

Документ подписан простой электронной подписью

Информация о владельце:

ФИО: Ястребов Олег Александрович

Должность: Ректор

Дата подписания: 27.05.2026 15:21:29

Уникальный программный ключ:

ca953a0120d891083f939673078ef1a989dae18a

Федеральное государственное автономное образовательное учреждение высшего образования

«Российский университет дружбы народов имени Патриса Лумумбы»

Филологический факультет

(наименование основного учебного подразделения (ОУП) – разработчика ОП ВО)

РАБОЧАЯ ПРОГРАММА ДИСЦИПЛИНЫ

КВАНТИТАТИВНАЯ ЛИНГВИСТИКА И НОВЫЕ ИНФОРМАЦИОННЫЕ ТЕХНОЛОГИИ

(наименование дисциплины/модуля)

Рекомендована МССН для направления подготовки/специальности:

45.04.02 ЛИНГВИСТИКА

(код и наименование направления подготовки/специальности)

Освоение дисциплины ведется в рамках реализации основной профессиональной образовательной программы высшего образования (ОП ВО):

КОГНИТИВНАЯ И ЦИФРОВАЯ ЛИНГВИСТИКА

(наименование (профиль/специализация) ОП ВО)

1. ЦЕЛЬ ОСВОЕНИЯ ДИСЦИПЛИНЫ

Дисциплина «Квантитативная лингвистика и новые информационные технологии» входит в программу магистратуры «Когнитивная и цифровая лингвистика» по направлению 45.04.02 «Лингвистика» и изучается во 2 семестре 1 курса. Дисциплину реализует Кафедра прикладного искусственного интеллекта. Дисциплина состоит из 3 разделов и 21 тема и направлена на изучение количественных методов анализа языка и текста с использованием современных цифровых инструментов: корпусная лингвистика, частотный анализ, статистические методы исследования текста, стилометрия, компьютерные инструменты для лингвистов и переводчиков (без программирования), основы обработки естественного языка (NLP), применение ИИ-технологий в лингвистических исследованиях, переводоведении и преподавании языков.

Целью освоения дисциплины является формирование у магистрантов практических навыков количественного анализа текста для лингвистических исследований и прикладных задач: работа с корпусами текстов, расчёт и интерпретация частотных характеристик, сравнительный анализ оригиналов и переводов, использование цифровых инструментов для изучения и преподавания языка, применение ИИ-сервисов для исследовательских и переводческих задач

2. ТРЕБОВАНИЯ К РЕЗУЛЬТАТАМ ОСВОЕНИЯ ДИСЦИПЛИНЫ

Освоение дисциплины «Квантитативная лингвистика и новые информационные технологии» направлено на формирование у обучающихся следующих компетенций (части компетенций):

Таблица 2.1. Перечень компетенций, формируемых у обучающихся при освоении дисциплины (результаты освоения дисциплины)

Шифр	Компетенция	Индикаторы достижения компетенции (в рамках данной дисциплины)
УК-1	Способен осуществлять поиск, критический анализ проблемных ситуаций на основе системного подхода, вырабатывать стратегию действий.	УК-1.1 Анализирует проблемную ситуацию как систему, выявляя ее составляющие и связи между ними; УК-1.2 Определяет пробелы в информации, необходимой для решения проблемной ситуации, и проектирует процессы по их устранению; УК-1.3 Критически оценивает надежность источников информации, работает с противоречивой информацией из разных источников; УК-1.4 Разрабатывает и содержательно аргументирует стратегию решения проблемной ситуации на основе системного и междисциплинарных подходов; УК-1.5 Использует логико-методологический инструментарий для критической оценки современных концепций философского и социального характера в своей предметной области;
УК-7	Способен: искать нужные источники информации и данные, воспринимать, анализировать, запоминать и передавать информацию с использованием цифровых средств, а также с помощью алгоритмов при работе с полученными из различных источников данными с целью эффективного использования полученной информации для решения задач; проводить оценку информации, ее достоверность, строить логические умозаключения на основании поступающих информации и данных.	УК-7.1 Использует цифровые средства для поиска нужных источников информации; УК-7.2 Решает задачи различной сложности с помощью цифровых алгоритмов при работе с полученными данными; УК-7.3 Оценивает источник информации и ее достоверность; УК-7.4 Строит логические умозаключения на основании поступающих информации и данных;
ОПК-6	Способен применять современные технологии при осуществлении сбора, обработки и интерпретации данных эмпирического исследования; составлять и оформлять научную документацию;	ОПК-6.1 Использует современные технологии сбора, обработки и интерпретации полученных экспериментальных данных в профессиональной деятельности; ОПК-6.2 Применяет различные приемы составления

