

МИНОБРНАУКИ РОССИИ



Федеральное государственное бюджетное образовательное учреждение
высшего образования
«Российский государственный гуманитарный университет»
(ФГБОУ ВО «РГГУ»)

ИНСТИТУТ ЛИНГВИСТИКИ

Методы лингвистической статистики
РАБОЧАЯ ПРОГРАММА ДИСЦИПЛИНЫ

Направление 45.03.02 Лингвистика
Профиль Перевод и переводоведение/Лингвистика
Уровень квалификации выпускника бакалавр

Форма обучения очная

РПД адаптирована для лиц
с ограниченными возможностями
здоровья и инвалидов

Москва 2021

«Методы лингвистической статистики»

Рабочая программа дисциплины

Составитель(и):

Авторы (составители): к.ф.н. С.И. Гиндин
 д.ф.н. А.Я. Шайкевич

УТВЕРЖДЕНО

Протокол заседания кафедры ТиПЛ

№ 1 от 31.08.21

ОГЛАВЛЕНИЕ

1. Пояснительная записка

1.1 Цель и задачи дисциплины

1.2. Формируемые компетенции, соотнесённые с планируемыми результатами обучения по дисциплине

1.3. Место дисциплины в структуре образовательной программы

2. Структура дисциплины

3. Содержание дисциплины

4. Образовательные технологии

5. Оценка планируемых результатов обучения

5.1. Система оценивания

5.2. Критерии выставления оценок

5.3. Оценочные средства (материалы) для текущего контроля успеваемости, промежуточной аттестации обучающихся по дисциплине

6. Учебно-методическое и информационное обеспечение дисциплины

6.1. Список источников и литературы

6.2. Перечень ресурсов информационно-телекоммуникационной сети «Интернет»

7. Материально-техническое обеспечение дисциплины

8. Обеспечение образовательного процесса для лиц с ограниченными возможностями здоровья и инвалидов

9. Методические материалы

9.1. Планы практических (семинарских, лабораторных) занятий

9.2. Методические рекомендации по подготовке письменных работ

Приложения

Приложение 1. Аннотация дисциплины

Приложение 2. Лист изменений

ПОЯСНИТЕЛЬНАЯ ЗАПИСКА

1. Пояснительная записка

1.1. Цель и задачи дисциплины

Предметом изучения являются статистические и измерительные методы исследования, применяемые в теории языка при синхронном, сравнительном и историческом описании явлений языка и речевой коммуникации¹.

Цель курса - дать слушателям по возможности полную и систематизированную картину всего многообразия существующих количественных методов, подходов, приемов.

Задача курса - показать специфику применения количественных методов в языкознании и поэтике, научить грамотной постановке лингвистических задач для количественного исследования и привить навыки самостоятельного пользования количественными приемами и процедурами.

Содержательными и методологическими особенностями построения курса являются центральное место, занимаемое в нем методами дистрибутивно-статистического анализа, стремление к исторической полноте охвата материала и специальное внимание, уделяемое проблемам текста и поэтической речи.

Форма занятий в основном лекционная, но вопросы внедрения в лингвостатистику новых информационных технологий целиком выносятся на семинары. Большой разброс публикаций и отсутствие учебников и обобщающих монографий предопределяют объем требований к самостоятельной работе слушателей.

2. Формируемые компетенции, соотнесённые с планируемыми результатами обучения по дисциплине:

Коды компетенции	Содержание компетенций	Перечень планируемых результатов обучения по дисциплине
ПК-17	способностью моделировать возможные ситуации общения между представителями различных культур и социумов	знать: знать современную лингвистическую литературу по количественным методам описания и моделирования языка, хорошо представлять историю применения статистических и математических методов в языкознании, знать выгоды и возможные опасности их применения
ПК-23	способностью использовать понятийный аппарат философии, теоретической и прикладной лингвистики, переводоведения, лингводидактики и теории межкультурной коммуникации для решения профессиональных задач	уметь: уметь проводить статистическое обследование лингвистического материала, грамотно применять другие методы количественного и математического моделирования, работать с лингвистическими корпусами и другими репрезентативными массивами данных владеть: всем арсеналом методов количественного исследования языка и речи и современными компьютерными программами такого исследования

¹ Статистические процедуры, применяемые в прикладных разработках, предполагается рассматривать в курсе «Основные проблемы прикладной лингвистики», который аспиранты, обучающиеся по специальности «Теория языка», могут выбрать как факультатив.

