

**ГОУ ВПО «ДОНЕЦКИЙ НАЦИОНАЛЬНЫЙ УНИВЕРСИТЕТ»
ФИЛОЛОГИЧЕСКИЙ ФАКУЛЬТЕТ
КАФЕДРА СЛАВЯНСКОЙ ФИЛОЛОГИИ И ПРИКЛАДНОЙ
ЛИНГВИСТИКИ**



УТВЕРЖДАЮ:

проректор по научно-методической
и учебной работе

Е.И. Скафа

23 апреля 2020 г.

МП

**РАБОЧАЯ ПРОГРАММА УЧЕБНОЙ ДИСЦИПЛИНЫ
«Автоматическая обработка естественных языков»**

Направление подготовки:	45.03.03 Фундаментальная и прикладная лингвистика
Профиль подготовки:	
Образовательная программа:	Бакалавриат
Квалификация:	Академический бакалавр
Форма обучения:	очная, заочная

Донецк 2020

УТВЕРЖДАЮ:

И.о. декана филологического факультета

Квашина Л.П.

«22» апреля 2020 г.



Программа учебной дисциплины «Автоматическая обработка естественных языков» составлена на основе ГОС ВПО по направлению подготовки 45.03.03 Фундаментальная и прикладная лингвистика, утверждённого приказом Министерства образования и науки Донецкой Народной Республики № 452 от 20.04.2016 г., Порядка организации учебного процесса в образовательных организациях высшего профессионального образования Донецкой Народной Республики, утверждённого приказом Министерством образования и науки ДНР № 1171 от «10» ноября 2017 г.; учебных планов по направлению подготовки 45.03.03 Фундаментальная и прикладная лингвистика (формы обучения: очная и заочная), разработанных в ГОУ ВПО «Донецкий национальный университет».

Разработчик:

старший преподаватель кафедры славянской филологии и
прикладной лингвистики

С.А. Могила

Программа учебной дисциплины утверждена на заседании кафедры славянской филологии и прикладной лингвистики

Протокол № 9 от «16» апреля 2020 г.

И.о. заведующего кафедрой

И.А. Кудрейко

Программа учебной дисциплины одобрена учебно-методической комиссией филологического факультета.

Протокол № 9 от «20» апреля 2020 года

Председатель учебно-методической комиссии факультета

Н.А. Шокотко

1. ОБЛАСТЬ ПРИМЕНЕНИЯ И МЕСТО ДИСЦИПЛИНЫ В УЧЕБНОМ ПРОЦЕССЕ

Учебная дисциплина «Автоматическая обработка естественных языков» относится к циклу базовой части профессионального блока и состоит из двух модулей: Содержательный модуль 1. Научные основы автоматической обработки естественной речи. Содержательный модуль 2. Синтез и распознавания речи в компьютерных системах.

Для изучения данной учебной дисциплины необходимы знания, умения и навыки, формируемые предшествующими и сопутствующими дисциплинами «Основы информатики», «Введение в языкознание», «Русский язык и культура речи», «Основы программирования», «Введение в прикладную лингвистику», «Современный украинский язык».

Дисциплины, для которых данная дисциплина является базовой: «Основы создания компьютерных словарей», «Компьютерная лингвистика», «Методика преподавания информатики», «Компьютерные технологии в филологии», «Компьютерный анализ текста».

2. СТРУКТУРА ДИСЦИПЛИНЫ

<i>Характеристика учебной дисциплины</i>		
Направление подготовки	45.03.03 Фундаментальная и прикладная лингвистика	
Профиль		
Образовательная программа	Бакалавриат	
Квалификация	Академический бакалавриат	
Количество содержательных модулей	2	
Дисциплина базовой / вариативной части образовательной программы	Дисциплина базовой части, профессиональный блок	
Формы контроля	2 модульных контроля, 1 зачет, 1 экзамен	
Показатели	очная форма обучения	заочная форма обучения
Количество зачетных единиц (кредитов)	10	10
Год подготовки	2	2
Семестр	3,4	3,4
Количество часов	360	360
- лекционных	68	14
- практических, семинарских		
- лабораторных	118	22
- самостоятельной работы	174	324
в т.ч. индивидуальное задание		
Недельное количество часов,	20/22,5	
в т.ч. аудиторных	5/6	

3. ОПИСАНИЕ ДИСЦИПЛИНЫ

Цели и задачи

Цель курса: ознакомить студентов с современным состоянием и достижениями в области автоматического обработки текста, распознавание, анализа и синтеза речи в звуковой и текстовой форме; научить пользоваться отечественными и зарубежными лингвистическими программными продуктами (экспертными, диалоговыми, учебными, информационными поисковыми системами).