Шифр	Компетенция	Индикаторы достижения компетенции (в рамках данной дисциплины)
		и оформления научной документации;
ОПК-7	Способен работать с основными информационно-поисковыми и экспертными системами и другими системами представления знаний и обработки вербальной информации.	ОПК-7.1 Работает с основными информационно-поисковыми и экспертными системами; ОПК-7.2 Владеет различными системами представления знаний и обработки вербальной информации;
ОПК-8	Способен использовать цифровые технологии и методы в профессиональной деятельности для: изучения и моделирования объектов профессиональной деятельности, анализа данных, представления информации и пр.	ОПК-8.1 Использует современные цифровые технологии и методы поиска, обработки, анализа, хранения и представления информации; ОПК-8.2 Ориентируется в условиях цифровой экономики и современной корпоративной информационной культуры;
ПК-2	Способен разрабатывать современные методики сбора хранения и представления баз данных и знаний в интеллектуальных системах различного назначения.	ПК-2.1 Разрабатывает современные методики сбора, хранения и представления баз данных и знаний в интеллектуальных системах различного назначения;

3. МЕСТО ДИСЦИПЛИНЫ В СТРУКТУРЕ ОП ВО

Дисциплина «Квантитативная лингвистика и новые информационные технологии» относится к обязательной части блока 1 «Дисциплины (модули)» образовательной программы высшего образования.

В рамках образовательной программы высшего образования обучающиеся также осваивают другие дисциплины и/или практики, способствующие достижению запланированных результатов освоения дисциплины «Квантитативная лингвистика и новые информационные технологии».

Таблица 3.1. Перечень компонентов ОП ВО, способствующих достижению запланированных результатов освоения дисциплины

Шифр	Наименование компетенции	Предшествующие дисциплины/модули, практики*	Последующие дисциплины/модули, практики*
УК-7	Способен: искать нужные источники информации и данные, воспринимать, анализировать, запоминать и передавать информацию с использованием цифровых средств, а также с помощью алгоритмов при работе с полученными из различных источников данными с целью эффективного использования полученной информации для решения задач; проводить оценку информации, ее достоверность, строить логические умозаключения на основании поступающих информации и данных.	Информационные базы данных;	Основы перевода устойчивых единиц генетический и ареально далеких языков; Преддипломная практика; Методика обучения переводу;
УК-1	Способен осуществлять поиск, критический анализ проблемных ситуаций на основе системного подхода, вырабатывать стратегию действий.	Новые лингвистические теории; Современные дискурсивные практики; Педагогика и психология высшей школы; Теория и практика перевода; Язык - текст: теория, методология, практика интерпретаций; Информационные базы данных;	Новые лингвистические теории; Методика обучения переводу; Цифровые технологии в переводе; Современные дискурсивные практики; Язык - текст: теория, методология, практика интерпретаций;

Шифр	Наименование компетенции	Предшествующие дисциплины/модули, практики*	Последующие дисциплины/модули, практики*
ОПК-6	Способен применять современные технологии при осуществлении сбора, обработки и интерпретации данных эмпирического исследования; составлять и оформлять научную документацию;	Педагогика и психология высшей школы; Язык - текст: теория, методология, практика интерпретаций; Новые лингвистические теории; История и методология науки "Лингвистика";	Язык - текст: теория, методология, практика интерпретаций; Проблемы современной философии; Новые лингвистические теории;
ОПК-7	Способен работать с основными информационно-поисковыми и экспертными системами и другими системами представления знаний и обработки вербальной информации.	Теория и практика перевода; Язык - текст: теория, методология, практика интерпретаций; История и методология науки "Лингвистика";	Язык - текст: теория, методология, практика интерпретаций; Проблемы современной философии; Методика обучения переводу; Цифровые технологии в переводе; Основы перевода устойчивых единиц генетический и ареально далеких языков;
ОПК-8	Способен использовать цифровые технологии и методы в профессиональной деятельности для: изучения и моделирования объектов профессиональной деятельности, анализа данных, представления информации и пр.	Педагогика и психология высшей школы; Теория и практика перевода; Основы устного перевода; Язык делового общения (второй иностранный язык); Практика перевода (второй иностранный язык);	Language of Business Communication; Цифровые технологии в переводе; Язык делового общения (второй иностранный язык); Практика перевода (второй иностранный язык); Преддипломная практика;
ПК-2	Способен разрабатывать современные методики сбора хранения и представления баз данных и знаний в интеллектуальных системах различного назначения.		Преддипломная практика; Цифровые технологии в переводе;

