

МИНИСТЕРСТВО НАУКИ И ВЫСШЕГО ОБРАЗОВАНИЯ РОССИЙСКОЙ ФЕДЕРАЦИИ

ФЕДЕРАЛЬНОЕ ГОСУДАРСТВЕННОЕ БЮДЖЕТНОЕ ОБРАЗОВАТЕЛЬНОЕ  
УЧРЕЖДЕНИЕ ВЫСШЕГО ОБРАЗОВАНИЯ

**СЕВЕРО-КАВКАЗСКАЯ ГОСУДАРСТВЕННАЯ АКАДЕМИЯ**

А. Б. Алиева  
Р. С. Кочкаров

## **ИНФОРМАЦИОННЫЕ ТЕХНОЛОГИИ В ЛИНГВИСТИКЕ**

Лабораторный практикум для обучающихся  
по направлению подготовки 45.03.02 Лингвистика

Черкесск  
2024

УДК 004:81  
ББК 16.2:81  
А 50

Рассмотрено на заседании кафедры «Общая информатика»  
Протокол № 1 от «06» сентября 2023 г.  
Рекомендовано к изданию редакционно-издательским советом СКГА.  
Протокол № 26 от «29» сентября 2023 г.

**Рецензенты:** Узденова Б. Х. – к. п. н., доцент кафедры «Общая информатика»

А 50            **Алиева, А.Б.** Информационные технологии в лингвистике. лабораторный практикум для обучающихся по направлению подготовки 45.03.02 Лингвистика /А. Б. Алиева, Кочкаров Р. С. – Черкесск: БИЦ СКГА, 2024. – 64 с.

Учебное-методическое пособие содержит теоретический материал и лабораторный практикум для изучения дисциплины «Информационные технологии в лингвистике». В данном пособии представлены лабораторные работы, которые позволяют сформировать практические навыки работы с компьютером и набором программных продуктов, предназначенными для автоматизации деятельности лингвистов и переводчиков. Лабораторные работы могут использоваться для освоения на практических занятиях, а также в процессе самостоятельной работы студентов.

Одобрено к изданию кафедрой «Общая информатика».

**УДК 004:81**  
**ББК 16.2:81**

© Алиева А. Б., Кочкаров Р. С. 2024  
© ФГБОУ ВО СКГА, 2024

## ВВЕДЕНИЕ

Учебно-методическое пособие «Информационные технологии в лингвистике» подготовлено в соответствии с требованиями ФГОС ВО по направлению подготовки 45.03.02 «Лингвистика» и программе курса «Информационные технологии в лингвистике».

Целью пособия является помощь в освоении дисциплины «Информационные технологии в лингвистике» и содействии формированию и развитию у студентов общекультурных и специальных компетенций, позволяющих им в дальнейшем осуществлять профессиональную деятельность посредством использования информационных технологий в лингвистических исследованиях и своей будущей профессиональной деятельности.

### Планируемые результаты обучения

Планируемые результаты освоения образовательной программы (ОП) – компетенции обучающихся определяются требованиями стандарта по направлению подготовки 45.03.02 Лингвистика и формируются в соответствии с матрицей компетенций ОП.

№ п/п	Номер/индекс компетенции	Наименование компетенции (или ее части)	В результате изучения дисциплины, обучающиеся должны:
1	2	3	4
1	ОПК-5	Способен работать с компьютером как средством получения, обработки и управления информацией для решения профессиональных задач	ОПК-5.1 осуществляет поиск и обработку необходимой информации, содержащейся в специальной литературе ОПК-5.2 Использует энциклопедические, толковые, исторические, этимологические словари, словари сочетаемости, включая профильные электронные ресурсы для решения профессиональных задач ОПК-5.3 соблюдает правила составления и оформления ссылок и библиографии, принятые в научном дискурсе

В пособии представлены девять лабораторных работ, отражающих возможности современных информационных технологий в области лингвистики. Каждая лабораторная работа содержит краткие теоретические сведения, описывает ход работы для достижения поставленных целей. В конце каждой работы имеется список контрольных вопросов и задания для самостоятельного выполнения, которые можно использовать в качестве текущего контроля. Для проверки уровня сформированности компетенций в пособии приведены примерные задания для промежуточного контроля.