Задача курса:

- ознакомление студентов с историей, теоретическими и практическими вопросами, задачами прикладной морфологии как составной части прикладной лингвистики;
- ознакомление студентов с актуальными проблемами и перспективами информационных технологий и систем, связанных с автоматизированной обработкой лингвистических данных, в частности, на морфологическом уровне;

- формирование у студентов понимания основ компьютеризации и современного информационного обеспечения и системы знаний, которая помогла бы ориентироваться в информационно-аналитической деятельности;

- формирование теоретических знаний и практических навыков профессиональной работы, которые необходимы для использования универсальных и специализированных информационных технологий, и систем в сфере профессиональной деятельности: работа с системами автоматического перевода, распознавания текста, поисковыми системами и т.п.

Требования к результатам освоения дисциплины: Процесс изучения дисциплины направлен на формирование элементов следующих компетенций в соответствии с ГОС ВПО по данному направлению подготовки (профилю):

а) общекультурных (ОК): (ОК-1, ОК-2, ОК-5, ОК-7)

- способность использовать основы философских знаний для формирования мировоззренческой позиции (ОК-1);
- способность анализировать основные этапы и закономерности исторического развития общества для формирования гражданской позиции (ОК-2);
- способность к коммуникации в устной и письменной формах на русском и иностранном языках для решения задач межличностного и межкультурного взаимодействия (ОК-5);
- способность к самоорганизации и самообразованию (ОК-7);

б) общепрофессиональных (ОПК): (ОПК-1, ОПК-2, ОПК-4, ОПК-5, ОПК-6)

- способность использовать теоретические знания и методы исследования на практике (ОПК-1);
- владение базовыми знаниями в области информационных технологий (программные продукты, используемые в управлении документами, системы электронного документооборота, технологии сканирования документов) (ОПК-2);
- владение навыками использования компьютерной техники и информационных технологий в поиске источников и литературы, использовании правовых баз данных, составлении библиографических и архивных обзоров (ОПК-4);
- владение знаниями в области правил публикации исторических источников и оперативного издания документов (ОПК-5);
- способность решать стандартные задачи профессиональной деятельности на основе информационной и библиографической культуры с применением информационно-коммуникационных технологий и с учетом основных требований информационной безопасности (ОПК-6).

в) профессиональных (ПК): (ПК-1, ПК-2, ПК-11, ПК-12)

научно-исследовательская деятельность:

- способность применять научные методы при исследовании объектов профессиональной деятельности (ПК-1);
- владение основами информационно-аналитической деятельности и способностью применять их в профессиональной сфере (ПК-2);
- владение навыками реферирования и аннотирования научной литературы, навыками редакторской работы (ПК-11);
- способность выявлять и отбирать документы для разных типов и видов публикаций (ПК-12);

В результате изучения учебной дисциплины студент должен
знать:

- основные задачи автоматической обработки звукового вещания;
- характеристики аппаратного звукового обеспечения компьютера;
- типы носителей звуковой информации;
- принципы образования звуков;
- способы представления речевого сигнала;
- акустические свойства звуков;

- основные единицы сегментации речевого потока;
- принципы спектрографии речевого сигнала;
- структуру автоматизированной системы обработки текста
- принципы оптимизации процесса построения конкретной АСОТ;
- процедуры экспликационного и информационного анализа текста;
- виды стандартных морфологических, синтаксических, семантических алгоритмов;
- основные методы синтеза текста;
- принципы автоматического реферирования, индексирования, аннотирования, редактирования.

уметь:

- настраивать аппаратное и программное обеспечение компьютера;
- характеризовать акустические свойства звуков;
- членить речевой поток на единицу, определять тип сегментной единицы;
- устанавливать варианты реализации фонем различных типов;
- строить алгоритмы морфологического, синтаксического, семантического анализа текста с использованием различных методов;
- создавать типовые модели диалоговых, учебных, экспертных систем, баз лингвистических данных;
- использовать программы распознавания, порождение звукового речи и текста;
- создавать и анализировать частотные словари к текстам различных типов;
- строить алгоритмы автоматического реферирования, индексирования, аннотирования текста.

