1	Тема «Количественные и статистические методы в исследовании лексики» /Лек/	2	2	ОПК-2	Л1.1 Л1.2Л2.2 Л2.1
2.2	Тема «Частотные словари, модель «ранг- частота» /Лек/	2	2	ОПК-2	Л1.1 Л1.2Л2.2 Л2.1
2.3	Тема «Закон Ципфа-Мандельброта. Закономерности функционирования лексики» /Лек/	2	2	ОПК-2	Л1.1 Л1.2Л2.2 Л2.1
2.4	Тема «Искусственный интеллект и обработка текстов на естественных языках» /Пр/	2	2	ОПК-2	Л1.1 Л1.2Л2.2 Л2.1
2.5	Тема «Автоматическое реферирование и аннотирование текстов» /Пр/	2	2	ОПК-2	Л1.1 Л1.2Л2.2 Л2.1
2.6	Тема «Системы автоматической обработки текстов» /Пр/	2	2	ОПК-2	Л1.1 Л1.2Л2.2 Л2.1

2.7	Темы рефератов: 1. Структурно-вероятностная модель языка. 2. Статистические инструменты в лингвистике. 3. Проблема репрезентативности лингвистической выборки. 4. Основные исторические вехи возникновения и уточнения закона Ципфа. 5. Современный текстовый корпус. 6. Квазисинонимы. Метод квазисинонимов. 7. Направления обработки текстов на естественных языках. /Ср/	2	12	ОПК-2	Л1.1 Л1.2Л2.2 Л2.1
2.8	/Экзамен/	2	36	ОПК-2	Л1.1 Л1.2Л2.2 Л2.1

#### 4. ФОНД ОЦЕНОЧНЫХ СРЕДСТВ

Структура и содержание фонда оценочных средств для проведения текущей и промежуточной аттестации представлены в Приложении 1 к рабочей программе дисциплины.

#### 5. УЧЕБНО-МЕТОДИЧЕСКОЕ И ИНФОРМАЦИОННОЕ ОБЕСПЕЧЕНИЕ ДИСЦИПЛИНЫ

##### 5.1. Основная литература

	Авторы, составители	Заглавие	Издательство, год	Колич-во
Л1.1	Моисеева И. Ю.	Квантитативная лингвистика и новые информационные технологии: учебное пособие	Оренбург: Оренбургский государственный университет, 2017	<a href="https://biblioclub.ru/index.php?page=book&amp;id=481797">https://biblioclub.ru/index.php?page=book&amp;id=481797</a> неограниченный доступ для зарегистрированных пользователей
Л1.2	Моисеева, И. Ю.	Квантитативная лингвистика и новые информационные технологии: учебное пособие	Оренбург: Оренбургский государственный университет, ЭБС АСВ, 2017	<a href="http://www.iprbookshop.ru/71281.html">http://www.iprbookshop.ru/71281.html</a> неограниченный доступ для зарегистрированных пользователей

##### 5.2. Дополнительная литература

	Авторы, составители	Заглавие	Издательство, год	Колич-во
Л2.1	Котельникова Е. В.	Квантитативная лингвистика: учеб.-метод. пособие	Ростов н/Д: Изд-во РГЭУ (РИНХ), 2018	68
Л2.2		Вестник Московского института лингвистики: журнал	Москва: Московский институт лингвистики, 2015	<a href="https://biblioclub.ru/index.php?page=book&amp;id=434674">https://biblioclub.ru/index.php?page=book&amp;id=434674</a> неограниченный доступ для зарегистрированных пользователей

##### 5.3 Профессиональные базы данных и информационные справочные системы

Консультант +

Гарант

Cambridge Dictionary <https://www.dictionary.cambridge.org/ru/>

##### 5.4. Перечень программного обеспечения

LibreOffice

##### 5.5. Учебно-методические материалы для студентов с ограниченными возможностями здоровья

При необходимости по заявлению обучающегося с ограниченными возможностями здоровья учебно-методические материалы предоставляются в формах, адаптированных к ограничениям здоровья и восприятия информации. Для лиц с нарушениями зрения: в форме аудиофайла; в печатной форме увеличенным шрифтом. Для лиц с нарушениями слуха: в форме электронного документа; в печатной форме. Для лиц с нарушениями опорно-двигательного аппарата: в форме электронного документа; в печатной форме.