## Лабораторная работа №1

### ТЕМА: ИНФОРМАЦИОННЫЕ ТЕХНОЛОГИИ. СТРУКТУРА ИНФОРМАЦИОННЫХ ТЕХНОЛОГИЙ.

Лингвистика как наука о закономерностях строения и развития естественного языка. Понятие теоретической и прикладной лингвистики. Соотношение прикладной и компьютерной лингвистики.

Язык как знаковая система. Понятие естественного и искусственного языка. Виды искусственных языков.

Изучение возможностей применения информационных технологий в лингвистике предполагает знание основных понятий соответствующей области знания, среди которых можно выделить понятия из сферы лингвистики (язык, лингвистика, компьютерная лингвистика и т.п.) и информатики (информация, алгоритм, модель и др.). Знакомство с этими понятиями начнем с лингвистических терминов, характеризующих непосредственную профессиональную область деятельности лингвистов, преподавателей иностранных языков и переводчиков.

Лингвистика (или языкознание) традиционно понимается как наука о естественном человеческом языке [9, 28]. Лингвистов занимают вопросы строения языка (выделение в нем фонетического, лексического, грамматического уровня и уровня текста), социального варьирования языка, вопросы порождения и понимания языковых высказываний, принципы функционирования языка в обществах разных типов, происхождения и развития языка и другие его аспекты [13, 618-622].

В зависимости от изучаемого аспекта языка, национальной традиции и научной методологии выделяются различные разделы лингвистики, например структурная лингвистика, социолингвистика, психолингвистика и т.п. Чтобы определить раздел лингвистики, наиболее тесно связанный с использованием информационных технологий, целесообразно обратиться к разграничению теоретической и прикладной лингвистики.

Теоретическая (или фундаментальная) лингвистика – это область языкознания, направленная на объективное установление состояния отдельного языка, его истории и закономерностей. Эта область лингвистики призвана ответить на вопрос «Каков язык?» [37,214-215].

Прикладная лингвистика развивается с конца 20-х годов XX в. и является областью языкознания, связанной с разработкой методов решения практических задач использования языка [13, 397]. Прикладная лингвистика отвечает на вопрос «Как лучше использовать язык?». Следует отметить, что в России и за рубежом сложились разные интерпретации понятия прикладной лингвистики. Если за рубежом в 1930-1940-е годы под прикладной лингвистикой (AppliedLinguistics) прежде всего, понимается процесс обучения иностранному языку, методика его преподавания, особенности описания грамматики для учебных целей, то в России начиная с 1950-х

годов, прикладная лингвистика ассоциируется с компьютерными технологиями и автоматическими системами обработки информации [4, 6].

В связи с этим в русскоязычной научной традиции прикладная лингвистика нередко рассматривается как синоним компьютерной / вычислительной / автоматической / инженерной лингвистики.

На современном этапе развития науки в рамках прикладной лингвистики выделяется несколько направлений по оптимизации использования языка, которые объединяются исследователями в две большие группы: традиционные («вечные») и новые.

К традиционным направлениям и соответствующим задачам прикладной лингвистики относятся:

- создание и совершенствование письменностей;
  - создание систем транскрипции устной речи;
  - создание систем транслитерации иноязычных слов;
  - создание систем стенографии;
  - создание систем письма для слепых;
  - упорядочение, унификация и стандартизация научно-технической терминологии;
  - изучение процессов и создание правил образования названий новых изделий, товаров, химических веществ;
  - разработка методов адекватного преобразования текстов в иноязычную форму (перевода);
  - совершенствование методики преподавания языков и др.
- Новыми задачами прикладной лингвистики считаются:
- разработка лингвистических основ машинного перевода;
  - автоматическое индексирование и аннотирование документов;
  - автоматический анализ текстов;
  - автоматический синтез текстов;
  - создание словарей-тезаурусов для автоматического поиска информации и др.

