

# МИНОБРНАУКИ РОССИИ

Федеральное государственное бюджетное образовательное учреждение  
высшего образования  
«Нижегородский государственный лингвистический университет имени  
Н.А. Добролюбова»



УТВЕРЖДАЮ

Проректор по учебной  
деятельности

Р.И. Иванова

29 августа 2023 г.

Компьютерная лингвистика. Основы программирования на языке Python и обработки устной (онлайн-курс)  
(наименование дисциплины)

## дополнительная профессиональная программа повышения квалификации

Форма обучения                      **заочная**

Общая трудоемкость                **2 ЗЕТ**

Часов по учебному плану            72  
*в том числе:*  
аудиторные занятия                18  
самостоятельная работа            53,95  
часов на контроль                    0,05

Виды контроля:

зачет

### Распределение часов дисциплины по семестрам

Вид занятий	Кол-во часов
Лекции	<b>18</b>
Практические (в том числе интеракт.)	
Семинарские (в том числе интеракт.)	
<b>Итого ауд.</b>	<b>18</b>
Часы на контроль	<b>0,05</b>
Контактная работа	<b>18</b>
Самостоятельная работа	<b>53,95</b>
<b>Итого</b>	<b>72</b>

Программу составил(и):  
*уч. степень, уч. звание, Ф.И.О.* канд. филол. наук, доц. М.Б. Чиков

Рабочая программа  
«Компьютерная лингвистика. Основы программирования на языке Python и обработки устной речи» (онлайн-курс)

разработана в соответствии с ФГОС:

Федеральный государственный образовательный стандарт высшего образования по направлению подготовки 45.03.02 Лингвистика, утвержденный приказом МИНОБРНАУКИ РОССИИ от 12 августа 2020 г., № 969.

Рабочая программа одобрена на заседании кафедры  
романо-германских языков, перевода, зарубежной литературы и межкультурной коммуникации

Протокол от 29 августа 2023 г. № 1

Зав. кафедрой канд. филол. наук, доц. Е.А. Сакулина  
*(уч. степень, уч. звание, Ф.И.О)*

### 1. ЦЕЛИ И ЗАДАЧИ ОСВОЕНИЯ КУРСА

1.1	Цель освоения дисциплины: Сформировать сквозные цифровые компетенции у студентов гуманитарных специальностей, направлений и профилей подготовки
1.2	Задачи освоения дисциплины:
1.2.1	Дать представление о базовых инструментах цифровой обработки естественного языка
1.2.2	Дать представление о программных средствах обработки естественного языка
1.2.3	Дать представление о научно-исследовательских задачах, решаемых при помощи цифровых инструментов обработки естественного языка

### 2. КОМПЕТЕНЦИИ ОБУЧАЮЩЕГОСЯ, ФОРМИРУЕМЫЕ В РЕЗУЛЬТАТЕ ОСВОЕНИЯ КУРСА

*ОПК-5. Способен работать с компьютером как средством получения, обработки и управления информацией для решения профессиональных задач.*

#### Знать:

Уровень Пороговый	На пороговом уровне: принципы междисциплинарного взаимодействия цифровых технологий и лингвистики, принципы функционирования компьютерных программ и сетевых ресурсов для обработки естественного языка, возможности применения цифровых технологий для решения учебных и научно-исследовательских лингвистических задач
Уровень Высокий	На высоком уровне: принципы междисциплинарного взаимодействия цифровых технологий и лингвистики, принципы функционирования компьютерных программ и сетевых ресурсов для обработки естественного языка, возможности применения цифровых технологий для решения учебных и научно-исследовательских лингвистических задач
Уровень Повышенный	На повышенном уровне: принципы междисциплинарного взаимодействия цифровых технологий и лингвистики, принципы функционирования компьютерных программ и сетевых ресурсов для обработки естественного языка, возможности применения цифровых технологий для решения учебных и научно-исследовательских лингвистических задач

#### Уметь:

Уровень Пороговый	На пороговом уровне: ставить научно-исследовательские и учебные лингвистические задачи, решаемые с использованием цифровых технологий, и применять цифровые технологии для решения тех или иных поставленных задач
Уровень Высокий	На высоком уровне: ставить научно-исследовательские и учебные лингвистические задачи, решаемые с использованием цифровых технологий, и применять цифровые технологии для решения тех или иных поставленных задач
Уровень Повышенный	На повышенном уровне: ставить научно-исследовательские и учебные лингвистические задачи, решаемые с использованием цифровых технологий, и применять цифровые технологии для решения тех или иных поставленных задач

