

Документ подписан простой электронной подписью
Информация о владельце:
ФИО: Наумова Наталия Александровна
Должность: Ректор
Дата подписания: 18.08.2025 18:08:57
Уникальный программный ключ:
6b5279da4e034bff679172803da5b7b559fc69e2

МИНИСТЕРСТВО ПРОСВЕЩЕНИЯ РОССИЙСКОЙ ФЕДЕРАЦИИ
Федеральное государственное автономное образовательное учреждение высшего образования
«ГОСУДАРСТВЕННЫЙ УНИВЕРСИТЕТ ПРОСВЕЩЕНИЯ»
(ГОСУДАРСТВЕННЫЙ УНИВЕРСИТЕТ ПРОСВЕЩЕНИЯ)

Лингвистический факультет
Кафедра переводоведения и когнитивной лингвистики

Согласовано
деканом лингвистического факультета
«18» августа 2025 г.
/Веквищева С.Н./

Рабочая программа дисциплины

Формализация лингвистической информации

Направление подготовки
45.04.02 Лингвистика

Программа подготовки:
Теория перевода и межкультурная/межъязыковая коммуникация (английский язык)

Квалификация
Магистр

Форма обучения
Очная

Согласовано учебно-методической комиссией
лингвистического факультета
Протокол «18» августа 2025 г. № 4
Председатель УМКом _____
/Горбачева О.А./

Рекомендовано кафедрой переводоведения
и когнитивной лингвистики
Протокол от «18» августа 2025 г. № 5
Зав. кафедрой _____
/Жирова И.Г./

Москва
2025

Автор-составитель:

Улиткин Илья Алексеевич – доцент кафедры переводоведения и когнитивной лингвистики, кандидат филологических наук

Рабочая программа дисциплины «Формализация лингвистической информации» составлена в соответствии с требованиями Федерального государственного образовательного стандарта высшего образования по направлению подготовки 45.04.02 Лингвистика, утвержденного приказом МИНОБРНАУКИ РОССИИ от 12.08.2020, № 992.

Дисциплина входит в обязательную часть Блока 1 «Дисциплины (модули)» и является обязательной для изучения.

Реализуется в формате электронного обучения с применением дистанционных образовательных технологий.

Год начала подготовки (по учебному плану) 2025

СОДЕРЖАНИЕ

1. Планируемые результаты обучения.....	4
2. Место дисциплины в структуре образовательной программы.....	4
3. Объем и содержание дисциплины.....	4
4. Учебно-методическое обеспечение самостоятельной работы обучающихся.....	6
5. Фонд оценочных средств для проведения текущей и промежуточной аттестации по дисциплине.....	7
6. Учебно-методическое и ресурсное обеспечение дисциплины.....	16
7. Методические указания по освоению дисциплины.....	16
8. Информационные технологии для осуществления образовательного процесса по дисциплине.....	16
9. Материально-техническое обеспечение дисциплины.....	17

1. ПЛАНИРУЕМЫЕ РЕЗУЛЬТАТЫ ОБУЧЕНИЯ

1.1. Цель и задачи дисциплины

Образовательная цель освоения дисциплины «Формализация лингвистической информации» предполагает повышение уровня культуры и образования студентов путем приобщения их к общенаучному знанию, пониманию необходимости формализации лингвистических методов и использования современных информационных технологий для научных исследований.

Практическая цель состоит в формировании у студентов компетенций, необходимых для анализа явлений языка формальными методами, понимания включенности исследований языка в новые информационные технологии, реализация полученных компетенций в научно-исследовательской и научно-практической деятельности.

Задачи дисциплины:

- актуализация и развитие знаний в области количественных методов исследования языка и новых информационных технологий;
- формирование навыков анализа количественных методов на разных уровнях языка;
- систематизация знаний о языке с позиций формализованной лингвистики.