1.3. Место дисциплины в структуре образовательной программы

Дисциплина относится к вариативной части блока дисциплин учебного плана.

Для освоения дисциплины необходимы знания, умения и владения, сформированные в ходе изучения следующих дисциплин и прохождения практик: Философия.

В результате освоения дисциплины формируются знания, умения и владения, необходимые для изучения следующих дисциплин и прохождения практик: ГИА.

2. Структура дисциплины (тематический план)

№ п/ п	Раздел дисциплины	Полу- годие обуче- ния	Виды учебной работы, включая самостоятельную работу аспирантов и трудоемкость (в часах)				Формы текущего контроля успеваемости
			Лек- ции	Практ. занятия	Самостоятельна я работа	Контроль	Форма промежуточно й аттестации
1.	Цели специфика и применения количественных методов в языковедении	2	1	2	11 Реферирование русской и зарубежной литературы		
2.	Частотные словари как инструмент лингвистического исследования		1	2	11 Реферирование русской и зарубежной литературы		
3.	Статистическое сравнение лексики как метод обнаружения индивидуальной специфики идиомов		1	4	11 Реферирование русской и зарубежной литературы		
4	Измерение и статистика в исследовании фонетики и поэтической фоники		1	4	11 Реферирование русской и зарубежной литературы		
5.	Теория вероятностей и статистика в изучении метрики и ритмики		1	4	11 Реферирование русской и зарубежной литературы		
6.	Измерение и статистика в изучении морфологии и синтаксиса		1	4	11 Реферирование русской и зарубежной литературы		
7.	Количественные методы в изучении целого текста		1	4	11 Реферирование русской и зарубежной литературы		

8.	Ритмическая и фонетическая структура целого текста		1	4	11 Реферирование русской и зарубежной литературы		
9.	Опыты квантификации в лингвистической типологии		2	4	11 Реферирование русской и зарубежной литературы		Реферат
	ИТОГО:		10	32			Зачет с оценко

СОДЕРЖАНИЕ КУРСА

Тема 1. Цели и специфика применения количественных методов в языковедении

Взаимоотношения «качественного» и «количественного» в научном исследовании. История применения количественных методов в языковедении и в смежных областях гуманитарного знания. Дискуссии о применимости «меры и числа» в исследованиях языка.

Основные направления и области применения математики в современном языковедении. Лингвистические приложения неколичественных разделов математики (общей алгебры, математической логики и т. п.). Основные разновидности проблем, требующих применения количественных методов измерения, массовые наблюдения, сравнения, проверка гипотез. Разделы и методы математики, применимые в количественных исследованиях.

Тема 2. Частотные словари как инструмент лингвистического исследования

Понятия частоты, ранга, распространенности. Алфавитно-частотный и ранговый словники.

Степень покрытия текста различными секторами рангового словника. Методы сравнения частотных словарей.

Статистические закономерности распределения лексики. Статистический спектр словника текста и его структура. Закон Ципфа и его обобщения. Связь ципфовских закономерностей с явлениями отдельности и целостности текста (Ю.А. Шрейдер, Ю.К. Орлов и др.).

Проблемы адекватности построения выборки для создания частотного словаря. Частотные словари и корпусная лингвистика.

Частота слова и его значимость в системе языка и в тексте. Объективные частоты и субъективные представления о них носителей языка. Статистика и отбор лексических минимумов для обучения языку (Г. Гугенейм, И.В. Рахманов и др.).

Тема 3. Статистическое сравнение лексики как метод обнаружения индивидуальной специфики идиомов

Идея отклонения наблюдаемых частот от ожидаемых как основа для выделения диагностических элементов. Формула П. Гиро для выявления «ключевых слов» в художественных идиомах. Ее родство с формулой, лежащей в основе статистического реферирования по Х.П. Луну. Опыты статистической характеристики «ключевых слов» художественного произведения, автора, литературного направления (П. Гиро, Ю.И. Левин, Г. Хетсо, Н.И. Абрамова и др.).

