

Министерство науки и высшего образования
Российской Федерации

Федеральное государственное бюджетное
образовательное учреждение высшего образования
«Донецкий государственный университет»

Филологический факультет
Кафедра русского языка

УТВЕРЖДАЮ
проректор

_____ П.А. Машаров
«16» апреля 2026 г.
МП

РАБОЧАЯ ПРОГРАММА ДИСЦИПЛИНЫ
КОМПЬЮТЕРНЫЕ ТЕХНОЛОГИИ СБОРА ЯЗЫКОВОГО
МАТЕРИАЛА

Укрупненная группа направлений подготовки	45.00.00 Языкознание и литературоведение
Программа высшего образования	Программа магистратуры
Направление подготовки	45.04.03 Фундаментальная и прикладная лингвистика
Направленность образовательной программы (профиль)	Фундаментальная и прикладная лингвистика
Квалификация	Магистр
Форма обучения	Очная

Рабочая программа может быть адаптирована для лиц
с ограниченными возможностями здоровья и инвалидов

Донецк 2026

Рабочая программа дисциплины **«Компьютерные технологии сбора языкового материала»** для обучающихся по направлению подготовки 45.04.03 Фундаментальная и прикладная лингвистика (Профиль: Фундаментальная и прикладная лингвистика) составлена на основании Федерального государственного образовательного стандарта высшего образования – магистратура по направлению подготовки 45.04.03 Фундаментальная и прикладная лингвистика, утвержденного приказом Министерства науки и высшего образования Российской Федерации от 12 августа 2020 г. № 993 (с изм. и доп.), Порядка организации и осуществления образовательной деятельности по образовательным программам высшего образования – программам бакалавриата, программам специалитета, программам магистратуры, утвержденного приказом Министерства науки и высшего образования Российской Федерации от 06 апреля 2021 г. № 245 (с изм. и доп.), в соответствии с учебным планом, утвержденным Ученым советом ФГБОУ ВО «ДонГУ» для набора 2026 года.

Разработчик:

зав. кафедрой русского языка,
д-р филол. наук, проф.

В.И. Теркулов

Рабочая программа одобрена на заседании кафедры русского языка
Протокол от 09.04.2026 № 8.

Заведующий кафедрой

В.И. Теркулов

СОГЛАСОВАНО:

Декан филологического факультета
15.04.2026

Н.А. Ярошенко

Учебно-методическая комиссия филологического факультета.
Протокол от 15.04.2026 № 4.
Председатель

А.Н. Стебунова

Руководитель основной
образовательной программы,
канд. филол. наук, доцент
15.04.2026

Н.А. Ярошенко

1. МЕСТО ДИСЦИПЛИНЫ В СТРУКТУРЕ ОБРАЗОВАТЕЛЬНОЙ ПРОГРАММЫ

1.1. Требования к предварительной подготовке обучающихся, предшествующие и сопутствующие дисциплины, на которых основывается изучение данной:

Компьютерная лингвистика, Основы прикладной лингвистики, Методология лингвистических исследований.

1.2. Дисциплины, курсовые работы и практики, для которых освоение данной дисциплины необходимо как предшествующее: Производственная практика: научно-исследовательская работа, Подготовка выпускной квалификационной работы.

2. ОПИСАНИЕ ДИСЦИПЛИНЫ

2.1. Общая характеристика

Наименование показателя	Значение показателя
Название образовательной программы	45.04.03 Фундаментальная и прикладная лингвистика (Профиль: Фундаментальная и прикладная лингвистика)
Шифр и название в соответствии с учебным планом	Б1.Б.М2.3 Компьютерные технологии сбора языкового материала
Часть образовательной программы	Базовая часть
Количество зачетных единиц / всего часов	3,5 / 126

В случае предъявления от обучающегося или его родителя (законного представителя) заявления на обучение по адаптированной образовательной программе высшего образования, подкрепленного заключением психолого-медико-педагогической комиссии (ПМПК) или медико-социальной экспертизы (МСЭ) с рекомендациями создания индивидуальной программы реабилитации и абилитации (ИПРА), данная рабочая программа может быть адаптирована с учетом индивидуальных особенностей здоровья обучающегося.