1.2. Планируемые результаты обучения

В результате освоения дисциплины у студентов формируются следующие компетенции:

УК-1. Способен осуществлять критический анализ проблемных ситуаций на основе системного подхода, вырабатывать стратегию действий;

ОПК-5. Способен осуществлять межкультурное и межкультурное взаимодействие с носителями изучаемого языка в соответствии с правилами и традициями межкультурного профессионального общения, правилами речевого общения в иноязычном социуме;

ОПК-6. Способен применять современные технологии при осуществлении сбора, обработки и интерпретации данных эмпирического исследования; составлять и оформлять научную документацию.

2. МЕСТО ДИСЦИПЛИНЫ В СТРУКТУРЕ ОБРАЗОВАТЕЛЬНОЙ ПРОГРАММЫ

Дисциплина входит в обязательную часть Блока 1 «Дисциплины (модули)» и является обязательной для изучения.

Программа создана на междисциплинарной основе. Она предусматривает интеграцию как общекультурных, так и профессиональных компетенций, академическую автономию студентов, аутентичность учебных материалов, реальные задания и реализацию современных методологических принципов и концепций.

Полученные в процессе обучения знания будут использованы при изучении таких дисциплин как «Прагмалингвистика», «Дискурс художественного текста».

Основные положения дисциплины будут использованы магистрами:

- в дальнейшей образовательной деятельности с учетом соотношения традиций и инноваций в развитии системы обучения переводу;
- при организации образовательной деятельности во время производственной и преддипломной практик;
- при организации собственной научно-исследовательской и учебно-методической деятельности.

3. ОБЪЕМ И СОДЕРЖАНИЕ ДИСЦИПЛИНЫ

3.1. Объем дисциплины

Показатель объема дисциплины	Кол-во часов
Объем дисциплины в зачетных единицах	4
Объем дисциплины в часах	144
Контактная работа	72,2
Лекции	18

Практические занятия	54 (26) ¹
Контактные часы на промежуточную аттестацию:	
Зачет с оценкой	0,2
Самостоятельная работа	64
Контроль	7,8

Форма промежуточной аттестации: зачет с оценкой во 2 семестре на 1 курсе.

3.2. Содержание дисциплины по очной форме обучения

Наименование разделов (тем) Дисциплины с кратким содержанием	Виды занятий	
	Лекции	Практические занятия
Тема 1. Введение и основные понятия квантитативной лингвистики. Основные статистические категории. Выборка, типы переменных в зависимости от уровня лингвистического анализа, элементарное событие. Частота, распределение частот встречаемости единиц разного уровня. Критерии значимости. Типология распределений.	4	8
Тема 2. Применение математических методов на разных уровнях языка. Использование математических методов в фонетике. Фонотактика. Законы Ципфа, новые трактовки. Статистико-комбинаторные, дистрибутивно-статистические и дешифровочные методы в грамматике.	4	10
Тема 3. Квантитативные методы в лексикографии. Частотные словари, словари-индексы, конкордансы. Машинные фонды лексики (корпуса). Зависимость «словарь- текст». Теория вероятностных систем	2	10 (10) ²
Тема 4. Прикладные аспекты квантитативной лингвистики. Задачи атрибуции текста. Математико-статистические методы в стилистике. Семантические сети. Кластерный анализ.	2	8
Тема 5. Системы анализа, представления и обработки лингвистических данных. Data mining. Системы извлечения знаний с машинным обучением. Сентимент-анализ. Количественный контент-анализ.	2	10 (8) ³
Тема 6. Новые информационные и коммуникационные технологии. Интернет-лингвистика. Системы обмена	4	8 (8) ⁴

¹ Реализуется в формате электронного обучения с применением дистанционных образовательных технологий

² Реализуется в формате электронного обучения с применением дистанционных образовательных технологий

³ Реализуется в формате электронного обучения с применением дистанционных образовательных технологий

⁴ Реализуется в формате электронного обучения с применением дистанционных образовательных технологий