* - заполняется в соответствии с матрицей компетенций и СУП ОП ВО

** - элективные дисциплины /практики

4. ОБЪЕМ ДИСЦИПЛИНЫ И ВИДЫ УЧЕБНОЙ РАБОТЫ

Общая трудоемкость дисциплины «Квантитативная лингвистика и новые информационные технологии» составляет «2» зачетные единицы.
Таблица 4.1. Виды учебной работы по периодам освоения образовательной программы высшего образования для очной формы обучения.

Вид учебной работы	ВСЕГО, ак.ч.		Семестр(-ы)
			2
<i>Контактная работа, ак.ч</i>	39		39
Лекции (ЛК)	13		13
Лабораторные работы (ЛР)	0		0
Практические/семинарские занятия (СЗ)	26		26
<i>Самостоятельная работа обучающихся, ак.ч.</i>	15		15
<i>Контроль (экзамен/зачет с оценкой), ак.ч.</i>	18		18
Общая трудоемкость дисциплины ак.ч.	ак.ч.	72	72
	зач.ед.	2	2

5. СОДЕРЖАНИЕ ДИСЦИПЛИНЫ

Таблица 5.1. Содержание дисциплины (модуля) по видам учебной работы*

Номер раздела	Наименование раздела дисциплины	Наименование темы		Содержание темы	Вид учебной работы*
Раздел 1	Корпусная лингвистика и частотный анализ	1.1	Введение в количественную лингвистику	Что такое количественная лингвистика: количественные закономерности в языке. Закон Ципфа, закон Хипса. Зачем лингвисту считать: объективность, воспроизводимость, большие данные. Области применения: стилометрия, переводоведение, методика преподавания, когнитивная лингвистика.	ЛК
		1.2	Корпусная лингвистика: принципы и ресурсы	Корпус текстов: типы (основной, параллельный, учебный, диахронический). НКРЯ: структура и возможности. Параллельные корпуса. Английские корпуса: COCA, BNC. Google Ngram Viewer. Разметка: лемматизация, POS-tagging. Частотный анализ, ключевые слова, коллокации.	ЛК
		1.3	Работа с НКРЯ	Практика: поиск в основном и параллельном подкорпусах НКРЯ. График частотности по годам. Сравнение синонимов. Поиск переводных соответствий. Задание: исследовать варианты перевода заданного выражения, представить результаты. Обсуждение: что корпус показывает, а что — нет.	СЗ
		1.4	Частотный анализ в AntConc	Практика: AntConc — частотный список, конкордансы, коллокации. Задание: построить частотный профиль текста, выявить ключевые коллокации. Обсуждение: интерпретация результатов, сравнение текстов разных жанров.	СЗ
		1.5	Voyant Tools для анализа текста	Практика: Voyant Tools — облако слов, графики частот, сравнение текстов. Задание: сравнить оригинал и перевод по частотным характеристикам. Обсуждение: какие различия объяснимы переводческими трансформациями.	СЗ
		1.6	Параллельные корпуса в переводоведении	Практика: Reverso Context, Linguee, параллельный подкорпус НКРЯ. Поиск переводных эквивалентов. Задание: собрать варианты перевода 10 единиц, классифицировать стратегии. Дискуссия: норма vs. вариативность в переводе.	СЗ
		1.7	Презентация мини-исследования	Защита индивидуальных мини-исследований: корпусный анализ языкового явления или сравнительный анализ оригинала и перевода. Формат: 5 мин презентация, вопросы, обсуждение методологии.	СЗ
Раздел 2	Статистические методы и стилометрия	2.1	Описательная статистика текста	Количественные характеристики: объём, длина слова и предложения, лексическое разнообразие (TTR, MATTR). Индексы читабельности. Применение: оценка сложности текстов, анализ стиля переводов, сравнение регистров. Статистическая значимость: p-value, хи-квадрат.	ЛК
		2.2	Стилометрия и атрибуция авторства	Стилометрия: количественное описание индивидуального стиля. Признаки: частотные слова, служебные слова, длина предложений. Методы атрибуции: Delta-метрика. Примеры: споры об авторстве. Применение в переводоведении: «голос переводчика». Сравнение переводов одного текста.	ЛК
		2.3	Расчёт статистик текста	Практика в Google Sheets: подсчёт слов, предложений, TTR, индекс читабельности. Задание: рассчитать показатели для оригинала и перевода, объяснить различия. Обсуждение: что влияет на сложность текста.	СЗ
		2.4	Сравнение частот и ключевые слова	Практика: AntConc — Keyword List. Хи-квадрат онлайн. Задание: выявить ключевые слова перевода на фоне оригинала, интерпретировать добавления/опущения. Дискуссия: о чём говорят ключевые слова.	СЗ
		2.5	Лексическое разнообразие	Практика: TTR и MATTR в AntConc и Sheets. Проблема зависимости от объёма.	СЗ