Тема 4. Измерение и статистика в исследовании фонетики и поэтической фоники

Измерения фонетического расстояния и сходства в фонологии. Меры системности и сложности фонологических систем и подсистем.

Измерение сходства рифменных и аллитерационных элементов. Статистический подход к описанию фоники поэтического текста (А.М. Пешковский, А.М. Финкель и др.).

Количественные методы изучения звукового символизма.

Тема 5. Теория вероятностей и статистика в изучении метрики и ритмики

Сравнение статистических распределений ритмических форм с расчетными вероятностными моделями. Модель языковая и модель речевая. Примеры статистических законов в метрике и ритмике.

Тема 6. Измерение и статистика в изучении морфологии и синтаксиса

Статистическое изучение морфологического уровня языков, распределения частей речи. Их использование в задачах атрибуции.

Исследование распределения длин предложений (В. Фукс, Г.А. Лесскис и др.).

Количественные характеристики синтаксической структуры предложения и меры его сложности. Их распределение в тексте как инструмент диагностики и атрибуции (И.П. Севбо и др.).

Тема 7. Количественные методы в изучении целого текста

Меры интенсивности межфразовых связей и их использование для сегментации текста.

Дифференциальные частотные словари в исследовании структуры текста (стилевой, жанровой, персонажной и т. п.).

Контент-анализ и другие методы анализа текста, основанные на повторяемости элементов семантических элементов.

Методы статистической классификации текстов.

Тема 8. Ритмическая и фонетическая структура целого текста

Исследование ритма посредством сравнения с вероятностной моделью. Языковая и речевая модель, правила их построения. Примеры диахронических законов ритма.

Нестатистические числовые модели ритмической и фонетической структуры текстов (А. Белый, Л.Г. Джоунз и др.).

Тема 9. Опыты квантификации в лингвистической типологии

Индексы Дж. Гринберга и результаты их использования. Методика квантитативной типологии С.Е. Яхонтова. Квантификации в синтаксической и семантической типологии.

4. Образовательные технологии

В учебном процессе широко используются как традиционные технологии обучения (лекции, семинары, практические занятия), так и интенсивные инновационные методы обучения (дискуссии, круглые столы, тренинги, «мозговой штурм», обучение с использованием мультимедийной техники и др.).

Самостоятельная работа аспирантов проводится в виде аннотирования и реферирования научной литературы, статей отечественных и зарубежных авторов.

5. Формы контроля и оценочные средства

для текущего контроля успеваемости и промежуточной аттестации по итогам освоения дисциплины

Система текущего и промежуточного контроля знаний аспирантов по дисциплине включает реферат и зачет с оценкой.

Объем реферата по дисциплине - 15-25 страниц печатного текста. При защите реферата аспирант кратко излагает концепцию реферата и основные выводы, отвечает на поставленные вопросы.

Критерии оценки за реферат

Оценка	Содержание
Отлично	Реферат написан четко и грамотно. Тема реферата хорошо раскрыта. Приведена качественно подобранная российская и зарубежная литература. Ответы на дополнительные вопросы по реферату правильные.
Хорошо	Реферат написан четко и грамотно. Тема реферата раскрыта не полностью. Приведена российская и зарубежная литература. Ответы на дополнительные вопросы по реферату правильные.
Удовлетворительно	Тема реферата раскрыта не полностью. Ответы на дополнительные вопросы по реферату правильные, но неполные.
Неудовлетворительно	Тема реферата не раскрыта. Ответы на дополнительные вопросы по реферату неправильные.

Критерии оценки уровня знаний аспирантов по итогам промежуточной аттестации

Оценка	Содержание
Отлично	Аспирант способен обобщить материал, сделать собственные выводы, выразить свое мнение, привести иллюстрирующие примеры.

Хорошо	Ответ аспиранта правильный, но неполный. Не приведены иллюстрирующие примеры, обобщающее мнение аспиранта недостаточно четко выражено.
Удовлетворительно	Ответ правильный в основных положениях, отсутствуют иллюстрирующие примеры, собственное мнение аспиранта, имеются ошибки в деталях.
Неудовлетворительно	В ответе аспиранта существенные ошибки в основных аспектах темы.