информацией. Проблемы обработки данных.		
Итого	18	54 (26) ⁵

4. УЧЕБНО-МЕТОДИЧЕСКОЕ ОБЕСПЕЧЕНИЕ САМОСТОЯТЕЛЬНОЙ РАБОТЫ ОБУЧАЮЩИХСЯ

Темы для самостоятельного изучения	Исучаемые вопросы	Количество часов	Формы самостоятельной работы	Методическое обеспечение	Формы отчетности
Тема 1. Введение и основные понятия квантитативной лингвистики	Элементы лингвостатистики. Выборка, типы переменных в зависимости от уровня лингвистического анализа. Частота, типы частот встречаемости языковых единиц. Типология распределений	12	Чтение литературы по обсуждаемым вопросам Подготовка доклада Подготовка проекта	Рекомендуемая литература Интернет-ресурсы	Доклад Проект
Тема 2. Применение математических методов на разных уровнях языка	Использование математических методов в фонетике, морфологии, синтаксисе, лексикологии. Стилеметрия	12	Чтение литературы по обсуждаемым вопросам Подготовка доклада Подготовка проекта	Рекомендуемая литература Интернет-ресурсы	Доклад Проект
Тема 3. Квантитативные методы в лексикографии	Частотные словари, словари-индексы, конкордансы. Машинные фонды лексики (корпуса) разных языков.	10	Чтение литературы по обсуждаемым вопросам Подготовка доклада Подготовка проекта	Рекомендуемая литература Интернет-ресурсы	Доклад Проект
Тема 4. Прикладные аспекты квантитативной лингвистики	Контент-анализ (количественный и качественный). Семантические сети. Кластерный анализ	10	Чтение литературы по обсуждаемым вопросам Подготовка доклада	Рекомендуемая литература Интернет-ресурсы	Доклад Проект

⁵ Реализуется в формате электронного обучения с применением дистанционных образовательных технологий

			Подготовка проекта		
Тема 5. Системы анализа, представления и обработки лингвистических данных	Извлечение данных. Извлечение мнений, сентимент-анализ. Системы извлечения знаний с машинным обучением	10	Чтение литературы по обсуждаемым вопросам Подготовка доклада Подготовка проекта	Рекомендуемая литература Интернет-ресурсы	Доклад Проект
Тема 6. Новые информационные и коммуникационные технологии	Интернет-лингвистика. Системы обмена информацией. Проблемы распознавания и обработки данных.	10	Чтение литературы по обсуждаемым вопросам Подготовка доклада Подготовка проекта	Рекомендуемая литература Интернет-ресурсы	Доклад Проект
Итого		64			

5. ФОНД ОЦЕНОЧНЫХ СРЕДСТВ ДЛЯ ПРОВЕДЕНИЯ ТЕКУЩЕЙ И ПРОМЕЖУТОЧНОЙ АТТЕСТАЦИИ ПО ДИСЦИПЛИНЕ

5.1. Перечень компетенций с указанием этапов их формирования в процессе освоения образовательной программы: УК-1, ОПК-5, ОПК-6

Код компетенции	Наименование компетенции	Этапы формирования компетенции
УК-1	Способен осуществлять критический анализ проблемных ситуаций на основе системного подхода, вырабатывать стратегию действий	1. Работа на учебных занятиях 2. Самостоятельная работа
ОПК-5	Способен осуществлять межъязыковое и межкультурное взаимодействие с носителями изучаемого языка в соответствии с правилами и традициями межкультурного профессионального общения, правилами речевого общения в иноязычном социуме.	1. Работа на учебных занятиях 2. Самостоятельная работа
ОПК-6	Способен применять современные технологии при осуществлении сбора, обработки и интерпретации данных эмпирического	1. Работа на учебных занятиях 2. Самостоятельная работа