Именно те области прикладной лингвистики, которые связаны с привлечением компьютеров для решения практических задач использования языка, являются предметом компьютерной лингвистики, оформившейся в 1960-е годы как особое научное направление.

Компьютерную лингвистику можно определить как область использования компьютерных инструментов – программ, технологий организации и обработки данных – для моделирования функционирования языка в тех или иных условиях, а также сферу применения компьютерных моделей языка в лингвистике и смежных с ней дисциплинах.

В связи с тем, что язык представляет собой весьма сложное образование, в компьютерной лингвистике сложились и развиваются различные направления, примерно сопоставимые с отдельными уровнями языка, с процессами порождения и восприятия языковых сообщений или другими видами человеческой деятельности, связанной с языком.

Соответственно, к направлениям компьютерной лингвистики относятся:

- автоматический анализ текстов;
- автоматический синтез текстов;
- создание и поддержка автоматических словарей;
- создание автоматизированных информационно-поисковых систем;
- машинный перевод;
- создание автоматических систем обучения языку;
- автоматическая атрибуция и дешифровка анонимных текстов;
- создание лингвистических баз данных;
- разработка программных инструментов для решения задач теоретической и прикладной лингвистики и т.д.

Лингвистика в целом и компьютерная лингвистика в частности имеют дело с языками различного типа и их отдельными уровнями. Язык в наиболее общем виде определяется как знаковая система, используемая для общения в некотором социуме.

Различают естественные и искусственные языки. Естественный язык – это исторически сложившаяся и используемая в определенной этнической группе или национальном государстве знаковая система. Примерами естественных языков выступают русский и английский (принадлежащие к индоевропейской языковой семье) или финский и эстонский (принадлежащие к финно-угорской языковой семье). Искусственные языки представляют собой знаковые системы, искусственно создаваемые в тех областях, где применение естественных языков менее эффективно или невозможно.

Среди искусственных выделяются неспециализированные (или международные) языки (эсперанто, волапюк и др.) и специализированные языки. К последним относятся языки науки (математики, логики, химии и т.д., создание которых началось в XV в.) и языки человеко-машинного общения (получающие распространение в специальных областях человеческой деятельности, связанной с облегчением диалога человека и компьютера, начиная с 1940-х годов).

Примеры языков человеко-машинного общения простираются от простейших систем символического кодирования (ассемблеров) до специализированных языков программирования (C++, Java, Python, Erlang и др.). К 1980-м годам в мире насчитывалось около 500 языков программирования.

В настоящее время активно используется примерно столько же, хотя общее количество известных языков программирования достигает нескольких тысяч. Эти факты свидетельствуют об остроте проблемы человеко-машинного общения и о множестве подходов к ее решению. Подводя итог разделу, констатируем, что лингвистикой следует считать науку о закономерностях происхождения, строения и функционирования естественного человеческого языка. Предметом лингвистики и компьютерной лингвистики как ее особого раздела выступает язык – знаковая система, используемая с различными целями.

В целом, во всей своей совокупности целей, направлений, методов и подходов, прикладная лингвистика – это прежде всего комплексная дисциплина, постоянно соприкасающаяся с философией, психологией, физиологией, математикой, логикой, социологией, информатикой.

Прилагаемый ниже список основных терминов (понятий) математической и прикладной лингвистики дает определенное представление о предмете, направлениях, методах этой науки сегодня.

**Алгоритм** – последовательность действий, выполнение которых необходимо для достижения конкретной цели.

**База данных** – совокупность данных, представленных в формализованном виде, удобном для автоматической передачи, интерпретации или обработки.

**База знаний** – база данных, элементами которых являются формализованные представления знаний об элементах внеязыковой действительности.

**Байт** – в вычислительной технике- обрабатываемый как единое целое элемент данных, представляющий собой последовательность двоичных разрядов (обычно-8). Используется и как единица измерения памяти. 1 килобайт= 1024 байта. 1 мегабайт=1024 килобайта. При обработке и хранении текстов обычно одному байту соответствует один символ текста.

**Грамматика зависимостей** – формальное представление строя предложения в виде иерархии компонентов, между которыми установлено отношение зависимости.