#### 6. МАТЕРИАЛЬНО-ТЕХНИЧЕСКОЕ ОБЕСПЕЧЕНИЕ ДИСЦИПЛИНЫ (МОДУЛЯ)

Помещения для проведения всех видов работ, предусмотренных учебным планом, укомплектованы необходимой специализированной учебной мебелью и техническими средствами обучения:

- столы, стулья;

- персональный компьютер (переносной);

- проектор, экран/интерактивная доска.

#### **7. МЕТОДИЧЕСКИЕ УКАЗАНИЯ ДЛЯ ОБУЧАЮЩИХСЯ ПО ОСВОЕНИЮ ДИСЦИПЛИНЫ (МОДУЛЯ)**

Методические указания по освоению дисциплины представлены в Приложении 2 к рабочей программе дисциплины.

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

1.1 Показатели и критерии оценивания компетенций:

ЗУН, составляющие компетенцию	Показатели оценивания	Критерии оценивания	Средства оценивания
ОПК-2 - Способен осуществлять поиск и применять перспективные информационно-коммуникационные технологии и программные средства для комплексной постановки и решения задач профессиональной деятельности			
<b>Знать:</b> основные понятия квантитативной лингвистики, стилистики, атрибуции; принципы построения частотных словарей; принципы работы с экспертными программами в области психолингвистики и корпусными технологиями	умение работать со специализированным программным обеспечением, проводить статистический анализ текстовой информации, использовать методики работы с лингвистическим материалом в профессиональных целях.	Полнота и содержательность ответа, правильность выполнения практических задач, употребления лексических единиц и грамматических конструкций.	Тесты 1-3, Устный опрос, практические задания (1-5), доклад с презентацией, Вопросы к экзамену (1-10)
<b>Уметь:</b> применять новые информационные технологии в процессе изучения иностранного языка и профессиональной переводческой деятельности	Решение задач при помощи программных продуктов статистического анализа, математического моделирования дискурса	Выполнение поставленной задачи в полном объеме;	Тесты 1-3, Устный опрос, практические задания (1-5), контрольные работы, Вопросы к экзамену (1-10)
<b>Владеть:</b> основами применения количественных методов в лингвистической практике в условиях информационной среды	Уверенное ориентирование в теме, способность решать поставленную задачу с помощью информационных и квантитативных методов.	Выполнение коммуникативной задачи в полном объеме; Корректное употребление лексических единиц; Понимание грамматических и стилистических конструкций. Самостоятельность выводов и суждений с соответствующим правильным использованием информационных и	Тесты 1-3, Устный опрос, практические задания, Вопросы к экзамену (1-10)

		квантитативных методов предложенной теме	
--	--	--	--

### 1.2 Шкалы оценивания:

Текущий контроль успеваемости и промежуточная аттестация осуществляется в рамках накопительной балльно-рейтинговой системы в 100-балльной шкале:

84-100 баллов (оценка «отлично»)

67-83 балла (оценка «хорошо»)

50-66 баллов (оценка «удовлетворительно»)

0-49 баллов (оценка «неудовлетворительно»)

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

Вопросы к экзамену:

1. Что понимается под термином «квантитативная лингвистика»?
2. Какие математические методы используются в квантитативной лингвистике?
3. Что такое структурно-вероятностная модель языка?
4. В чем заключается принципиальная цель лингвостатистического анализа?
5. Что обозначают аббревиатуры ГЛС и ВЛС?
6. По какой формуле рассчитывается среднее квадратичное отклонение?
7. Изложите суть закона Дж. Ципфа. Какой вклад в развитие закона Дж. Ципфа внес Б. Мандельброт?
8. Что такое национальный корпус? Приведите примеры существующих национальных корпусов.
9. Что изучает стилеметрия как наука?
10. В чем сущность методов количественного анализа квазисинонимичных лексем?

### Критерии оценивания:

- оценка «отлично» (84-100 баллов) выставляется, если даны полные ответы на все вопросы, использованный словарный запас, грамматические структуры и фонетическое оформление соответствуют поставленной задаче; если перевод сделан полностью с соблюдением стиля высказывания, без нарушения языковых норм; допускается одна суммарная ошибка, кроме искажения смысла.

- оценка «хорошо» (67-83 балла) выставляется, если даны неполные ответы на все вопросы, использованный словарный запас, грамматические структуры, фонетическое оформление высказывания соответствует поставленной задаче, небольшие нарушения использования средств логической связи; если перевод сделан полностью с соблюдением стиля высказывания, без нарушения языковых норм. Допускаются две полные суммарные ошибки, в т.ч. не более одного искажения смысла.