2.2. Распределение часов по формам и периодам обучения

Форма обучения	курс	семестр	Общее количество часов					Форма контроля
			лекционных	лабораторных	практических	самостоятельной работы + контроль	всего	
Очная	1	2	14	–	14	98	126	зачет

3. ЦЕЛИ ДИСЦИПЛИНЫ

Ознакомить студентов с основами корпусной лингвистики в рамках эмпирического подхода к изучению языка с предметом, задачами, методами и понятийным (терминологическим) аппаратом корпусной лингвистики; сформировать знания о компьютерном инструментарии корпусной лингвистики; научить пользоваться Национальным корпусом русского языка; дать студентам первичное представление о возможностях применения корпусов в лингвистических исследованиях.

Задачи:

- познакомить студентов с особенностями использования специализированных

- программ и ресурсов для обработки, анализа и хранения лингвистических данных.
- на примерах показать, какие новые возможности в исследовании грамматики и лексики языка дает использование корпусных, а также применение современных методов обработки этих данных;
 - изучить типологии и особенности использования программ и компьютерных инструментов лингвистических исследований;
 - познакомить с технологиями и проблемами разметки языковых данных.

4. КОМПЕТЕНЦИИ ОБУЧАЮЩЕГОСЯ, ФОРМИРУЕМЫЕ В РЕЗУЛЬТАТЕ ОСВОЕНИЯ КОМПОНЕНТА ОБРАЗОВАТЕЛЬНОЙ ПРОГРАММЫ, ИХ ИНДИКАТОРЫ И ПЛАНИРУЕМЫЕ РЕЗУЛЬТАТЫ ОБУЧЕНИЯ

Компетенции	Индикаторы	Результаты обучения
ОПК-3. Способен выбирать оптимальные подходы и методы решения конкретных научных и прикладных задач в области лингвистики и информационных технологий	ОПК-3.2. Демонстрирует знание существующих подходов и методов решения конкретных научных и прикладных задач в области лингвистики и информационных технологий	ОПК-3.2.1. Знает типологию программ и компьютерных инструментов сбора языкового материала; схемы проведения лингвистических исследований с использованием специализированных компьютерных программ. ОПК-3.2.2. Умеет анализировать существующие программы и компьютерные технологии сбора языкового материала; проводить лингвистические исследования с использованием компьютерных технологий. ОПК-3.2.3. Владеет навыками работы с современными компьютерными инструментами лингвистических исследований; навыками подготовки исходных данных для проведения лингвистических исследований с использованием компьютерных технологий.

5. ПРОГРАММА ДИСЦИПЛИНЫ

Название темы	Краткое содержание темы (вопросы темы)
Раздел 1.	
Тема 1. Лингвистические программы и ресурсы	1. Понятие и виды лингвистических программ и ресурсов. 2. Программы анализа и лингвистической обработки текстов. 3. Программы преобразования текстов. Психолингвистические программы. Генераторы текстов и "говорящие" программы. 4. Программы и утилиты лингвистического анализа текста. Программы и утилиты статистического анализа текста.
Тема 2. Лингвистические базы данных и знаний	1. Теория знаний. Способы представления знаний в компьютерных системах, работающих с естественным языком. Язык как механизм представления и обработки знаний. 2. Процедурный и декларативный способы представления знаний. 3. Формальные языки представления текста как инструмент его автоматической обработки. Метаязыки.