	исследования; составлять и оформлять научную документацию	
--	---	--

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

Оцениваемые компетенции	Уровень сформированности	Этап формирования	Показатели	Критерии оценивания	Шкала оценивания
УК-1	<i>Пороговой</i>	1. Работа на учебных занятиях 2. Самостоятельная работа	Знает методологию критического анализа проблемных ситуаций. Умеет осуществлять системный подход к анализу критических ситуаций.	Тестирование, доклад	Шкала оценивания тестирования Шкала оценивания доклада
	<i>Продвинутой</i>	1. Работа на учебных занятиях 2. Самостоятельная работа	Знает методологию критического анализа проблемных ситуаций. Умеет осуществлять системный подход к анализу критических ситуаций. Владеет стратегией действий в проблемных ситуациях на основе их системного анализа	Тестирование, доклад, проект	Шкала оценивания тестирования Шкала оценивания доклада Шкала оценивания проекта
ОПК-5	<i>пороговой</i>	1. Работа на учебных занятиях. 2. Самостоятельная работа.	Знает правила и традиции межкультурного общения. Умеет применять их на практике.	Тестирование, доклад	Шкала оценивания тестирования Шкала оценивания доклада
	<i>продвинутой</i>	1. Работа на учебных занятиях. 2. Самостоятельная работа.	Знает правила и традиции межкультурного общения. Умеет применять их на практике. Владеет системой необходимых условий для профессионального общения с носителями иностранного языка.	Тестирование, доклад, проект	Шкала оценивания тестирования Шкала оценивания доклада Шкала оценивания проекта
ОПК-6	<i>пороговой</i>	1. Работа на учебных занятиях. 2.	Знает современные технологии сбора, обработки эмпирического	Тестирование, доклад	Шкала оценивания тестирования

		Самостоятел ьная работа.	материала. Умеет грамотно оформлять научную документацию		Шкала оценивания доклада
	<i>продвину тый</i>	1. Работа на учебных занятиях. 2. Самостоятел ьная работа.	Знает современные технологии сбора, обработки эмпирического материала. Умеет грамотно оформлять научную документацию Владеет способностью интерпретации собранных. данных с учетом знания теории языка и развития лингвистики на современном этапе развития науки.	Тестирование, доклад, проект	Шкала оценивания тестировани я Шкала оценивания доклада Шкала оценивания проекта

Описание шкал оценивания

Вид работы	Критерии оценивания
1. Проект	10 баллов , если представленный на защите проект свидетельствует о проведенном самостоятельном исследовании с привлечением различных источников информации, отражает теоретическую и практическую направленность проекта; во время защиты проекта продемонстрированы коммуникативные и рефлексивные умения, а также навыки работы в команде.
	8 баллов , если представленный на защиту проект свидетельствует о проведенном исследовании с привлечением различных источников информации, отражает теоретическую и практическую направленность проекта.
	4 балла , если представленный на защиту проект свидетельствует о проведенном исследовании, отражает теоретическую направленность проекта.
	2 балла , если представленный на защиту проект свидетельствует о проведенном исследовании с привлечением одного источника информации; слабо отражает теоретическую и практическую направленность проекта.
2. Тест	10 баллов , если из всех заданий студент выполнил как минимум 80%
	8 баллов , если из всех заданий студент выполнил как минимум 60%
	4 балла , если из всех заданий студент выполнил 40%
3. Доклад	10 баллов , если представленный доклад свидетельствует о проведенном самостоятельном исследовании с привлечением различных источников информации; логично, связно и полно раскрывается тема; заключение содержит логично вытекающие из содержания выводы; правильно (уместно и достаточно) используются разнообразные средства речи

	8 балла , если представленный доклад свидетельствует о проведенном самостоятельном исследовании с привлечением двух-трех источников информации; логично, связно и полно раскрывается тема; заключение содержит логично вытекающие из содержания выводы; правильно (уместно и достаточно) используются разнообразные средства речи
	4 балла , если представленный доклад свидетельствует о проведенном исследовании с привлечением одного источника информации; тема раскрыта не полностью; логичный вывод не сделан
	0 баллов , если доклад не представлен

5.3. 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

Образец тестовых заданий

1. База данных это
 - а. Корпус текстов, организованных по внутренним правилам
 - б. Именованная совокупность данных, организованных по определенным правилам, предусматривающим общие принципы описания, хранения и манипулирования данными
 - в. Именованная совокупность файлов, организованных по определенным правилам
 - г. Именованный корпус любых данных с возможностью хранения и оперирования данными
2. Дистрибутивно-статистический анализ это
 - а. формальный метод анализа лексики
 - б. формальный метод анализа фонетики
 - в. формальный метод анализа распределения разных языковых единиц
 - г. формальный метод подсчета разных языковых единиц