- оценка «удовлетворительно» (50-66 баллов) выставляется, если даны неполные ответы на два вопроса, использован ограниченный словарный запас, грамматические структуры и фонетическое оформление высказывания имеют небольшие нарушения; , если перевод сделан адекватно; допускаются 3 полные суммарные ошибки.

- оценка «неудовлетворительно» (0-49 баллов) выставляется, если ответы не связаны с вопросами, наличие грубых ошибок в ответе, непонимание сущности излагаемого вопроса, неумение применять знания на практике, неуверенность и неточность ответов на

дополнительные и наводящие вопросы, если перевод реферирование текста не соответствует требованиям, установленных для оценки «удовлетворительно».

## Тесты

### Тест 1

1. \_\_\_\_\_ - это раздел общей лингвистики и, в частности, математической лингвистики, исследующей язык при помощи статистических методов и имеющей своей конечной целью формулирование законов функционирования языка и построение общей теории языка.

- A) Когнитивная лингвистика
- B) Квантитативная лингвистика
- C) Компьютерная лингвистика
- B) Нейролингвистика

2. \_\_\_\_\_ - это массив текстов, собранных в единую систему по определенным признакам (языку, жанру, времени создания текста, автору и т.д.) и снабженных поисковой системой.

- A) Энциклопедический словарь
- B) Лингвистический корпус
- C) Тезаурус
- D) Электронная библиотека

3. Методы \_\_\_\_\_ используются не только в цифровой гуманитарике, но и в компьютерных системах, помогающих переводчикам.

- A) Психолингвистики
- B) Квантитативной лингвистики
- C) Корпусной лингвистики
- D) Контрастивной лингвистики

\_\_\_\_\_ - это набор слов языка с указанием информации о частоте их встречаемости.

- A) Толковый словарь
- B) Частотный словарь
- C) Этимологический словарь
- D) Фразеологический словарь

5. Как называется программа для формирования конкорданса?

- A) Текстовый редактор
- B) Издательская система
- C) Текстовый процессор
- D) Конкордансер

6. \_\_\_\_\_ - метод качественно-количественного анализа содержания документов с целью выявления и оценки специфических характеристик текстов.

- A) Семантический анализ
- B) Контент-анализ
- C) Лингвостилистический анализ
- D) Контекстный анализ

7. \_\_\_\_\_ - информационная модель предметной области, имеющая вид ориентированного графа, вершины которого соответствуют объектам предметной области, а дуги (рёбра) задают отношения между ними. Объектами могут быть: понятия, события, свойства, процессы.

- A) Фрейм
- B) Контентная сеть
- C) Социальная сеть
- D) Семантическая сеть

8. \_\_\_\_\_ - это список всех употреблений заданного языкового выражения (например, слова) в контексте, возможно, со ссылками на источник.

- A) Словарная статья



- В) Конкорданс
- С) Толковый словарь
- Д) Тезаурус

9. \_\_\_\_\_ - процесс определения подлинности или подложности рукописного текста и установление его автора.

- А) Стилистический анализ
- В) Атрибуция текста
- С) Контент-анализ
- Д) Герменевтический анализ

10. \_\_\_\_\_ — процесс перевода текстов с одного естественного языка на другой при помощи специализированной компьютерной программы, направление научных исследований, связанных с построением подобных систем.

- А) Последовательный перевод
- В) Машинный перевод
- С) Коммунальный перевод
- Д) Авторизированный перевод

11. \_\_\_\_\_ - выполняемые на компьютере действия по преобразованию речи одного естественного языка в эквивалентную по содержанию речь другого, а также результаты такого действия.

- А) Машинный перевод
- В) Автоматический анализ текста

12. \_\_\_\_\_ собрание текстов в электронной форме, представляющих данный язык на определенном этапе его существования.

- А) Национальный корпус
- В) Массив лингвистических данных

13. \_\_\_\_\_ - приписывание текстам и их компонентам специальных меток.

- А) Разметка
- В) Семантическая сеть

14. \_\_\_\_\_ — порядковый номер слова в частотном словаре (списке).

- А) Ранг
- В) Частота

15. \_\_\_\_\_ — связный текст, который кратко выражает не только центральную тему или предмет какого-либо документа, но и цель, применяемые методы, основные результаты описанного исследования или разработки.

- А) Реферат
- В) Аннотация

16. \_\_\_\_\_ - краткое изложение содержания документа, дающее общее представление об его теме.