#### Владеть:

Уровень Пороговый	На пороговом уровне: основными компьютерными программами и сетевыми ресурсами для обработки естественного языка, их функциональным инструментарием
Уровень Высокий	На высоком уровне: основными компьютерными программами и сетевыми ресурсами для обработки естественного языка, их функциональным инструментарием
Уровень Повышенный	На повышенном уровне: основными компьютерными программами и сетевыми ресурсами для обработки естественного языка, их функциональным инструментарием

### 3. СТРУКТУРА И СОДЕРЖАНИЕ КУРСА

Код занятия	Наименование разделов и тем /вид занятия/	Семестр / Курс	Объем в часах	Компетенции
	<b>Раздел 1. Язык программирования Python</b>	1/1	14	ОПК-5
4.1	<i>Возможности языка Python для обработки текста. /Лек/</i>	1/1	2	ОПК-5
4.2	<i>Изучение основных функций обработки текста с помощью языка Python. /Лек/</i>	1/1	2	ОПК-5
4.3	Самостоятельная работа	1/1	12	ОПК-5
	<b>Раздел 2. Анализ и обработка устной речи</b>	1/1	14	ОПК-5
5.1	<i>Из истории автоматизированной обработки устной речи. /Лек/</i>	1/1	2	ОПК-5
5.2	<i>Анализ и обработка устной речи. /Лек/</i>	1/1	2	ОПК-5
5.3	Самостоятельная работа	1/1	10	ОПК-5

### 4. ФОНД ОЦЕНОЧНЫХ СРЕДСТВ

#### 5.1. Контрольные вопросы и задания

Раздел 1.

1. В чем причина большой произносительной вариативности?

- a) В дефектах органов артикуляции
  - b) Вариативность зависит от канала связи
  - c) Произносительная вариативность является лингвистически несущественной
  - d) В характере ситуации общения
2. Кто является создателем теории акустических дистинктивных признаков фонем?
- a) Теорию разработал Фердинанд де Соссюр
  - b) Вольфганг фон Кемпелен
  - c) Р. Якобсон, Г. Фант, М. Халле
  - d) Разработчики устройства VODER
3. Для каких целей создана программа Praat?
- a) Для анализа спектральных характеристик звуков
  - b) Для анализа спектральных и просодических характеристик звуков
  - c) Для анализа просодических характеристик звуков
  - d) Для анализа просодических и невербальных характеристик речи
4. Какова максимальная продолжительность отрезка речи для анализа?
- a) 5 сек
  - b) 10 сек
  - c) 10 мсек
  - d) Ограничений нет
5. Что такое Pitch?
- a) Резкое изменение частоты основного тона
  - b) Усиление частоты тона голоса при ударе
  - c) Самая высокая частота тона голоса
  - d) Частота основного тона голоса
6. Что такое форманта?
- a) Акустическая характеристика мелодии голоса
  - b) Самое сильное усиление энергии в спектре звука
  - c) Область усиления энергии в спектре звука
  - d) Характеристика интенсивности звука
7. Что такое Intensity?
- a) Характеристика громкости звуков речи
  - b) Напряжение органов артикуляции
  - c) Напряжение голосовых связок
  - d) Громкость пения
8. Какие спектральные характеристики можно получить с помощью программы Praat?
- a) Частоты формант, темп, ритм, интенсивность
  - b) Частоты формант, частота основного тона
  - c) Частоты формант, интенсивность, частота основного тона, длительность
  - d) Частоты формант, интенсивность, частота основного тона
9. Чем отличается осциллограмма от спектрограммы?
- a) Осциллограмма показывает только длительность звука
  - b) Осциллограмма показывает формантную структуру звука

- c) Осциллограмма показывает те же характеристики, что и спектрограмма
- d) Осциллограмма показывает частоту и амплитуду колебаний на отрезке времени

10. Что такое суперсегментный уровень?