владеть: навыками применения различного программного обеспечения для реализации поставленных прикладных задач в области автоматической обработки естественных языков.

4. СОДЕРЖАНИЕ ДИСЦИПЛИНЫ И ФОРМЫ ОРГАНИЗАЦИИ УЧЕБНОГО ПРОЦЕССА

Курс дисциплины «Автоматическая обработка естественных языков» предусматривает следующие формы организации учебного процесса: лекции, лабораторные работы, самостоятельная работа студента.

Материал излагается с использованием объяснительно-иллюстративных, эвристических и исследовательских методов преподавания. При проведении лекций для обсуждения материала широко используются мультимедийные презентации, анимации, а также раздаточные материалы.

В учебном процессе широко применяются активные и интерактивные формы проведения занятий (разбор конкретных ситуаций, дискуссия, полемика), внеаудиторная самостоятельная работа, балльно-рейтинговая система оценки успеваемости, личностно-ориентированное обучение, проблемное обучение, блочно-модульное обучение.

Использование в учебном процессе Интернет-ресурсов по данному курсу; прикладного программного обеспечения; рассмотрение задач, максимально приближенных к конкретным научно-исследовательским ситуациям; лабораторные работы.

Самостоятельная работа студентов предусматривает изучение учебной и методической литературы, работа с ПО, поиск альтернативного решения типовых задач.

Порядковый номер и тема	Краткое содержание темы
	<i>Содержательный модуль 1. Научные основы автоматической обработки письменного текста</i>
<i>Тема 1.</i>	<i>Язык как знаковая система. Язык – средство передачи информации.</i>

	Формы коммуникации. Естественные и формальные языки. Теория автоматической обработки текста. Роль компьютера как инструмента познания действительности. Основные направления компьютерной лингвистики.
Тема 2.	Техническая обработка текста (word processing). Система автоматического распознавания текстов. Алгоритмы распознавания. Сканирование текста.
Тема 3.	Автоматическая обработка текста как совокупности слов (word matching). Автоматическое индексирование. Автоматическая классификация документов. Создание словарей, конкордансов, частотных словарей.
Тема 4.	Система OCR оптического распознавания символов. Основные принципы работы системы. Порядок распознавания текста. Система оптического распознавания текстов ABBYY FineReader
Тема 5.	Автоматизация лингвистического анализа текста. Морфологический анализ текста с использование приложений MS Office. Синтаксический анализ текста с использование приложений MS Office.
Тема 6.	Автоматизация письменного перевода текста. Основные виды письменного перевода. Основные подходы к переводу. Алгоритм автоматизированного перевода письменного текста. Перевод письменного текста в среде Microsoft Office. Использование автоматизированных переводчиков для письменного перевода текста. Перевод и локализация сайтов
Тема 7.	Компрессия письменного текста. Информационная компрессия. Основные варианты компрессии письменного текста. Семиотические и коммуникативные способы информационной компрессии. Автоматическое реферирование и аннотирование. Алгоритм построения реферата и аннотации. Системы автоматического реферирования и аннотирования.
Тема 8.	Тезаурусы. История создания тезаурусов. Основные понятия. Основные отношения в тезаурусе. Основные функции и задачи тезауруса. Использование тезауруса в приложениях MS Office.
	Содержательный модуль 2. Синтез и распознавания речи в компьютерных системах.
Тема 9.	Синтез речи. Основные понятия синтеза речи. История синтеза речи. Области применения синтеза речи. Способы синтеза речи. Упрощенная технология синтеза речи.
Тема 10.	Распознавание речи. История систем распознавания речи. Область применения систем распознавания речи. Звуковое оборудование компьютера. Системы распознавания речи.
Тема 11.	Шумоочистка речевых сигналов. Основные задачи. Общая схема шумоочистки. Программные комплексы редактирования, анализа и шумоочистки звуковых (речевых) сигналов.
Тема 12.	Оценочная лексика. Основные группы стилистической окраски лингвистической единицы. Функция воздействия и степень выразительности языковых средств. Фоносемантическая оценка словоформ и письменного текста. Программа ВААЛ.