- А) Реферат
- В) Аннотация

17. \_\_\_\_\_ - соотнесение произведению соответствующих атрибутов, к которым причисляются имя создателя, жанр, время и место создания.

- А) Атрибуция (авторизация)
- В) Идиостиль

18. \_\_\_\_\_ - перевод с использованием как традиционной, так и машинной переработки текста, где большую часть стереотипной деятельности по отысканию эквивалентов и соответствий на основании формализованных методик фреймов-сценариев осуществляет искусственный интеллект

- А) Смешанный перевод
- В) Машинный перевод

19. \_\_\_\_\_ - прикладная филологическая дисциплина, занимающаяся измерением стиливых характеристик с целью систематизации и упорядочения (типологии, атрибуции, датировки, диагностики, реконструкции и т.д.) текстов и их частей.

А) Стилеметрия

В) Атрибуция

20. \_\_\_\_\_ - вид словаря, в котором приводятся числовые характеристики употребительности слов (словоформ, словосочетаний) какого-либо языка, в том числе языка писателя, какого-либо произведения.

А) Частотный словарь

В) Translation Memory

Ответы: 1. А; 2. D; 3. В; 4. В; 5. D; 6. В; 7. D; 8. В; 9. В; 10. В; 11. А; 12. А; 13. А; 14. А; 15. А; 16. В; 17. А; 18. А; 19. А; 20. А.

## Тест 2

1. Заполните пропуски.

\_\_\_\_\_ — несплошное наблюдение, при котором исследуется не вся совокупность языковых единиц, а лишь определенная часть.

А) Выборочное наблюдение

В) Случайное наблюдение

2. \_\_\_\_\_ — это функция, приближающая плотность вероятности некоторого распределения, построенная на основе выборки из него.

А) Граф

В) Гистограмма

3. \_\_\_\_\_ — вероятность, с которой доверительный интервал «захватит» истинное значение параметра генеральной совокупности.

А) Кумулята

В) Доверительная вероятность

4. \_\_\_\_\_ — интервал вокруг выборочного значения определенного параметра, который с заданной (достаточно высокой) вероятностью накрывает истинное значение параметра. Величина такого интервала зависит как от доверительной вероятности (т.е. надежности оценивания), так и от объема выборки.

А) Доверительный интервал

В) Выборочный интервал

5. Закон \_\_\_\_\_ — закономерность функционирования лексики, сутью которой является то, что чем шире сфера употребления слова, тем беднее его содержание.

А) Крушевского-Куриловича

В) Мандельброта

6. Закон \_\_\_\_\_ — эмпирическая закономерность распределения частоты слов естественного языка, суть которого: если к какому-либо достаточно большому тексту составить список всех встретившихся в нем слов, расположить эти слова в порядке убывания частоты их встречаемости в данном тексте и пронумеровать в порядке от 1 (номер наиболее часто встречающегося слова) до  $g$ , то для любого слова произведение его порядкового номера-ранга ( $g$ ) в таком списке и частоты встречаемости слова в тексте ( $f$ ) будет величиной постоянной, имеющей примерно одинаковое значение для любого словарного списка.

А) Ципфа

В) Крушевского-Куриловича

7. \_\_\_\_\_ — отнесение объектов базы данных к заранее определенным категориям.

А) Классификация

В) Кластеризация

8. \_\_\_\_\_ — выделение из базы данных компактных подгрупп объектов с близкими свойствами.

А) Кластеризация

В) Классификация

9. Метод \_\_\_\_\_ анализа квазисинонимичных лексем — метод атрибуции произведений, который заключается в выявлении авторских предпочтений в выборе из группы квазисинонимов — близких по значению слов или устойчивых словосочетаний (фразеологизмов).

А) количественного

В) качественного

10. \_\_\_\_\_ (систематическая) выборка — разновидность случайной выборки, упорядоченная по какому-либо признаку. Первый элемент отбирается случайно, затем, с шагом  $n$  отбирается каждый  $k$  элемент. Размер генеральной совокупности, при  $m N = nk$ .

А) Механическая

В) Автоматическая

11. Модель \_\_\_\_\_ слова А.А. Поликарпова — тенденция к увеличению степени абстрактности значений у многозначного слова — чем позже возникает значение, тем оно более абстрактно.

А) жизненного цикла

В) изменения смысла

12. \_\_\_\_\_ (степень точности) — величина, которая характеризует ширину доверительного интервала, в который попадает относительная частота исследуемого свойства лингвистической единицы.