Номер раздела	Наименование раздела дисциплины	Наименование темы		Содержание темы	Вид учебной работы*
				Задание: сравнить лексическое разнообразие текстов разных переводчиков одного произведения. Обсуждение: богатство vs. простота стиля	
		2.6	Стилометрический анализ	Практика: частотный профиль по служебным словам в Google Sheets. Сравнение текстов. Задание: определить, какой из анонимных текстов ближе к известному автору. Дискуссия: надёжность стилометрии, границы метода.	СЗ
		2.7	Сравнительный анализ переводов	Комплексная практика: два перевода одного произведения. Частотный анализ, ключевые слова, TTR, стилистический профиль. Групповая работа и презентация выводов. Обсуждение: какой перевод «лучше» и можно ли это измерить.	СЗ
Раздел 3	Новые информационные технологии в лингвистике	3.1	Обработка естественного языка (NLP)	Что такое NLP. Задачи: токенизация, лемматизация, POS-tagging, NER, анализ тональности. Инструменты: spaCy, Natasha. Машинный перевод: от статистического к нейронному. Ограничения NLP.	ЛК
		3.2	Большие языковые модели в лингвистике	LLM (GPT, GigaChat, Claude): как работают. Возможности: генерация примеров, аннотирование, перевод, резюмирование. ИИ в переводе: постредактирование. ИИ в преподавании: адаптивные упражнения. Ограничения: галлюцинации, предвзятость.	ЛК
		3.3	Этика ИИ в лингвистике и переводе	Авторство и ИИ: кто автор текста, созданного с помощью ИИ. Академическая честность. Качество машинного перевода: когда доверять, когда проверять. Будущее профессии переводчика и лингвиста. Регулирование ИИ.	ЛК
		3.4	NLP-анализ текста	Практика: spaCy demo / Natasha. Токенизация, лемматизация, POS-tagging, NER. Задание: проанализировать текст, извлечь именованные сущности. Обсуждение: где NLP ошибается и почему.	СЗ
		3.5	Машинный перевод и его оценка	Практика: сравнение Google Translate, DeepL, Яндекс на одном тексте. Критерии оценки. Задание: перевести профессиональный текст тремя системами, оценить, предложить редактуру. Дискуссия: будущее переводчика	СЗ
		3.6	ИИ-ассистент для лингвиста	Практика с ChatGPT / GigaChat: генерация примеров, объяснение различий синонимов, создание упражнений, проверка перевода. Задание: подготовить дидактический материал с помощью ИИ, оценить качество. Обсуждение: где ИИ помогает, где мешает.	СЗ
		3.7	Итоговый проект	Защита комплексных мини-исследований: (а) корпусный анализ; (б) стилометрическое сравнение переводов; (в) оценка машинного перевода; (г) NLP-анализ. Презентация (7 мин), вопросы, обсуждение. Подведение итогов курса.	СЗ

* - заполняется только по ОЧНОЙ форме обучения: ЛК – лекции; ЛР – лабораторные работы; СЗ – практические/семинарские занятия.