Тематический план

Содержательный модуль 1. <i>Научные основы автоматической обработки письменного текста</i>												
Названия содержательных модулей и тем	Количество часов											
	Очная форма						Заочная форма					
	Всего	лекции	практические	лабораторные	самостоятельная работа	индивидуальная работа	всего	лекции	практические	лабораторные	самостоятельная работа	индивидуальная работа
<i>Тема 1. Язык как знаковая система.</i>	20	4		6	10		22	1		1	20	
<i>Тема 2. Техническая обработка текста (word processing).</i>	24	4		8	12		22	1		1	20	
<i>Тема 3. Автоматическая обработка текста как совокупности слов (word matching).</i>	24	4		8	12		23	1		2	20	
<i>Тема 4. Система OCR оптического распознавания символов.</i>	22	4		8	10		28	1		2	25	
<i>Тема 5. Автоматизация лингвистического анализа текста.</i>	24	4		8	12		23	1		2	20	
<i>Тема 6. Автоматизация письменного перевода текста.</i>	30	6		12	12		28	1		2	25	
<i>Тема 7. Компрессия письменного текста.</i>	30	6		12	12		28	1		2	25	
<i>Тема 8. Тезаурусы.</i>	22	4		8	10		28	1		2	25	
Итого по содержательному модулю 1:	196	36		70	90		202	8		14	180	

Содержательный модуль 2. <i>Синтез и распознавания речи в компьютерных системах.</i>												
Названия содержательных модулей и тем	Количество часов											
	Очная форма						Заочная форма					
	Всего	лекции	практические	лабораторные	самостоятельная работа	индивидуальная работа	всего	лекции	практические	лабораторные	самостоятельная работа	индивидуальная работа
<i>Тема 9. Синтез речи.</i>	42	8		12	22		44	2		2	40	

Тема 10. Распознавание речи.	42	8		12	22		44	2		2	40	
Тема 11. Шумоочистка речевых сигналов	40	8		12	20		43	1		2	40	
Тема 12. Оценочная лексика.	40	8		12	20		27	1		2	24	
Итого по содержательному модулю 2:	164	32		48	84		158	6		8	144	
ВСЕГО по дисциплине	360	68		118	174		360	14		22	324	

5. МЕТОДИЧЕСКИЕ РЕКОМЕНДАЦИИ ДЛЯ ПРОВЕДЕНИЯ ЛЕКЦИОННЫХ, ПРАКТИЧЕСКИХ И ЛАБОРАТОРНЫХ ЗАНЯТИЙ

Темы лекционных занятий

№ п/п	Название темы	Количество часов
1.	Язык как знаковая система.	4
2.	Техническая обработка текста (word processing).	4
3.	Автоматическая обработка текста как совокупности слов (word matching).	4
4.	Система OCR оптического распознавания символов.	4
5.	Автоматизация лингвисти-ческого анализа текста.	4
6.	Автоматизация письменно-го перевода текста.	6
7.	Компрессия письменного текста.	6
8.	Тезаурусы.	4
9.	Синтез речи.	8
10.	Распознавание речи.	8
11.	Шумоочистка речевых сигналов.	8
12.	Оценочная лексика.	8
	ВСЕГО	68

Темы лабораторных занятий

№ п/п	Название темы	Количество часов
1.	Язык как знаковая система.	6
2.	Техническая обработка текста (word processing).	8
3.	Автоматическая обработка текста как совокупности слов (word matching).	8
4.	Система OCR оптического распознавания символов.	8
5.	Автоматизация лингвисти-ческого анализа текста.	8
6.	Автоматизация письменно-го перевода текста.	12
7.	Компрессия письменного текста.	12
8.	Тезаурусы.	8
9.	Синтез речи.	12
10.	Распознавание речи.	12

11.	Шумоочистка речевых сигналов.	12
12.	Оценочная лексика.	12
	ВСЕГО	118

6. МЕТОДИЧЕСКИЕ РЕКОМЕНДАЦИИ ПО ОРГАНИЗАЦИИ САМОСТОЯТЕЛЬНОЙ РАБОТЫ СТУДЕНТОВ

Организация самостоятельной работы студентов

(соответственно данным в таблице тематического плана)