Образец проекта

1. Провести анализ любого языкового материала (лексической выборки, текстовой выборки и пр.) одним из изученных формальных методов (контент-анализ, дистрибутивно-статистический анализ, сентимент-анализ)

Примерные темы докладов

1. Описание и анализ прикладной математической программы MatLab
 2. Описание и анализ прикладной математической программы Cubicalc.
 3. Описание и анализ автоматической экспертной системы по выбору.
 4. Описание и анализ автоматической обучающей системы по выбору.
 5. Описание и анализ информационно-поисковой системы по выбору.
- Анализ проводится с целью последующего обсуждения на семинарских занятиях. Особое внимание уделяется проблеме эффективности обработки естественного языка такими программами с позиций пользователя и разработчика подобных систем.

Примерные вопросы к зачету с оценкой:

1. Основные статистические категории. Понятие выборки и совокупности, типы переменных в зависимости от уровня лингвистического анализа. Понятие элементарного события.
2. Понятие частоты и распределения частот встречаемости единиц разного уровня. Случайная выборка. Методы оценки представительности распределения. Критерии значимости.

3. Нормальное распределение. Интерпретация лингвистических распределений.
4. Планирование статистического эксперимента и проверка гипотез.
5. Критерий χ^2 .
6. Системная и прагматическая выборка лингвистических данных. Частоты: индивидуальная, частота класса, абсолютная, относительная, условная и позиционная.
7. Понятие и методы лингвистической дешифровки. Эвристические модели.
8. Применение статистических методов в фонетике: построение системы фонем с использованием критерия частоты встречаемости, квантитативные модели слогоделения. Фонотактика. Фонотактические типы слов. Длина слов.
9. Статистико-комбинаторные, дистрибутивно-статистические и дешифровочные методы в грамматике.
10. Метод дистрибутивно-статистического анализа: традиционный и автоматизированный варианты. Использование ДСА для автоматического распознавания жанров текстов в информационном потоке.

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

Промежуточная аттестация по изучаемой дисциплине учитывает уровень результатов обучения, общее качество работы магистранта, дисциплинированность, самостоятельность. Освоение дисциплины оценивается по балльной шкале.

Общее количество баллов по дисциплине – 100 баллов.

Максимальное количество баллов, которое может набрать магистрант в течение семестра за текущий контроль, равняется 70 баллам.

Промежуточная аттестация проводится в форме зачета с оценкой.

Зачет с оценкой проводится в устной форме, магистрант должен дать развернутый ответ на вопросы. Максимальное число баллов, которые выставляются магистранту по итогам зачета с оценкой равняется 30 баллам.

Шкала оценивания зачета с оценкой

Критерии оценивания	Баллы
• Глубокое и прочное усвоение программного материала. Умение выделить существенное, сделать обобщение и выводы. • Полное, логически последовательное, грамотное изложение ответа. • Правильность формулировки понятий и использования терминологии по проблеме. • Умение применить на практике теоретические знания при анализе примеров.	30-11
• Полное и прочное усвоение программного материала. • Грамотное изложение ответа по существу. • Отсутствие грубых неточностей в формулировке понятийного аппарата. Умение применить на практике теоретические знания при анализе примеров.	10 - 6
• Общие знания основного материала без усвоения некоторых существенных теоретических положений. • Формулировка основных лингвистических понятий, с рядом неточностей. • Затруднения в практическом применении теоретических знаний при анализе примеров. • Недостаточные сведения библиографического характера.	5 - 1

• Незнание основной части программного материала. • Существенные ошибки в изложении ответа. • Неумение сделать выводы и обобщения. • Неумение применить на практике теоретические знания. • Отсутствие или пассивное присутствие на практических занятиях.	0
--	---

Итоговая шкала оценивания результатов освоения дисциплины

Итоговая оценка по дисциплине формируется из суммы баллов по результатам текущего контроля и промежуточной аттестации и выставляется в соответствии с приведенной ниже таблицей.