	4. Виды лингвистических баз данных (ЛБД). Компьютерный архив текстов, корпус текстов, лингвистический (лингвистически размеченный / аннотированный) корпус. Особенности их создания и использования в работе филолога.
Тема 3. Корпусная лингвистика как наука	1. Корпус как поисковая система. 2. История лингвистических корпусов. 3. Корпусная лингвистика как наука. 4. Цели и задачи корпусной лингвистики. 5. Соотношение корпусной лингвистики и компьютерной лингвистики 6. История становления и основные направления корпусной лингвистики.
Тема 4. Структура и особенности составления корпусов	1. Общая характеристика корпусов. 2. Классификация (типология) корпусов по различным основаниям. 3. Понятие разметки 4. Развитие лингвистических корпусов.
Тема 5. Особенности Национального корпуса русского языка	1. История и общая характеристика национального корпуса русского языка. 2. Структура Национального корпуса русского языка. 3. Методики поиска в корпусе. 4. Система корпусов Национального корпуса русского языка. Принципы создания подкорпусов в корпусах Национального корпуса русского языка.
Тема 6. Использование поисковых машин в качестве лингвистических корпусов	1. Общая характеристика поисковых машин Google и Яндекс. 2. Использование поисковых машин Google и Яндекс в качестве лингвистических корпусов. 3. Операторы поисковых машин Google и Яндекс. 4. Использование Google Ngram Viewer в филологических исследованиях
Тема 7. Использование корпусов в научной и педагогической практике	1. Филологические исследования на базе корпусов. 2. Использование корпусов в учебной практике.

6. СТРУКТУРА И СОДЕРЖАНИЕ ДИСЦИПЛИНЫ

6.1. Форма обучения – очная, курс – 1, семестр – 2

Наименования разделов и тем	Количество часов				
	Лекц.	Лабор.	Практ.	СРС+К	Всего
Тема 1. Лингвистические программы и ресурсы	2		2	14	18
Тема 2. Лингвистические базы данных и знаний	2		2	14	18
Тема 3. Корпусная лингвистика как наука.	2		2	14	18
Тема 4. Структура и особенности составления корпусов	2		2	14	18
Тема 5. Особенности Национального	2		2	14	18

корпуса русского языка					
Тема 6. Использование поисковых машин в качестве лингвистических корпусов	2		2	14	18
Тема 7. Использование корпусов в научной и педагогической практике	2		2	14	18
ИТОГО ПО КОМПОНЕНТУ ОПОП	14	0	14	98	126

7. ОЦЕНОЧНЫЕ МАТЕРИАЛЫ (СРЕДСТВА) ДЛЯ ТЕКУЩЕГО КОНТРОЛЯ УСПЕВАЕМОСТИ, ПРОМЕЖУТОЧНОЙ АТТЕСТАЦИИ

7.1. Контрольные вопросы

Раздел 1

Тема 1. Лингвистические программы и ресурсы

1. Понятие и виды лингвистических программ и ресурсов.
2. Программы анализа и лингвистической обработки текстов.
3. Программы преобразования текстов. Психоллингвистические программы. Генераторы текстов и "говорящие" программы.
4. Программы и утилиты лингвистического анализа текста. Программы и утилиты статистического анализа текста.

Тема 2. Лингвистические базы данных и знаний

1. Способы представления знаний в компьютерных системах, работающих с естественным языком. Язык как механизм представления и обработки знаний.
2. Процедурный и декларативный способы представления знаний.
3. Формальные языки представления текста как инструмент его автоматической обработки. Метаязыки.
4. Виды лингвистических баз данных (ЛБД). Компьютерный архив текстов, корпус текстов, лингвистический (лингвистически размеченный / аннотированный) корпус. Особенности их создания и использования в работе филолога.

Тема 3. Корпусная лингвистика как наука

1. Корпус как поисковая система.
2. История лингвистических корпусов.
3. Корпусная лингвистика как наука.
4. Цели и задачи корпусной лингвистики.
5. Соотношение корпусной лингвистики и компьютерной лингвистики
6. История становления и основные направления корпусной лингвистики.

Тема 4. Структура и особенности составления корпусов

1. Общая характеристика корпусов.
2. Классификация (типология) корпусов по различным основаниям.
3. Понятие разметки
4. Развитие лингвистических корпусов.

Тема 5. Особенности национального корпуса русского языка

1. История и общая характеристика национального корпуса русского языка.
2. Структура национального корпуса русского языка.
3. Методики поиска в корпусе.
4. Система корпусов национального корпуса русского языка.
5. Принципы создания подкорпусов в корпусах национального корпуса русского языка.