А) Относительная величина

В) Относительная ошибка

13. \_\_\_\_\_ — характеристика какого-либо параметра в выборке; является не точным, а приближенным значением этого же параметра в генеральной совокупности.

А) Оценка

В) Определение

14. \_\_\_\_\_ — один из способов графического представления плотности вероятности случайной величины. Представляет собой ломаную, соединяющую точки, соответствующие средним значениям интервалов группировки и частотам этих интервалов.

А) Гистограмма

В) Полигон частот

15. \_\_\_\_\_ — отбор единиц в выборочную совокупность производится из генеральной совокупности, разбитой на равные интервалы (группы). При этом размер интервала в генеральной совокупности равен обратной величине доли выборки. В соответствии с принятой долей отбора генеральная совокупность механически разбивается на равновеликие группы. Из каждой такой группы в выборку отбирается лишь одна единица.

А) Механическая выборка

В) Случайная выборка

16. \_\_\_\_\_ — двусторонняя единица, представленная внешне (цепочка фонем, ударение) и внутренне (значение слова).

А) Словарная статья

В) Словоформа

17. \_\_\_\_\_ — такая выборка, которая предполагает однородность генеральной совокупности, одинаковую вероятность доступности всех элементов, наличие полного списка всех элементов.

А) Случайная выборка

В) Сплошная выборка

18. \_\_\_\_\_ - вид словаря, в котором приводятся числовые характеристики употребительности слов (словоформ, словосочетаний) какого-либо языка, в том числе языка писателя, какого-либо произведения.

А) Частотный словарь

В) Translation Memory

19. \_\_\_\_\_ собрание текстов в электронной форме, представляющих данный язык на определенном этапе его существования.

А) Национальный корпус

В) Массив лингвистических данных

20. \_\_\_\_\_ - большой, представленный в машиночитаемом виде, унифицированный, структурированный, размеченный, филологически компетентный массив языковых данных, предназначенный для решения конкретных лингвистических задач.

А) Лингвистический, языковой корпус текстов

В) База данных

Ответы: 1. А; 2. В; 3. В; 4. А; 5. А; 6. А; 7. А; 8. А; 9. А; 10. А; 11. А; 12. В; 13. А; 14. В; 15. А; 16. В; 17. А; 18. А; 19. А; 20. А.

### Тест 3

1. \_\_\_\_\_ — направление лингвистики, в рамках которого изучаются и эксплицируются лингвистические явления с помощью методов количественной математики (теория вероятностей, математическая статистика, теория информации, математический анализ и др.).

А) Квантитативная лингвистика

В) Когнитивная лингвистика

2. \_\_\_\_\_ — численная мера возможности наступления некоторого события.

А) вероятность или вероятностная мера

В) лингвистическая частота

3. \_\_\_\_\_ - часть генеральной лингвистической совокупности, объединенная общим признаком совокупность языковых единиц, выбранная из генеральной совокупности при проведении выборочного наблюдения.

В) выборочная совокупность

В) выборочная лингвистическая совокупность

4. \_\_\_\_\_ — часть объектов генеральной совокупности, отобранная с помощью специальных приемов для получения надежной информации обо всей генеральной совокупности.

А) выборочная совокупность

В) выборочная лингвистическая совокупность

5. \_\_\_\_\_ - несплошное наблюдение, при котором исследуется не вся совокупность языковых единиц, а лишь определенная часть.

А) контент-анализ

В) выборочное наблюдение

6. \_\_\_\_\_ - частоты, показанные отдельными лингвистическими выборками.

А) выборочные частоты

В) ключевые слова

7. \_\_\_\_\_ - множество всех возможных однородных объектов, обладающих признаками (признаком), составляющими предмет анализа.

А) Генеральная совокупность

В) Генеральная лингвистическая совокупность

8. \_\_\_\_\_ - совокупность однородных лингвистических объектов (лингвистических единиц), обладающих признаком/признаками, составляющим/и предмет лингвистического анализа.

А) Генеральная совокупность

В) Генеральная лингвистическая совокупность

9. \_\_\_\_\_ - вероятность, с которой доверительный интервал «захватит» истинное значение параметра генеральной совокупности.

А) Доверительная вероятность

В) Доверительный интервал

10. \_\_\_\_\_ — интервал вокруг выборочного значения определенного параметра, который с заданной (достаточно высокой) вероятностью покрывает истинное значение параметра. Величина доверительного интервала зависит как от доверительной вероятности (т.е. надежности оценивания), так и от объема выборки.