№ п/п	Название темы	Количество часов
	Содержательный модуль 1. Научные основы автоматической обработки письменного текста.	
1.	Сканирование в среде программного обеспечения сканера.	10
2.	Система оптического распознавания текстов ABBYY FineReader.	12
3.	Поиск текста в документе Word.	12
4.	Виды переводов.	10
5.	Методы автоматического реферирования и аннотирования текстов.	12
6.	Краткое изложение информации в реферате и аннотации.	12
7.	Построение реферата человеком.	12
8.	Функция «Тезаурус» в Microsoft Word.	10
	Содержательный модуль 2. Использование информационных технологий в лингвистике.	
9.	История синтеза речи. Звуковое оборудование компьютера.	22
10.	Сферы деятельности с использованием системы распознавания речи.	22
11.	Поиск в среде Интернет программ для синтеза речи.	20
12.	Поиск в среде Интернет программ для распознавания речи и шумоочистки.	20
	ВСЕГО	174

7. ИНДИВИДУАЛЬНЫЕ ЗАДАНИЯ

Содержательный модуль 1.

1. Ознакомиться с основными функциями пункта Вставка основного меню текстового редактора Word.
2. Освоить работу с разделом Стили ленты Главная текстового редактора Word.
3. Изучить порядок работы с формулами в табличном процессоре Excel.
4. Найти в среде Интернет систему OCR и описать порядок работы с ней.
5. Основные составляющие лингвистического анализа произведения.
6. Найти в среде Интернет онлайн-переводчик и описать работу с ним.
7. Ознакомиться с работой конверторов в среде Интернет.
8. Выучить форматы файлов основных приложений MS Office.

Содержательный модуль 2.

1. Выполнить настройку программы ОРФО для выполнения реферирования письменного текста.
2. КТ для построения алфавитно-частотного словаря.
3. Найти в среде Интернет систему синтеза речи и описать ее работу.
4. Найти в среде Интернет систему распознавания речи и описать ее работу.

8. КОНТРОЛЬНЫЕ ВОПРОСЫ К ПРОМЕЖУТОЧНОЙ АТТЕСТАЦИИ

Вопросы к зачету

1. Основные понятия языка как знаковой системы для порождения, передачи и хранения информации.
2. Формальные языки. Основное отличие формальных языков от естественных.
3. Компьютерная лингвистика как новый этап в развитии теоретической и прикладной лингвистики.
4. Основные этапы автоматической обработки письменного текста.
5. Техническая обработка текста (word processing). Основные этапы.
6. Основные принципы работы системы OCR (Optical Character Recognition).
7. Общая характеристика системы оптического распознавания текстов ABBYY FineReader.
8. Основные этапы работы системы оптического распознавания текстов ABBYYFineReader.
9. Поиск текста в документе. Работа с панелью Навигация.
10. Замена текста в документе. Использование окна «Найти и Заменить».
11. Основные виды письменного перевода.
12. Критерий «прозрачности перевода». Основные понятия.
13. Основные разновидности письменного перевода.
14. Виды переводов.
15. Автоматизация лингвистического анализа текста.
16. Автоматизация морфологического анализа.
17. Автоматизация синтаксического анализа.
18. Тезаурусы. Основные понятия.
19. Основные функции тезаурусов.
20. Компрессия текста. Основные понятия.
21. Компрессия текста.
22. Основные типы компрессированных вариантов представления текста.
23. Компрессия текста с использованием набора ключевых слов.
24. Семиотические способы информационной компрессии.
25. Коммуникативные способы информационной компрессии.
26. Этапы построения реферата человеком.
27. Автоматическое реферирование и аннотирование. Основные виды рефератов.
28. Методы автоматического реферирования и аннотирования.
29. Порядок составления реферата или аннотации при помощи компьютера.
30. Характеристика различных групп методов автоматического реферирования и аннотирования текстов.
31. Статистическая группа методов автоматического реферирования и аннотирования текстов.
32. Модель линейных весовых коэффициентов при автоматическом реферировании и аннотировании.
33. Модель реферирующих систем с использованием базы знаний при автоматическом реферировании и аннотировании.
34. Алгоритм автоматического построения реферата и аннотации письменного текста.
35. Использование программы ОРФО для проведения лингвистического анализа текста.