Примерная тематика рефератов

5.2. Критерии выставления оценки по дисциплине

Баллы/ Шкала ECTS	Оценка по дисциплине	Критерии оценки результатов обучения по дисциплине
100-83/ A,B	«отлично»/ «зачтено (отлично)»/ «зачтено»	<p>Выставляется обучающемуся, если он глубоко и прочно усвоил теоретический и практический материал, может продемонстрировать это на занятиях и в ходе промежуточной аттестации. Обучающийся исчерпывающе и логически стройно излагает учебный материал, умеет увязывать теорию с практикой, справляется с решением задач профессиональной направленности высокого уровня сложности, правильно обосновывает принятые решения. Свободно ориентируется в учебной и профессиональной литературе.</p> <p>Оценка по дисциплине выставляется обучающемуся с учётом результатов текущей и промежуточной аттестации. Компетенции, закреплённые за дисциплиной, сформированы на уровне – «высокий».</p>
82-68/ C	«хорошо»/ «зачтено (хорошо)»/ «зачтено»	<p>Выставляется обучающемуся, если он знает теоретический и практический материал, грамотно и по существу излагает его на занятиях и в ходе промежуточной аттестации, не допуская существенных неточностей. Обучающийся правильно применяет теоретические положения при решении практических задач профессиональной направленности разного уровня сложности, владеет необходимыми для этого навыками и приёмами. Достаточно хорошо ориентируется в учебной и профессиональной литературе.</p> <p>Оценка по дисциплине выставляется обучающемуся с учётом результатов текущей и промежуточной аттестации. Компетенции, закреплённые за дисциплиной, сформированы на уровне – «хороший».</p>
67-50/ D,E	«удовлетворительно»/ «зачтено (удовлетворительно)»/ «зачтено»	<p>Выставляется обучающемуся, если он знает на базовом уровне теоретический и практический материал, допускает отдельные ошибки при его изложении на занятиях и в ходе промежуточной аттестации. Обучающийся испытывает определённые</p>

Баллы/ Шкала ECTS	Оценка по дисциплине	Критерии оценки результатов обучения по дисциплине
		затруднения в применении теоретических положений при решении практических задач профессиональной направленности стандартного уровня сложности, владеет необходимыми для этого базовыми навыками и приёмами. Демонстрирует достаточный уровень знания учебной литературы по дисциплине. Оценка по дисциплине выставляются обучающемуся с учётом результатов текущей и промежуточной аттестации. Компетенции, закреплённые за дисциплиной, сформированы на уровне – «достаточный».
49-0/ F,FX	«неудовлетворительно»/ не зачтено	Выставляется обучающемуся, если он не знает на базовом уровне теоретический и практический материал, допускает грубые ошибки при его изложении на занятиях и в ходе промежуточной аттестации. Обучающийся испытывает серьёзные затруднения в применении теоретических положений при решении практических задач профессиональной направленности стандартного уровня сложности, не владеет необходимыми для этого навыками и приёмами. Демонстрирует фрагментарные знания учебной литературы по дисциплине. Оценка по дисциплине выставляются обучающемуся с учётом результатов текущей и промежуточной аттестации. Компетенции на уровне «достаточный», закреплённые за дисциплиной, не сформированы.

5.3. Оценочные средства (материалы) для текущего контроля успеваемости, промежуточной аттестации обучающихся по дисциплине

Перечень вопросов к зачету

6. Учебно-методическое и информационное обеспечение дисциплины

6.1. Список источников и литературы

Литература

ИСТОЧНИКИ И ЛИТЕРАТУРА

Литература, имеющаяся в библиотеке РГГУ

ЛИТЕРАТУРА

Реформатский, А. А. Лингвистика и поэтика / А. А. Реформатский. — Москва : Издательство Юрайт, 2019. — 266 с. — (Антология мысли). — ISBN 978-5-534-06480-3. — Текст : электронный // ЭБС Юрайт [сайт]. — URL: <https://biblio-online.ru/bcode/441952> (дата обращения: 07.02.2020).7.

Материально-техническое обеспечение дисциплины

Аудитория для занятий должна быть оснащена компьютером и проектором. Для работы с лингвистическими электронными ресурсами необходим постоянный доступ к сети Интернет.