А) Доверительная вероятность

В) Доверительный интервал

11. \_\_\_\_\_ — раздел количественной лингвистики, целью которого является обнаружение факторов порядка в лексике, на основе ранжирования ее по системной и функциональной значимости.
- А) Количественная лексикология  
В) Когнитивная лингвистика
12. \_\_\_\_\_ — термин, относящийся к основному содержанию текста и повторяющийся в нем несколько раз (с учетом всех возможных синонимов).
- А) Ключевое (опорное) слово  
В) Кумулята
13. \_\_\_\_\_ — закономерность функционирования лексики, сутью которой является то, что слова, часто встречающиеся вместе в определенном отрезке текста, теснее связаны между собой по смыслу, чем слова, реже появляющиеся рядом в том же отрезке текста.
- А) Закон семантических связей  
В) Закон Ципфа
14. \_\_\_\_\_ - индивидуальный авторский стиль писателя или поэта, специфика речи художественных произведений, предмет изучения стилистики.
- А) Идиостиль  
В) Стилиевая характеристика
15. \_\_\_\_\_ - наука и технология создания интеллектуальных машин, особенно интеллектуальных компьютерных программ, связанная со сходной задачей использования компьютеров для понимания человеческого интеллекта.
- А) Компьютерная лингвистика  
В) Искусственный интеллект
16. \_\_\_\_\_ - качественно-количественный метод изучения документов, которое характеризуется объективностью выводов и строгостью процедуры и состоит из квантификационной обработки текста с дальнейшей интерпретацией результатов.
- А) контент-анализ  
В) стилистический анализ
17. \_\_\_\_\_ — раздел компьютерной лингвистики, занимающийся разработкой общих принципов построения и использования лингвистических корпусов (корпусов текстов) с применением компьютерных технологий.
- А) Корпусная лингвистика  
В) Количественная лингвистика
18. \_\_\_\_\_ - графическое представление ряда статистических данных, базы данных
- А) Лемма  
В) Кумулята
19. \_\_\_\_\_ - слово в основной, исходной форме, в качестве которой в русском языке выступает обычно именительный падеж единственного числа (для именных форм) и инфинитив (для глагольных форм).
- А) Лингвистическая единица  
В) Лемма
20. \_\_\_\_\_ - большой, представленный в машиночитаемом виде, унифицированный, структурированный, размеченный, филологически компетентный массив языковых данных, предназначенный для решения конкретных лингвистических задач.
- А) Лингвистический, языковой корпус текстов  
В) База данных

Ответы: 1. А; 2. В; 3. В; 4. А; 5. В; 6. А; 7. А; 8. В; 9. А; 10. В; 11. А; 12. А; 13. А; 14. А; 15. В; 16. А; 17. А; 18. В; 19. А; 20. А.

### **Инструкция по выполнению**

В процессе тестирования студент должен ответить на все вопросы. Время выполнения блока заданий составляет 80 минут. Максимальное количество баллов, которое студент может набрать за 1 блок заданий – 30 баллов. Общее количество баллов за семестр – 60.

### **Критерии оценивания:**

За каждый блок заданий студент может получить максимально 30 баллов. Количество баллов за каждый вопрос высчитывается в процентном соотношении от количества вопросов в тесте.

## **Практические задания**

### **Задание 1**

а) ознакомьтесь с опорными материалами соответствующего раздела Практикума; б) рассмотрите возможности использования словарей разных видов для решения Ваших исследовательских задач.

Задание 2. Интернет-ралли по теме «Национальные корпуса текстов». Вам предложены ссылки на наиболее известные национальные корпуса различных языков. Ознакомьтесь с предложенными ресурсами и ответьте на следующие вопросы по каждому корпусу.

1. Какое количество слов содержит данный национальный корпус?
  2. Какие подкорпусы включены в данный национальный корпус?
  3. Каков режим доступа к данному национальному корпусу?
  4. Каковы очевидные преимущества данного национального корпуса?
  5. Каковы очевидные недостатки данного национального корпуса?
- Результаты работы представьте в виде файла программы «Libre Office».