9. ОБРАЗЕЦ МОДУЛЬНОГО КОНТРОЛЯ

ГОУ ВПО «ДОНЕЦКИЙ НАЦИОНАЛЬНЫЙ УНИВЕРСИТЕТ»

Филологический факультет

<i>Направление подготовки:</i>	45.03.03 Фундаментальная и прикладная лингвистика
<i>Программа подготовки:</i>	бакалавриат
<i>Семестр:</i>	4
<i>Учебная дисциплина:</i>	Автоматическая обработка естественных языков

**МОДУЛЬНАЯ КОНТРОЛЬНАЯ РАБОТА
ВАРИАНТ №1**

1. Язык – это ... система, предназначенная для порождения, передачи и хранения информации.
2. В коммуникации участвуют две формы передачи информации:
 - а) звуковая и графическая
 - б) устная и письменная
 - в) цифровая и алфавитная
 - г) вербальная и невербальная
3. В качестве единицы анализа (письменного) текста в работах используются, прежде всего, такие стандартные единицы, как:
 - а) лексема
 - б) словоформа
 - в) морфема
 - г) фонема
4. Любой текст на естественном языке характеризуется информационной, в противном случае он не может быть воспринят и понят адресатом.
5. В состав тезауруса входят лексемы, относящиеся к частям речи:
 - а) прилагательное
 - б) существительное
 - в) глагол
 - г) местоимение
6. В качестве компрессированных вариантов представления текста рассматривается:
 - а) лакунарный (лингвистический) текст
 - б) набор опорных слов
7. Напишите коэффициент лексической важности
8. Воспроизведение исходного текста в единицах переводящего языка, по возможности, с сохранением даже порядка следования элементов это:
 - а) дословный перевод
 - б) буквальный перевод
 - в) пословный перевод
9. Первый синтезатор «текст-речь» был создан в:
 - а) 1968 г.
 - б) 1939 г.
 - в) 1979 г.
10. Основным критерием оценки качества синтеза речи является ... синтезированной речи.

Утверждено на заседании кафедры славянской филологии и прикладной лингвистики, протокол № ____ от «__» _____ 20__ г.

Заведующий кафедрой
Преподаватель

Критерии оценивания модульного контроля

<i>Номер задания</i>	<i>Количество баллов</i>
1.	1
2.	1
3.	1
4.	1
5.	1
6.	1

7.	1
8.	1
9.	1
10.	1
<i>Всего</i>	<i>10</i>

10. ОБРАЗЕЦ ЭКЗАМЕНАЦИОННОГО БИЛЕТА

Теоретические вопросы к экзамену

1. Применение синтеза речи человек при получении информации.
2. Автоматический синтез речи. Основные понятия.
3. Основные принципы функционирования системы синтеза речи по тексту.
4. Основные способы синтеза речи. Краткая характеристика.
5. Параметрический синтез речи.
6. Общие сведения о форманте и ее роль в осуществлении синтеза речи.
7. Компиляционный синтез речи.
8. Полный синтез речи по правилам.
9. Предметно-ориентированный синтез речи.
10. Порядок синтеза текстовой информации.
11. Основная структура систем автоматического синтеза речи.
12. Направления использования систем распознавания речи.
13. Сферы деятельности с использованием системы распознавания речи
14. Аналого-цифровой преобразователь
15. Принцип работы АЦП
16. Биологический нейрон
17. Искусственный нейрон
18. Весовые коэффициенты и память нейрона
19. Модель нейрона
20. Сумматор в схеме искусственного нейрона
21. Функция активации
22. Структуры нейронных сетей
23. Однослойная нейронная сеть.
24. Многослойные нейронные сети.
25. Обучение нейронных сетей
26. Обучение нейронной сети с учителем
27. Обучение нейронной сети без учителя
28. Технология распознавания слитной речи на большом словаре
29. Группы помех и искажений
30. Примеры шумоочистки для разных типов искажений

ГОУ ВПО «ДОНЕЦКИЙ НАЦИОНАЛЬНЫЙ УНИВЕРСИТЕТ»

Кафедра славянской филологии и прикладной лингвистики

Направление подготовки: **45.03.03 – Фундаментальная и прикладная лингвистика**

Программа подготовки: **академический бакалавриат**

Семестр: **4**

Учебная дисциплина: **«Автоматическая обработка естественных языков»**

ЭКЗАМЕНАЦИОННЫЙ БИЛЕТ № 2

1. Применение синтеза речи человека при получении информации.
2. Основные характеристики моделей в системе распознавания слитной речи на большом словаре.