7. Материально-техническое обеспечение дисциплины

Занятия по курсу можно проводить с максимальной эффективностью в компьютерном классе или аудитории с доступом в Интернет, проектором и экраном для презентаций. Необходимо также наличие доски или флипчарта, чтобы преподаватель мог разбирать примеры по ходу объяснения и записывать задания. Для самостоятельной работы студентам необходимо рабочее место, оборудованное персональным компьютером с доступом в Интернет, аудио- и видеоплеером (Windows Media Player, MPC, WinAmp, VLC и т.п.) а также офисными программами (Microsoft Office, OpenOffice, LibreOffice, Zoho Office и т.п.).

8. Обеспечение образовательного процесса для лиц с ограниченными возможностями здоровья и инвалидов

В ходе реализации дисциплины используются следующие дополнительные методы обучения, текущего контроля успеваемости и промежуточной аттестации обучающихся в зависимости от их индивидуальных особенностей:

- для слепых и слабовидящих:
 - лекции оформляются в виде электронного документа, доступного с помощью компьютера со специализированным программным обеспечением;
 - письменные задания выполняются на компьютере со специализированным программным обеспечением, или могут быть заменены устным ответом;
 - обеспечивается индивидуальное равномерное освещение не менее 300 люкс;
 - для выполнения задания при необходимости предоставляется увеличивающее устройство; возможно также использование собственных увеличивающих устройств;
 - письменные задания оформляются увеличенным шрифтом;
 - экзамен и зачёт проводятся в устной форме или выполняются в письменной форме на компьютере.
- для глухих и слабослышащих:
 - лекции оформляются в виде электронного документа, либо предоставляется звукоусиливающая аппаратура индивидуального пользования;
 - письменные задания выполняются на компьютере в письменной форме;
 - экзамен и зачёт проводятся в письменной форме на компьютере; возможно проведение в форме тестирования.
- для лиц с нарушениями опорно-двигательного аппарата:
 - лекции оформляются в виде электронного документа, доступного с помощью компьютера со специализированным программным обеспечением;
 - письменные задания выполняются на компьютере со специализированным программным обеспечением;
 - экзамен и зачёт проводятся в устной форме или выполняются в письменной форме на компьютере.

При необходимости предусматривается увеличение времени для подготовки ответа.

Процедура проведения промежуточной аттестации для обучающихся устанавливается с учётом их индивидуальных психофизических особенностей. Промежуточная аттестация может проводиться в несколько этапов.

При проведении процедуры оценивания результатов обучения предусматривается использование технических средств, необходимых в связи с индивидуальными особенностями обучающихся. Эти средства могут быть предоставлены университетом, или могут использоваться собственные технические средства.

Проведение процедуры оценивания результатов обучения допускается с использованием дистанционных образовательных технологий.

Обеспечивается доступ к информационным и библиографическим ресурсам в сети Интернет для каждого обучающегося в формах, адаптированных к ограничениям их здоровья и восприятия информации:

- для слепых и слабовидящих:
 - в печатной форме увеличенным шрифтом;
 - в форме электронного документа;

- в форме аудиофайла.
- для глухих и слабослышащих:
 - в печатной форме;
 - в форме электронного документа.
- для обучающихся с нарушениями опорно-двигательного аппарата:
 - в печатной форме;
 - в форме электронного документа;
 - в форме аудиофайла.

Учебные аудитории для всех видов контактной и самостоятельной работы, научная библиотека и иные помещения для обучения оснащены специальным оборудованием и учебными местами с техническими средствами обучения:

- для слепых и слабовидящих:
 - устройством для сканирования и чтения с камерой SARA CE;
 - дисплеем Брайля PAC Mate 20;
 - принтером Брайля EmBraille ViewPlus;
- для глухих и слабослышащих:
 - автоматизированным рабочим местом для людей с нарушением слуха и слабослышащих;
 - акустический усилитель и колонки;
- для обучающихся с нарушениями опорно-двигательного аппарата:
 - передвижными, регулируемые эргономическими партами СИ-1;
 - компьютерной техникой со специальным программным обеспечением.