Список национальных корпусов:

1. Британский национальный корпус ([www.natcorp.ox.ac.uk/](http://www.natcorp.ox.ac.uk/) или <http://sara.natcorp.ox.ac.uk/>).
2. Венгерский национальный корпус ([corpus.nytud.hu/mnsz](http://corpus.nytud.hu/mnsz)).
3. Корпус испанского языка (исторический) ([www.corpusdelespanol.org](http://www.corpusdelespanol.org)).
4. Корпус современного датского языка ([www.korpus2000.dk](http://www.korpus2000.dk)).
5. Корпус современного итальянского языка CORIS/CODIS ([www.cilta.unibo.it/ricerca.htm](http://www.cilta.unibo.it/ricerca.htm)).
6. Корпус современного китайского языка (LIVAC Synchronous Corpus) ([www.rcl.cityu.edu.hk/livac](http://www.rcl.cityu.edu.hk/livac)).
7. Мангеймский корпус немецкого языка (Institut für Deutsche Sprache, Mannheim, Germany) ([corpora.ids-mannheim.de/~cosmas](http://corpora.ids-mannheim.de/~cosmas)).
8. Национальный корпус словенского языка ([www.fida.net/eng/](http://www.fida.net/eng/)).

Задание 3. Изучите основные аспекты работы с Национальным корпусом русского языка (НКРЯ). Ознакомьтесь с основными аспектами работы с Национальным корпусом русского языка (НКРЯ). Осуществите пошагово предложенные действия.

### **Задание 4.**

Темы для самостоятельного рассмотрения:

Машинный (автоматический) перевод и его программное обеспечение.

Ресурсы Интернета: программы анализа и лингвистической обработки текстов.

Возможности моделирования речевой коммуникации на основе материалов базы данных CHILDES.

Обзор программ обработки звукового сигнала.

Программа PRAAT.

Лингвистическая экспертиза текста.

### **Задание 5.**

Работа с системой интеллектуального поиска и анализа научных публикаций «Eхactus Expert».

Пройдите по ссылке <http://expert.exactus.ru>, зарегистрируйтесь в системе. Ознакомьтесь с основами работы системы, скачав руководство пользователя.

### **Критерии оценивания:**

Каждый блок заданий оценивается следующим образом

- студенту выставляется 9-10 баллов, если блок заданий выполнен верно.
- студенту выставляется 7-8 баллов, если блок заданий выполнен верно, но допущены незначительные ошибки.
- студенту выставляется 1-6 баллов, если блок заданий выполнен неверно, задание сделано с незначительными искажениями информации, в речи допускаются ошибки.

## Темы докладов с презентацией

1. Машинный (автоматический) перевод и его программное обеспечение.
2. Ресурсы Интернета: программы анализа и лингвистической обработки текстов.
3. Возможности моделирования речевой коммуникации на основе материалов базы данных CHILDES.
4. Обзор программ обработки звукового сигнала.
5. Программа PRAAT.
6. Лингвистическая экспертиза текста.
7. Модели речевой коммуникации в когнитивной науке: модели порождения речи.
8. Моделирование речевой деятельности в когнитивной науке: коннекционистские модели.
9. Модели восприятия речи.
10. Моделирование коммуникации на естественном языке в мультимодальных и мультимедийных системах.

### Инструкция по выполнению

Студент готовит 2 доклада в течение семестра. За 2 доклада студент может максимально набрать 30 баллов. Методические рекомендации по написанию и требования к оформлению содержатся в приложении 2.

### Критерии оценивания:

Каждый доклад оценивается по следующим критериям

- - студенту выставляется 15-13 баллов, если изложенный материал фактически верен, присутствует наличие глубоких исчерпывающих знаний по подготовленному вопросу, в том числе обширные знания в целом по дисциплине; грамотное и логически стройное изложение материала, широкое использование не только основной, но и дополнительной литературы;
- - студенту выставляется 12-11 баллов - изложенный материал верен, наличие полных знаний в объеме пройденной программы по подготовленному вопросу; грамотное и логически стройное изложение материала, широкое использование основной литературы;
- - студенту выставляется 10-8 балла – изложенный материал верен, наличие твердых знаний в объеме пройденной программы по подготовленному вопросу; изложение материала с отдельными ошибками, уверенно исправленными использованием основной литературы;
- - студенту выставляется менее 8 баллов – работа не связана с выбранной темой, наличие грубых ошибок, непонимание сущности излагаемого вопроса.

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

Процедуры оценивания включают в себя текущий контроль и промежуточную аттестацию.

**Текущий контроль** успеваемости проводится с использованием оценочных средств, представленных в п.2 данного приложения. Результаты текущего контроля доводятся до сведения студентов до промежуточной аттестации.