**Грамматика непосредственно составляющих** – формальное представление строя предложения в виде иерархии вложенных друг в друга линейно непересекающихся элементов, максимально независимых друг от друга.

**Грамматика представлений** – система правил, основанная на учете возможных окружений каждого слова.

**Граф** – математический объект, состоящий из множества вершин (точек) и множества ребер (связей), соединяющих пары вершин.

**Дерево предложения, дерево зависимостей** – представление структуры предложения в виде ориентированного графа, узлы которого соответствуют компонентам структуры зависимостей.

**Интерфейс** – средства и способы обмена информации. Обычно имеют в виду человеко-машинный интерфейс, т. е. средства и способы организации диалога человека и ЭВМ, хотя можно говорить и об интерфейсе между отдельными блоками вычислительной машины, программами, частями программ и т.д. Одной из лингвистических задач искусственного интеллекта является создание естественно-языкового интерфейса, т. е. организация диалога человека с ЭВМ на естественном языке.

**Квазиреферат** – реферат, составленный компьютером.

**Конфигурационный анализ** – анализ входного текста путем сравнения с набором предварительно выделенных синтаксических

конструкций (конфигураций, формул) переводимого языка. При совпадении установленной конфигурации входного текста с одной из конфигураций набора текстовая конфигурация считается распознанной, «свертывается» и в дальнейшем анализе фигурирует в свернутом виде. При синтезе полученная из анализа конфигурация сравнивается с соответствующей ей конфигурацией переводящего языка. Соответствие конфигураций устанавливается предварительно.

**Лингвистический процессор** – набор процедур для обработки текстов на естественном языке, которые являются входной информацией автоматизированных систем.

**Накопитель** – устройство в ЭВМ, предназначенное для считывания и записи информации на внешних носителях. Например, накопитель на магнитных дисках (дисковод), накопитель на магнитных лентах (лентопротяжное устройство) и т.п.

**Оконный интерфейс** – один из современных способов организации интерфейса (см.). При этом способе на экране дисплея выделяется прямоугольная область (окно), в которую человек вводит и/или ЭВМ выводит информацию. При необходимости организации диалога на различные темы может быть организовано несколько окон (многооконный интерфейс).

**Предсказуемый анализ** – процедура, основанная на просмотре текста слева направо, при этом для каждого очередного слова проверяется, является ли оно по тем или иным параметрам предсказанным предыдущими.

**Псевдограф** – ориентированный граф, имеющий и петли, и дуги.

**Семантическое представление данных** – формальная запись смысловой интерпретации информации, которая использует какую-либо модель семантического описания.

**Словарь-конкорданс** – специальный словарь текстовых словоформ с указанием всех контекстов для каждой словоформы.

**Тезаурус** – идеографический словарь, в котором показаны семантические отношения между его единицами. Структурной основой тезауруса является иерархическая система понятий конкретной предметной области.

**Файл** – совокупность связанных записей в памяти ЭВМ, рассматриваемая как единое целое.

**Фрейм** – способ описания структуры текста или ситуации, состоящий в перечислении возможных компонентов и типов связей между ними.

**Язык представления данных** – формальный способ структурирования записи информации, в основе которого лежит какая-либо формализованная (математическая) модель описания.

**Компьютерная сеть** – система связи компьютеров и компьютерного оборудования.

В зависимости от территориального расположения абонентских систем вычислительные сети можно разделить на три основных класса:



**Глобальная вычислительная сеть (WAN – Wide Area Network)** – это система объединенных в сеть компьютеров, расположенных на больших расстояниях друг от друга.

**Региональная вычислительная сеть (MAN – Metropolitan Area Network)** связывает абонентов, расположенных на значительном расстоянии друг от друга (десятки – сотни километров).

**Локальная вычислительная сеть (ЛВС, LAN – Local Area Network)** объединяет абонентов, расположенных в пределах небольшой территории (2 – 2,5 км.).