9. Методические материалы

Рекомендации по организации самостоятельной работы аспирантов

Самостоятельная работа аспирантов организуется в форме аннотирования и реферирования научной литературы, статей отечественных и зарубежных авторов. По итогам самостоятельной работы аспиранты готовят рефераты, лучшие из которых заслушиваются на научном семинаре кафедры. Организация самостоятельной работы аспирантов направлена на осуществление научно исследовательской работы, подготовку научных статей, диссертационной работы, подготовку к преподавательской деятельности.

АННОТАЦИЯ ДИСЦИПЛИНЫ

Дисциплина реализуется в Институте лингвистики *кафедрой* теоретической и прикладной лингвистики.

Цель курса - дать слушателям по возможности полную и систематизированную картину всего многообразия существующих количественных методов, подходов, приемов.

Задача курса - показать специфику применения количественных методов в языкознании и поэтике, научить грамотной постановке лингвистических задач для количественного исследования и привить навыки самостоятельного пользования количественными приемами и процедурами.

Дисциплина направлена на формирование следующих компетенций:

ПК-17 способностью моделировать возможные ситуации общения между представителями различных культур и социумов

ПК-23 способностью использовать понятийный аппарат философии, теоретической и прикладной лингвистики, переводоведения, лингводидактики и теории межкультурной коммуникации для решения профессиональных задач

В результате освоения дисциплины обучающийся должен:

знать:

знать современную лингвистическую литературу по количественным методам описания и моделирования языка, хорошо представлять историю применения статистических и математических методов в языкознании, знать выгоды и возможные опасности их применения

уметь:

уметь проводить статистическое обследование лингвистического материала, грамотно применять другие методы количественного и математического моделирования, работать с лингвистическими корпусами и другими репрезентативными массивами данных

владеть:

всем арсеналом методов количественного исследования языка и речи и современными компьютерными программами такого исследования

По дисциплине предусмотрена промежуточная аттестация в форме *зачета*.

Общая трудоемкость освоения дисциплины составляет 3 зачетных единиц.

ЛИСТ ИЗМЕНЕНИЙ

№	Текст актуализации или прилагаемый к РПД документ, содержащий изменения	Дата	№ протокола
4	Приложение №4	26.06.2020	9

2. Перечень БД и ИСС

Таблица 2

№п/п	Наименование
	Международные реферативные наукометрические БД, доступные в рамках национальной подписки в 2018 г. Web of Science Scopus
	Профессиональные полнотекстовые БД, доступные в рамках национальной подписки в 2018 г. Журналы Cambridge University Press ProQuest Dissertation & Theses Global SAGE Journals Журналы Taylor and Francis Электронные издания издательства Springer
	Профессиональные полнотекстовые БД JSTOR Издания по общественным и гуманитарным наукам
	Компьютерные справочные правовые системы Консультант Плюс, Гарант

2. Образовательные технологии (к п.4 на 2020 г.)

В период временного приостановления посещения обучающимися помещений и территории РГГУ. для организации учебного процесса с применением электронного обучения и дистанционных образовательных технологий могут быть использованы следующие образовательные технологии:

- видео-лекции;
- онлайн-лекции в режиме реального времени;
- электронные учебники, учебные пособия, научные издания в электронном виде и доступ к иным электронным образовательным ресурсам;
- системы для электронного тестирования;
- консультации с использованием телекоммуникационных средств.

3. Перечень БД и ИСС (к п. 6.2 на 2020 г.)

№п /п	Наименование
1	Международные реферативные наукометрические БД, доступные в рамках национальной подписки в 2020 г. Web of Science Scopus
2	Профессиональные полнотекстовые БД, доступные в рамках национальной подписки в 2020 г. Журналы Cambridge University Press ProQuest Dissertation & Theses Global SAGE Journals Журналы Taylor and Francis
3	Профессиональные полнотекстовые БД JSTOR Издания по общественным и гуманитарным наукам Электронная библиотека Grebennikon.ru
4	Компьютерные справочные правовые системы Консультант Плюс, Гарант

4. Состав программного обеспечения (ПО) (к п. 7 на 2020 г.)

№п /п	Наименование ПО	Производитель	Способ распространения (лицензионное или свободно распространяемое)
2	Microsoft Office 2010	Microsoft	лицензионное
3	Windows 7 Pro	Microsoft	лицензионное
9	Microsoft Office 2013	Microsoft	лицензионное
14	Microsoft Office 2016	Microsoft	лицензионное