**Промежуточная аттестация** проводится в форме контрольной работы/ экзамена.

Контрольные работы проводятся по окончании теоретического обучения до начала экзаменационной сессии в форме письменной работы по пройденным темам.

Экзамен проводится по расписанию промежуточной аттестации в устной форме. Количество вопросов в экзаменационном задании – 3. Проверка ответов и объявление результатов производится в день экзамена. Результаты аттестации заносятся в экзаменационную ведомость и зачетную книжку студента. Студенты, не прошедшие промежуточную аттестацию по графику сессии, должны ликвидировать задолженность в установленном порядке.

**МЕТОДИЧЕСКИЕ УКАЗАНИЯ ПО ОСВОЕНИЮ ДИСЦИПЛИНЫ**

Учебным планом предусмотрены следующие виды занятий:

- лекции
- практические занятия.

В ходе практических занятий углубляются и закрепляются знания студентов по ряду вопросов, развиваются навыки устной и письменной речи на иностранном языке.

При подготовке к практическим занятиям каждый студент должен:

- изучить рекомендованную учебную литературу;
- подготовить ответы на все вопросы по изучаемой теме.

По согласованию с преподавателем студент может подготовить доклад по теме занятия. В процессе подготовки к практическим занятиям студенты могут воспользоваться консультациями преподавателя.

Вопросы, не рассмотренные на практических занятиях, должны быть изучены студентами в ходе самостоятельной работы. Контроль самостоятельной работы студентов над учебной программой курса осуществляется в ходе занятий методом устного опроса или посредством тестирования. В ходе самостоятельной работы каждый студент обязан прочитать основную и по возможности дополнительную литературу по изучаемой теме, выписать определения основных понятий; законспектировать основное содержание; выписать ключевые слова; выполнить задания-ориентиры в процессе чтения рекомендуемого материала, проанализировать презентационный материал, осуществить обобщение, сравнить с ранее изученным материалом, выделить новое.

**Подготовка доклада с презентацией**

Одной из форм самостоятельной деятельности студента является написание докладов. Выполнение таких видов работ способствует формированию у студента навыков самостоятельной научной деятельности, повышению его теоретической и профессиональной подготовки, лучшему усвоению учебного материала. Темы докладов определяются преподавателем в соответствии с программой дисциплины. Конкретизация темы может быть сделана студентом самостоятельно.

Следует акцентировать внимание студентов на том, что формулировка темы (названия) работы должна быть:

- 1) ясной по форме (не содержать неудобочитаемых и фраз двойного толкования);
- 2) содержать ключевые слова, которые репрезентируют исследовательскую работу;
- 3) быть конкретной (не содержать неопределенных слов «некоторые», «особые» и т.д.);
- 4) содержать в себе действительную задачу;
- 5) быть компактной.

Выбрав тему, необходимо подобрать соответствующий информационный, статистический материал и провести его предварительный анализ. К наиболее доступным источникам литературы относятся фонды библиотеки, а так же могут использоваться электронные источники информации (в том числе и Интернет).

Важным требованием, предъявляемым к написанию докладов на английском языке, является грамотность, стилистическая адекватность, содержательность (полнота отражения и раскрытия темы).

Доклад должен включать такой элемент как выводы, полученные студентом в результате работы с источниками информации.

Доклады представляются строго в определенное графиком учебного процесса время, и их выполнение является обязательным условием для допуска к промежуточному контролю.

Презентация представляет собой публичное выступление, ориентированное на ознакомление, убеждение слушателей по определенной теме-проблеме. Обеспечивает визуально-коммуникативную поддержку устного выступления, способствует его эффективности и результативности.

Качественная презентация зависит от следующих параметров:

1. постановки темы, цели и плана выступления;
2. определения продолжительности представления материала;
3. учета особенностей аудитории, адресности материала;
4. интерактивных действий выступающего (включение в обсуждение слушателей);
5. манеры представления презентации: соблюдение зрительного контакта с аудиторией, выразительность, жестикуляция, телодвижения;
6. наличия иллюстраций (не перегружающих изображаемое на экране), ключевых слов,
7. нужного подбора цветовой гаммы;
8. использования указки.



Преподаватель должен рекомендовать студентам

- 1) не читать написанное на экране;
- 2) обязательно неоднократно осуществить представление презентации дома;
- 3) предусмотреть проблемные, сложные для понимания фрагменты и прокомментировать их;
- 4) предвидеть возможные вопросы, которые могут быть заданы по ходу и в результате предъявления презентации.