Тема 6. Использование поисковых машин в качестве лингвистических корпусов

1. Общая характеристика поисковых машин google и яндекс.

2. Использование поисковых машин google и яндекс в качестве лингвистических корпусов.

3. Операторы поисковых машин google и яндекс.

4. Использование google ngram viewer в филологических исследованиях

Тема 7. Использование корпусов в научной и педагогической практике

1. Филологические исследования на базе корпусов

2. Использование корпусов в учебной практике.

8. РАСПРЕДЕЛЕНИЕ БАЛЛОВ, КОТОРЫЕ ПОЛУЧАЮТ ОБУЧАЮЩИЕСЯ

Общая оценка знаний обучающихся по дисциплине проводится по 100-балльной шкале исходя из максимума, приведенного в таблице ниже.

Организационно-учебная работа в аудитории оценивается на основе таких критериев как посещаемость занятий, своевременное и качественное выполнение домашних заданий, активность во время проведения лекционных и практических занятий (участие в обсуждении текущего и пройденного материала, решение задач и т.п.).

Самостоятельная работа оценивается на основе предоставленных на проверку выполненных домашних, индивидуальных заданий с учетом своевременности их предоставления и соответствия требованиям к их выполнению.

Количество баллов за контрольную работу вычисляется как сумма баллов за все входящие в её состав задания. Каждое задание оценивается исходя из максимально возможного количества баллов с учетом правильности выполнения задания, полноты приводимых обоснований.

По результатам работы в семестре обучающийся, набравший не менее 60 баллов, имеет право получить оценку. Те, кто претендует на более высокий балл, проходят промежуточную аттестацию. Максимальное количество баллов на промежуточной аттестации – 100. Общее количество баллов за семестр вычисляется как максимальная из полученных за семестр и на промежуточной аттестации и выставляется согласно принятому порядку.

8.1. Семестр 2

Номера разделов	Виды работ	Максимальное количество баллов
1	Организационно-учебная работа в аудитории	20
	Самостоятельная работа	10
	Контрольная работа	30
Зачет		40
Общий итог за семестр		100

Соответствие баллов оценке

Количество баллов из 100	ECTS	Оценка по пятибалльной шкале	
		Экзамен, дифференцированный зачет	Зачет
90-100	A	отлично	зачтено
80-89	B	хорошо	зачтено
75-79	C		зачтено
70-74	D	удовлетворительно	зачтено
60-69	E		зачтено
35-59	FX	неудовлетворительно	не зачтено
0-34	F		не зачтено

9. МАТЕРИАЛЬНО-ТЕХНИЧЕСКОЕ ОБЕСПЕЧЕНИЕ УЧЕБНОГО ПРОЦЕССА

Учебные занятия проводятся в 1-м учебном корпусе ДонГУ (г. Донецк, ул. Университетская, 24). Для проведения занятий требуется аудитория, оборудованная меловой или маркерной доской, мультимедийный проектор и экран, ноутбук, комплект учебной мебели для студентов, рабочее место преподавателя, выход в Интернет – проводной или с использованием Wi-Fi.

Для самостоятельной работы используются текстовые и электронные ресурсы Научной библиотеки университета и других электронных библиотечных баз данных, учебно-методическое обеспечение, представленное на кафедре.

При изучении дисциплины могут применяться электронное обучение и дистанционные образовательные технологии.

10. РЕКОМЕНДУЕМАЯ ЛИТЕРАТУРА

10.1. Основная литература:

1. Теркулов, В. И. Основы корпусной лингвистики [Электронный ресурс] : учебное пособие / В. И. Теркулов ; ГОУ ВПО "Донецкий национальный университет", Филологический факультет, Кафедра русского языка. – Донецк : ДонГУ, 2019. – Электронные текстовые данные (1 файл).
2. Теркулов В.И. Методики работы с Национальным корпусом русского языка : учебно-методическое пособие / В.И. Теркулов. – Донецк, 2019. – 86 с.
3. Захаров В.П. Корпусная лингвистика: учебник. 3-е изд., перераб. / Захаров В.П., Богданова С.Ю. — СПб.: Изд-во С.-Петербур. ун-та, 2020. — 234 с.