#### **Аппаратное обеспечение компьютерной сети**

Аппаратное обеспечение локальной вычислительной сети включает **рабочие станции, сервер, коммуникационное оборудование**

**Рабочая станция** – персональный компьютер, подключенный к сети и обеспечивающий доступ пользователя к ее ресурсам.

**Сервер** – специализированный компьютер в сети, предназначенный для выполнения на нем сервисного программного обеспечения.

#### **Рассмотрим каналы связи.**

**Витая пара** характеризуется плохой помехозащищенностью и низкой скоростью передачи информации 0,25 – 1 Мбит/с. Низкая стоимость.

**Коаксиальный кабель** обладает высокой помехозащищенностью и обеспечивает скорость передачи информации до 10-50 Мбит/с.

**Оптоволоконный кабель** не подвержен действию электромагнитных полей и сам практически не имеет излучения. Скорость передачи информации более 50 Мбит/с.

**Беспроводные сети.** Технологии беспроводной связи в компьютерных сетях основаны на кодировании и передаче сигналов в виде электромагнитного излучения.

#### **Коммуникационное оборудование:**

**Сетевой концентратор или хаб** – сетевое устройство, предназначенное для объединения нескольких устройств Ethernet1 общий сегмент сети. Устройства подключаются при помощи витой пары, коаксиального кабеля или оптоволоконка.

В настоящее время **хабы** почти не выпускаются – им на смену пришли **сетевой коммутатор (свитчи)** – устройства, предназначенные для соединения нескольких узлов сети в пределах одного или нескольких сегментов сети.

**Роутер (маршрутизатор)** – устройство, обеспечивающее соединение административно независимых коммуникационных сетей.

**Мост** – это устройство, соединяющее две сети, использующие одинаковые методы передачи данных

**Шлюз** – это устройство, которое позволяет организовать обмен данными между двумя сетями, использующими различные протоколы взаимодействия.

**Internet** – это глобальная сеть, объединяющая многие локальные, региональные сети и включающая сотни миллионов ПК.

Основными составляющими любой глобальной сети являются компьютерные узлы и каналы связи. К узлам компьютерной сети подключаются ПК пользователей и организации через файл-сервер локальной сети.

**Поисковая служба заключается в услугах поисковых серверов.** Существует две разновидности поисковых серверов: *поисковые каталоги* и *поисковые указатели* (или *поисковые машины*).

– **поисковые каталоги** – поиск информации сводится к выбору определенного каталога, после чего пользователю будет предоставлен список ссылок на URL-адреса наиболее посещаемых Web-сайтов и Web-страниц.

– **поисковые указатели или поисковые машины** (www.yandex.ru, www.rambler.ru, www.google.com.ru, www.yahoo.com и др.) находят Web-сайты или Web-страницы по ключевым словам в базе данных.

Для того чтобы сделать поиск более продуктивным, во всех поисковых системах существует специальный *язык формирования запросов* со своим синтаксисом.

### **Вопросы для обсуждения на семинарских занятиях**

1. Информационные технологии. Структура информационных технологий.

2. Что такое лингвистика? Назовите ее разделы. В каком разделе лингвистика имеет дело с информационными технологиями?

3. Можно ли считать синонимами прикладную и компьютерную лингвистику? Аргументируйте свой ответ.

4. Перечислите основные направления компьютерной лингвистики. Расскажите об одном из направлений.

5. Сравните разные определения языка. Выделите в них ключевые слова. Составьте на основе повторяющихся ключевых слов свое определение языка.

6. Подумайте, с естественным или искусственным языком имеет дело компьютерная лингвистика?

7. Какие виды естественных и искусственных языков вам известны? Приведите примеры естественных и искусственных языков разных видов.

### **Упражнения**

1. Определите статистические показатели приведенного ниже текста смешанного языкового типа.