3. Выполнить компрессию произвольного текста в среде системы проверки правописания ОРФО. Аннотация.

Утверждено на заседании кафедры славянской филологии и прикладной лингвистики
 Протокол № _____ от « _____ » _____ 2020 г.

И.о.зав. кафедрой _____

доц. Кудрейко И. А.

Экзаменатор _____

ст. препод. Могила С.А.

Критерии оценивания экзамена

Номер задания	Количество баллов
1.	20
2.	20
3.	10
Всего	50 баллов

11. ОБРАЗЕЦ ТЕСТОВОГО ЗАДАНИЯ.

12. КРИТЕРИИ ОЦЕНИВАНИЯ

Текущий контроль (макс. 100 баллов)					Всего
Аудиторная работа студентов	Индивидуальная работа	Самостоят. работа	МКР	Зачет	
20	10	10	10	50	100

Шкала оценивания: национальная и ECTS

Сумма баллов по всем видам учебной деятельности	Оценка ECTS	Оценка по национальной шкале	
		для экзамена, курсового проекта (работы), практики	для зачета
90-100	A	отлично	зачтено
80-89	B	хорошо	
75-79	C		
70-74	D	удовлетворительно	
60-69	E		
35-59	FX	неудовлетворительно с возможностью повторной сдачи	не зачтено с возможностью повторной сдачи
0-34	F	неудовлетворительно с обязательным повторным изучением дисциплины	не зачтено с обязательным повторным изучением дисциплины

Критерии оценки знаний по ECTS (максимальная сумма баллов - 100)

90-100 баллов «А»

- выполнение обязательных и дополнительных заданий в полном объеме;
- активная и максимально продуктивная работа в аудитории и при подготовке к занятиям;
- фундаментальное владение знаниями по теории АОЕЯ;
- понимание и умение аргументировано пояснять выполняемые операции;
- квалифицированная работа со специальной и справочной литературой;
- осмысленное использование методов исследования;
- уверенное владение современным КТ;
- умение логично, непротиворечиво, последовательно разъяснять понятия, явления, делать выводы;
- отсутствие ошибок при выполнении лабораторных работ.

80-89 баллов «В»

- выполнение обязательных и дополнительных заданий в полном объеме при допущении несущественных неточностей;
- активная и продуктивная работа в аудитории и при подготовке к занятиям;
- качественное владение знаниями по теории;
- понимание и умение аргументировано пояснить использование КТ;
- осмысленное и успешное использование КТ при выполнении лабораторных работ;
- качественное владение современными КТ;
- умение дать без существенных ошибок определения, разъяснить понятия, делать непротиворечивые выводы;
- допущении 2-3 негрубых ошибок при выполнении лабораторных работ, хороший уровень речевой и языковой компетенции.

75-79 баллов «С»

- выполнение обязательных и некоторых дополнительных заданий в достаточном объеме при допущении нескольких неточностей;
- активная, но нерегулярная работа в аудитории и при подготовке к занятиям;
- хорошее владение знаниями по теории;
- умение достаточно аргументировано пояснить использование КТ;
- достаточный уровень умения работы с литературой;
- хорошее владение современными КТ;
- допущение 3-5 негрубых ошибок.

70-74 баллов «D»

- выполнение 70% обязательных заданий с допущением существенных погрешностей, которые могут быть устранены после их выявления;
- нерегулярная и не всегда продуктивная работа в аудитории, недостаточная подготовка к занятиям;
- знание содержания материала не в полном объеме и без требуемой глубины и осмысленности;
- умение в целом анализировать фактический материал, но бессистемно с допущением неточностей, неполноты или явных противоречий;
- поверхностное владение теоретическим материалом, отражающемся в неуверенном, не всегда качественном практическом применении;
- не вполне качественная работа с КТ;

60-69 баллов «E»

- выполнение 60% обязательных заданий при наличии некоторых ошибок и недостатков;
- пассивность на занятиях и нерегулярная подготовка к занятиям;
- пропуски занятий до 20%;
- наличие неотработанных тем;
- знание материала частичное или поверхностное;

- недостаточно выработанные умения по использованию КТ;
- недостаточное владение КТ с допущением некоторых грубых ошибок.