Оценка по 100-балльной системе	Оценка по 5-балльной системе
81 – 100	5 «отлично»
61 - 80	4 «хорошо»
41 - 60	3 «удовлетворительно»
0 - 40	2 «неудовлетворительно»

6. УЧЕБНО-МЕТОДИЧЕСКОЕ И РЕСУРСНОЕ ОБЕСПЕЧЕНИЕ ДИСЦИПЛИНЫ

6.1 Основная литература

1. Баранов, А.Н. Введение в прикладную лингвистику [Текст] / А. Н. Баранов. - 4-е изд., доп. - М. : ЛИБРОКОМ, 2013. - 368с.
2. Волосатова, Т.М. Информатика и лингвистика: Учебное пособие / Волосатова Т.М., Чичварин Н.В. М.: НИЦ ИНФРА-М, 2016. - 196 с. – Режим доступа: <http://znanium.com/catalog.php?bookinfo=508097>.

6.2. Дополнительная литература:

3. Марчук Ю.Н. Основы компьютерной лингвистики [Текст] : учеб.пособие / Ю. Н. Марчук. - М. : МПУ, 2000. - 226с.
4. Арапов М. В. Квантитативная лингвистика. М., 1988
5. Лапшин, В.А. Лекции по математической лингвистике [Текст] / В. А. Лапшин. - М. : Науч.мир, 2010. - 248с.
6. Максименко О.И. Формализованная лингвистика. М., 2013.
7. Максименко О. И. Формальные методы в современной прикладной лингвистике [Текст] / О. И. Максименко. - М. : МГОУ, 2002. - 255с.
8. Потапова Р.Г. Новые информационные технологии и лингвистика. М.: УРСС, 2012.
9. Белоногов Г.Г., Калинин Ю.П., Хорошилов А.А. Компьютерная лингвистика и перспективные информационные технологии. М.: Русский мир, 2004.
10. Корнеев В.В. и др. Базы данных. Интеллектуальная обработка информации. М.: Нолидж, 2000, 351 с.
11. Шайкевич А.Я.. Введение в лингвистику. Учебное пособие. 2-е изд., испр. М.: «Академия», 2010. 394 с.
12. Шемакин Ю.И. Семантика самоорганизующихся систем. – М.: Академический Проект, 2003.

6.3. Ресурсы информационно-телекоммуникационной сети Интернет

www.philology.ru
www.postnauka.ru
www.elibrary.ru
www.krugosvet.ru

7. МЕТОДИЧЕСКИЕ УКАЗАНИЯ ПО ОСВОЕНИЮ ДИСЦИПЛИНЫ

Методические рекомендации по организации самостоятельной работы магистрантов

8. ИНФОРМАЦИОННЫЕ ТЕХНОЛОГИИ ДЛЯ ОСУЩЕСТВЛЕНИЯ ОБРАЗОВАТЕЛЬНОГО ПРОЦЕССА ПО ДИСЦИПЛИНЕ

Лицензионное программное обеспечение:

Microsoft Windows

Microsoft Office

Kaspersky Endpoint Security

Информационные справочные системы:

Система ГАРАНТ

Система «КонсультантПлюс»

Профессиональные базы данных:

fgosvo.ru – Портал Федеральных государственных образовательных стандартов высшего образования

pravo.gov.ru - Официальный интернет-портал правовой информации

www.edu.ru – Федеральный портал Российское образование

Свободно распространяемое программное обеспечение, в том числе отечественного производства

ОМС Плеер (для воспроизведения Электронных Учебных Модулей)

7-zip

Google Chrome

9. МАТЕРИАЛЬНО-ТЕХНИЧЕСКОЕ ОБЕСПЕЧЕНИЕ ДИСЦИПЛИНЫ

Материально-техническое обеспечение дисциплины включает в себя:

- учебные аудитории для проведения занятий лекционного и семинарского типа, групповых и индивидуальных консультаций, текущего контроля и промежуточной аттестации, укомплектованные учебной мебелью, доской, демонстрационным оборудованием;
- помещения для самостоятельной работы, укомплектованные учебной мебелью, персональными компьютерами с подключением к сети Интернет и обеспечением доступа к электронным библиотекам и в электронную информационно-образовательную среду.