10.2. Дополнительная литература

Грудева Е.В. Корпусная лингвистика : учебное пособие – М. : ФЛИНТА, 2012. – 165 с.

Компьютерная обработка лингвистических данных : учебное пособие / А. В. Всеволодова ; Международная академия бизнеса и новых технологий (МУБиНТ). — Ярославль, 2014. — 88 с. https://elibrary.ru/download/elibrary_28371592_45856301.pdf

Харченко, В. К. Малоизвестные приемы и методики исследования языкового материала / В. К. Харченко. – Москва : Изд-во Лит. ин-та им. А. М. Горького, 2008. – 128с.

11. ИНФОРМАЦИОННЫЕ РЕСУРСЫ

1. **Научная электронная библиотека eLIBRARY.RU:** сайт / НЭБ (ООО). – Москва, 2000- . – URL: <https://elibrary.ru> (дата обращения: 08.04.2026). – Режим доступа: для авторизов. пользователей. –Текст: электронный.

2. **Научная электронная библиотека «КиберЛенинка»:** сайт / Ассоциация «Открытая наука». – Москва, 2014- . – URL: <https://cyberleninka.ru/> (дата обращения: 08.04.2026). – Режим доступа: свободный. – Текст: электронный.

3. **Национальная подписка:** сайт / РЦНИ. – Москва, 2020- . – URL: <https://podpiska.rcsi.science/> (дата обращения: 08.04.2026). – Режим доступа: для авторизов.пользователей ДонГУ. – Текст: электронный.

4. **Национальная электронная библиотека (НЭБ):** федеральная государственная информационная система / Министерство Культуры РФ; Российская государственная библиотека. – Москва, 2019- . – URL: <https://rusneb.ru/> (дата обращения: 08.04.2026). – Режим доступа: свободный. – Текст: электронный.

5. **Образовательная платформа Юрайт:** сайт. – Москва, 2013- . – URL: <https://urait.ru/library/svobodnyy-dostup/> (дата обращения: 08.04.2026). – Режим доступа: издания свободного доступа, для авторизов. пользователей. – Текст: электронный.
6. **Электронная библиотека РГБ:** раздел сайта / РГБ. – Текст: электронный // Российская государственная библиотека: официальный сайт. – <https://ldiss.rsl.ru/electroniclibrary>. – Режим доступа: поиск свободный, электронные документы – Виртуальный читальный зал РГБ на территории НБ ДонГУ.
7. **Электронно-библиотечная система Лань:** сайт / ЭБС Лань. – С.-Петербург, 2011- . – URL: <https://e.lanbook.com> (дата обращения: 08.04.2026). – Режим доступа: издания Сетевой электронной библиотеки, для авторизов. пользователей. – Текст: электронный.
8. **Электронно-библиотечная система ДонГУ:** сайт / ФГБОУ ВО «ДонГУ». – Донецк, 2016- . – URL: <http://library.donnu.ru/> (дата обращения: 08.04.2026). – Режим доступа: свободный. – Текст: электронный.
9. **Электронный каталог** Научной библиотеки ДонГУ: раздел сайта / НБ ДонГУ. – Текст: электронный // ЭБС ДонГУ: сайт. – URL: <http://library.donnu.ru/catalog/> (дата обращения: 08.04.2026). – Режим доступа: поиск свободный, электронные документы – для пользователей ДонГУ.
10. **Электронный архив ДонГУ:** раздел сайта / НБ ДонГУ. – Текст: электронный // ЭБС ДонГУ: сайт. – URL: <http://repo.donnu.ru/> (дата обращения: 08.04.2026). – Режим доступа: свободный.

12. ПРОГРАММНОЕ ОБЕСПЕЧЕНИЕ

1. РЕД ОС, Astra Linux (ОЕМ).
2. OpenOffice, LibreOffice.