35-59 баллов «FX»

- выполнение обязательных заданий на 30-50%;
- пропуск занятий от 20% до 30%;
- пассивное посещение занятий и нерегулярна подготовка;
- наличие неотработанных тем и задолженностей по лабораторным работам (более 20%);
- неумение использовать КТ;
- отсутствие логики в рассуждениях;
- частичное знание и слабое владение КТ;
- слабо развитая речь.

0-34 баллов «F»

- выполнено менее 30% обязательных заданий;
- пропущено свыше 40% занятий;
- не подготовлена ни одна лабораторная работа;
- не усвоено более 50% теоретического материала;
- не выявлено знаний и умений в использовании КТ;
- полное невладение терминологическим аппаратом;

13. Материально-техническое обеспечение учебного процесса

Для проведения лекционных занятий и выполнения лабораторных работ требуется аудитория на группу, оборудованная КТ, мультимедийным проектором и экраном.

14. РЕКОМЕНДОВАННАЯ ЛИТЕРАТУРА

№ п/п	Наименование	Количество экз. в библиотеке ДонНУ	Наличие электрон. версии в ЭБС
<i>Основная литература</i>			
1.	Баранов А. Н. Введение в прикладную лингвистику: Учеб. пособие / А. Н. Баранов ; Моск. гос. ун-т им. М. В. Ломоносова, Филол. фак. - М.: Эдиториал УРСС, 2003. - 358 с	9	
2.	Кузнецов, С. Д. Базы данных: модели и языки: учеб. пособ. для студентов вузов, обучающихся по специальности " Прикладная математика и информатика" и "Информационные технологии" / С. Д. Кузнецов. - М.: Бинوم, 2008. - 720 с.	49	
3.	Леонтьева Н. Н. Автоматическое понимание текстов: системы, модели, ресурсы: [учеб. пособие для лингв. фак. вузов] / Н. Н. Леонтьева. - М.: Academia, 2006. - 303 с.	1	
4.	Введение в прикладную лингвистику: учебное пособие / Е. П. Соснина. – 2-е изд., испр. и доп. – Ульяновск : УлГТУ, 2012. – 110 с.		+
<i>Дополнительная литература</i>			
5.	Генерация тонального контура фразы в системах автоматического синтеза речи // Труды международного семинара Диалог'2000 по компьютерной лингвистике и приложениям. Протвино, 2000. [С. 211-219]. www.dialog-21.ru/digest/2000/articles/krivnova/	-	+
6.	Маслова, В. А. Современные направления в лингвистике: учеб. пособие для студентов вузов, обучающихся по направлению "Филол. образование" / В. А. Маслова. - Москва : Академия, 2008. - 264 с.	1	
7.	Лобанов Б.М., Цирульник Л.И. Компьютерный синтез и		+

	клонирование речи. Минск, 2008. pselab.ru/Books/Lobanov_Cirulnik_2008.pdf		
8.	Автоматическая обработка текстов на естественном языке и анализ данных : учеб. пособие / Большакова Е.И., Воронцов К.В., Ефремова Н.Э., Клышинский Э.С., Лукашевич Н.В., Сапин А.С. — М.: Изд-во НИУ ВШЭ, 2017. — 269 с		+

15. ИНФОРМАЦИОННЫЕ РЕСУРСЫ

1. Грамота.Ру: справочно-информационный портал «Русский язык»
<http://www.gramota.ru/>.
2. Культура письменной речи <http://www.gramma.ru/>.
3. Справочная служба русского языка <http://spravka.gramota.ru>.

16. ПРОГРАММНОЕ ОБЕСПЕЧЕНИЕ

Рабочая программа рассмотрена и переутверждена на заседании кафедры славянской филологии и прикладной лингвистики с изменениями (без изменений) на 2020 год.

Протокол № ____ от ____ 2020 года.

И.о. зав.кафедрой _____ И.А. Кудрейко