6. МАТЕРИАЛЬНО-ТЕХНИЧЕСКОЕ ОБЕСПЕЧЕНИЕ ДИСЦИПЛИНЫ

Таблица 6.1. Материально-техническое обеспечение дисциплины

Тип аудитории	Оснащение аудитории	Специализированное учебное/лабораторное оборудование, ПО и материалы для освоения дисциплины (при необходимости)
Лекционная	Аудитория для проведения занятий лекционного типа, оснащенная комплектом специализированной мебели; доской (экраном) и техническими средствами мультимедиа презентаций.	
Семинарская	Аудитория для проведения занятий семинарского типа, групповых и индивидуальных консультаций, текущего контроля и промежуточной аттестации, оснащенная комплектом специализированной мебели и техническими средствами мультимедиа презентаций.	
Для самостоятельной работы	Аудитория для самостоятельной работы обучающихся (может использоваться для проведения семинарских занятий и консультаций), оснащенная комплектом специализированной мебели и компьютерами с доступом в ЭИОС.	

* - аудитория для самостоятельной работы обучающихся указывается **ОБЯЗАТЕЛЬНО!**

7. УЧЕБНО-МЕТОДИЧЕСКОЕ И ИНФОРМАЦИОННОЕ ОБЕСПЕЧЕНИЕ ДИСЦИПЛИНЫ

Основная литература:

1. Буторина, Е. П. Прикладная лингвистика : учебник для вузов / Е. П. Буторина. — Москва : Издательство Юрайт, 2026. — 148 с. — (Высшее образование). — ISBN 978-5-534-16994-2. — Текст : электронный // Образовательная платформа Юрайт [сайт]. — URL: <https://urait.ru/bcode/600290>

2. Шунейко, А. А. Квантитативная лингвистика и новые информационные технологии : учебник для вузов / А. А. Шунейко, И. А. Авдеенко. — Москва : Издательство Юрайт, 2026. — 347 с. — (Высшее образование). — ISBN 978-5-534-15446-7. — Текст : электронный // Образовательная платформа Юрайт [сайт]. — URL: <https://urait.ru/bcode/588475>

Дополнительная литература:

1. Колмогорова, А. В. Цифровые технологии в лингвистических исследованиях : учебник / А. В. Колмогорова. — Москва : Ай Пи Ар Медиа, 2024. — 177 с. — ISBN 978-5-4497-2373-4. — Текст : электронный // Цифровой образовательный ресурс IPR SMART : [сайт]. — URL: <https://www.iprbookshop.ru/133468.html>

Ресурсы информационно-телекоммуникационной сети «Интернет»:

1. ЭБС РУДН и сторонние ЭБС, к которым студенты университета имеют доступ на основании заключенных договоров

- Электронно-библиотечная система РУДН – ЭБС РУДН <https://mega.rudn.ru/MegaPro/Web>
- ЭБС «Университетская библиотека онлайн» <http://www.biblioclub.ru>
- ЭБС «Юрайт» <http://www.biblio-online.ru>
- ЭБС «Консультант студента» www.studentlibrary.ru
- ЭБС «Знаниум» <https://znanium.ru/>

2. Базы данных и поисковые системы

- Sage <https://journals.sagepub.com/>
- Springer Nature Link <https://link.springer.com/>
- Wiley Journal Database <https://onlinelibrary.wiley.com/>
- Наукометрическая база данных Lens.org <https://www.lens.org>

Учебно-методические материалы для самостоятельной работы обучающихся при освоении дисциплины/модуля*:

1. Курс лекций по дисциплине «Квантитативная лингвистика и новые информационные технологии».

* - все учебно-методические материалы для самостоятельной работы обучающихся размещаются в соответствии с действующим порядком на странице дисциплины **в ТУИС!**

РАЗРАБОТЧИКИ

Должность

РУКОВОДИТЕЛЬ БУП

Заведующий кафедрой

Должность

РУКОВОДИТЕЛЬ ОП ВО

Профессор

Должность

Широкова Е.П.

Фамилия И.О

Подолько П.М.

Фамилия И.О

Ломакина О.В.

Фамилия И.О