- a) Суперсегментный уровень – это просодия эмоциональной речи
- b) Суперсегментный уровень – это характеристики эмоциональной речи
- c) Суперсегментный уровень – это характеристики звука в зависимости от соседнего сегмента (звука)
- d) Суперсегментный уровень – это средства интонации

Задания для самостоятельной работы

Загрузите программу Praat с сайта программы, выберите команду New > Record mono sound, сделайте запись предложения «Фонетика будущего» и откройте запись в окне анализа: Record > Stop > Save to list > View&Edit. Определяем длительность звучания (t) и вычлняем первое слово для анализа: маркируем курсором, слушаем, нажав на клавишу под выделением, уточняем границы выделения и разворачиваем выделенный фрагмент на все окно анализа > sel.

Определяем значения просодических параметров

1) *Движение основного тона*. Нажимаем кнопку Pitch над окном анализа. Ставим курсор на максимумы огибающей основного тона (синяя линия), соответствующие слогам и записываем в таблицу данные в Гц. Затем можно построить график.

2) *Огибающая интенсивности*. Нажимаем кнопку Intensity и записываем данные по интенсивности в дБ.

Ответьте на вопрос, есть ли связь между высотой тона и интенсивностью?

Раздел 2.

Контрольные вопросы

1. Что такое невербальная коммуникация?

- a) Средства общения для невидящих или слабовидящих
- b) Жестовая коммуникация для неслышащих или слабослышащих
- c) Использование мимики, жестов для передачи смыслов вместо слов
- d) Средства общения на незнакомом языке

2. Что такое поликодовая коммуникация?

- a) Кодирование информации в каналах телефонной и радиосвязи
- b) Общение с использованием вербальных и невербальных кодов
- c) Электронное кодирование информации
- d) Общение неслышащих людей

3. Чем программа ELAN отличается от Praat?

- a) Программа ELAN в отличие от Praat может использоваться для компьютерного анализа устной речи
- b) Использование программы ELAN ограничено английским языком
- c) Программа ELAN позволяет анализировать не только устную, но и письменную речь
- d) Программа ELAN в отличие от Praat может использоваться для паралингвистических исследований

4. Можно с помощью программы ELAN изучать акустику речи?

- a) Да
- b) Нет

- c) Да, если перегрузить аудиозапись в анализатор акустики
- d) Да, если не анализировать видеоизображение

5. В каких направлениях науки можно использовать программу Praat?

- a) Только в компьютерной лингвистике
- b) При изучении поведения животных
- c) При изучении эмоций в психологии
- d) Во всех названных направлениях

Задания для самостоятельной работы

Загрузите программу ELAN 6.2 с сайта программы. Будьте внимательны: версии для разных операционных систем могут различаться! Загрузите из интернета небольшой видеоролик и сохраните его на компьютере в формате, например, Мр4. Откройте программу, выберите «Файл» > Новый > Откроется окно «Новый документ» > выберите «Добавить медиафайл» > найдите ваш видеоролик > Открыть > Ваш видеоролик в списке выбранных файлов > ОК > Открывается окно, в нем видео, звуковая дорожка и строки для аннотирования. Потренируйтесь выбирать различные фрагменты курсором, записывать на имеющемся слое аннотацию (текст) фрагмента, добавлять новые слои (команда «Слой» сверху окна программы). После этого можно ознакомиться со всеми возможностями программы в руководстве, которое можно найти в интернете, и работать с ней! Завершая работу, можно сохранить результат: Файл > Сохранить как (выбираете путь сохранения и название); можно сохранить только выделенный фрагмент, который вы анализировали > в формате \*.eaf (специальный формат программы).

#### 4.2. Фонд оценочных средств

Фонд оценочных средств представлен в Приложении 1

#### 4.3. Перечень видов оценочных средств

*Контрольные вопросы, задания для самостоятельной работы, автоматизированные тесты*

### 5. УЧЕБНО-МЕТОДИЧЕСКОЕ И ИНФОРМАЦИОННОЕ ОБЕСПЕЧЕНИЕ ДИСЦИПЛИНЫ (МОДУЛЯ)

#### 5.1. Рекомендуемая литература

##### 5.1.1. Основная литература

	Авторы, составители	Заглавие	Издательство, год
Л1.1	<i>Бессмертный, И. А.</i>	Интеллектуальные системы : учебник и практикум для вузов / И. А. Бессмертный, А. Б. Нугуманова, А. В. Платонов. (Высшее образование). — ISBN 978-5-534-01042-8. — Текст : электронный // Образовательная платформа Юрайт [сайт]. — URL: <a href="https://urait.ru/bcode/511999">https://urait.ru/bcode/511999</a>	Москва : Издательство Юрайт, 2023. — 243 с.
Л1.2	<i>Чернышев, С. А.</i>	Основы программирования на Python : учебное пособие для вузов / С. А. Чернышев. — 2-е изд., перераб. и доп. (Высшее образование). — ISBN 978-5-534-17139-6. — Текст : электронный // Образовательная платформа Юрайт [сайт]. — URL: <a href="https://urait.ru/bcode/532446">https://urait.ru/bcode/532446</a>	Москва : Издательство Юрайт, 2023. — 349 с.
Л1.3	<i>Шунейко, А. А.</i>	Корпусная лингвистика : учебник для вузов / А. А. Шунейко. (Высшее образование). — ISBN 978-5-534-13603-6. — Текст : электронный // Образовательная платформа Юрайт [сайт]. — URL: <a href="https://urait.ru/bcode/519477">https://urait.ru/bcode/519477</a>	Москва : Издательство Юрайт, 2023. — 222 с.

##### 5.1.2. Дополнительная литература

*(литература из электронных ресурсов и печатные варианты из библиотеки НГЛУ)*

	Авторы, составители	Заглавие	Издательство, год
Л2.1	<i>Шунейко, А. А.</i>	Квантитативная лингвистика и новые информационные технологии : учебник для вузов / А. А. Шунейко, И. А. Авдеенко. (Высшее образование). — ISBN 978-5-534-15446-7. — Текст : электронный // Образовательная платформа Юрайт [сайт]. — URL: <a href="https://urait.ru/bcode/519734">https://urait.ru/bcode/519734</a>	Москва : Издательство Юрайт, 2023. — 347 с.

#### 5.2. Перечень ресурсов информационно-телекоммуникационной сети "Интернет"

Э1	<a href="https://www.memsource.com/resources/">https://www.memsource.com/resources/</a>
Э2	<a href="https://studiorum.ruscorpora.ru/">https://studiorum.ruscorpora.ru/</a>
Э3	<a href="https://archive.mpi.nl/tla/elan">https://archive.mpi.nl/tla/elan</a>
Э4	<a href="https://www.fon.hum.uva.nl/praat/">https://www.fon.hum.uva.nl/praat/</a>

Э5	<a href="http://www.python.org">www.python.org</a>
<b>5.3. Перечень программного обеспечения</b>	
6.3.1.	Microsoft Windows 7, 10
6.3.2.	Microsoft Office - Word - Excel - Access - Power Point - Outlook
6.3.3.	Adobe Acrobat Reader DC
6.3.4.	ABBYY FineReader 11
6.3.5.	Microsoft Edge
6.3.6.	Mozilla Firefox
6.3.7.	Google Chrome
6.3.8.	CorelDraw
6.3.9.	Adobe InDesign cs 6
6.3.10.	Adobe Photoshop
6.3.11.	Open Broadcaster Software (OBS Studio)
6.3.12.	Контур.Толк
6.3.13.	Система «Антиплагиат»
6.3.14.	1С:Бухгалтерия 8. Учебная версия
6.3.15.	Антивирус Касперского
<b>5.4. Перечень информационных справочных систем</b>	
6.4.1	Справочная правовая система «Консультант Плюс»
6.4.2.	Справочная система «Гарант.Ру»

<b>6. МАТЕРИАЛЬНО-ТЕХНИЧЕСКОЕ ОБЕСПЕЧЕНИЕ КУРСА</b>	
7.1	<i>Помещения для самостоятельной работы, оснащенные компьютерной техникой с возможностью подключения к сети «Интернет» и обеспечением доступа в электронную информационно-образовательную среду вуза.</i>

<b>7. МЕТОДИЧЕСКИЕ УКАЗАНИЯ ДЛЯ ОБУЧАЮЩИХСЯ ПО ОСВОЕНИЮ КУРСА</b>	
<p>В дисциплине «Компьютерная лингвистика. Основы программирования на языке Python и обработки устной речи» (онлайн-курс) занятия требуют работы по освоению материала видеолекций и внеаудиторной работы, а именно:</p> <ul style="list-style-type: none"> <li>- прослушивания лекций;</li> <li>- самостоятельного решения практических задач;</li> <li>- осуществления самоконтроля знаний;</li> <li>- изучения рекомендуемой дополнительной литературы;</li> <li>- взаимодействия с куратором (проктором) по решению технических и содержательных вопросов,</li> <li>- выполнения автоматизированных тестов.</li> </ul>	