<p>Проекты Cibola/Oleada реализуют обширные компьютерные системы лингвистического анализа текстов, представленных в Unicode. Компоненты системы включают средства работы с мультязыковыми текстами (MUTT), построения конкорданса (XConcord) для текстов на более чем 16 языках, статистического анализа, автоматического перевода, различные словари и тезаурусы. Некоторые версии этих компонентов доступны для бесплатной загрузки после процедуры формальной регистрации. Все компоненты реализованы в среде XII Window System для SunOs и Solaris (источник: <i>Проекты Cibola/Oleada</i> <a href="http://rvb.ru/soft/catalogue/c01.html">http://rvb.ru/soft/catalogue/c01.html</a>).</p>
--

Слов	
Символов (без пробелов)	
Символов (с пробелами)	
Символов в латинской графике	
Чисел	
Средняя длина слов	

2. Какому языку соответствует средняя длина слов текста смешанного типа, приведенного в задании 1? Для выполнения задания вычислите среднюю длину слов русского языка из приведенного текста и среднюю длину слов в латинской графике.

3. Определите, каким языкам соответствуют следующие специфические буквы, буквосочетания и слова:

а) 0 б В р со ёis,

б) th sch sc.

в) et, the, der, och, icin.

4. Создайте диагностический словарь для определения языка на материале текстов на двух разных языках (на ваш выбор). Для этого заполните следующую таблицу 1.

Таблица 1

Критерий	Язык 1:	Язык 2:
Типичные артикли		
Указательные местоимения		
Местоимения 3-го лица		
Отдельные формы вспомогательных глаголов		
Основные предлоги и союзы		
Другие частотные слова		

5. Дополните таблицу встречаемости букв в распространенных европейских языках, добавив в нее данные по русскому языку. Используйте для этого любой текст на русском языке объемом не менее 100 символов.

6. Прочитайте несколько фраз на эсперанто. Назовите морфологические диагностические показатели этого языка, учитывая, что существительные и прилагательные на эсперанто всегда имеют одни и те же окончания.

Рус.	Эсперанто
зеленое дерево	verda arbo
старый человек	maljuna viro
хороший друг	bela amiko

## Лабораторная работа № 2

### Тема: АВТОМАТИЧЕСКИЙ АНАЛИЗ ТЕКСТА

**Цель:** изучить особенности автоматического анализа текста.

#### **Краткие теоретические сведения**

Автоматический анализ текста включает ряд этапов:

- 1) графематический анализ: выделение границ слов, предложений, абзацев и других элементов текста (например, врезок в газетном тексте);
- 2) морфологический анализ: определение исходной формы каждого использованного в тексте слова и набора морфологических характеристик этого слова;
- 3) синтаксический анализ: выявление грамматической структуры предложений текста;
- 4) семантический анализ: определение смысла фраз.

**Графематический анализ** определяется также как токенизация (от англ. *token* – отдельное слово, фраза или любой другой значимый элемент текста). Формальными сигналами границ текстовых элементов выступают разделители различного рода: пробелы, обозначающие границы между словами, прописные буквы и знаки препинания, обозначающие границы между предложениями и составными частями предложений, абзацные отступы, обозначающие границы между связанными по смыслу группами предложений и т.п.

При **морфологическом анализе** каждое использованное в тексте слово возводится к его исходной форме и определяется набор морфологических характеристик текстовой формы слова: часть речи; род, число и падеж для существительных, число и лицо для глаголов и т.п.

Для морфологического анализа важно не только понятие машинной основы, понимаемой как последовательность букв от начала словоформы, общая для всех словоформ, входящих в формообразовательную парадигму данного слова. Следующий шаг – это определение частеречной принадлежности слова (частеречный тегинг) и его морфологических характеристик, что чаще всего происходит с опорой на словоизменительные элементы слова (машинные окончания).

Машинные окончания – элементы, описывающие формоизменение конкретной лексемы и представляемые в виде парадигм. Все возможные наборы машинных окончаний зафиксированы в типовой парадигме лексемы. При этом, с одной стороны, можно наблюдать совпадения типовых парадигм (и, соответственно, машинных окончаний) разных лексем, например, ручка и кочка, а с другой – совпадения машинных основ лексем, имеющих разные типовые парадигмы, типовые парадигмы машинной основы лож#, относящейся к лексемам ложь и ложиться.

По машинным окончаниям, входящим в определенные типовые парадигмы, осуществляется полная морфологическая характеристика каждой словоформы, например:

Девочка {девочка = S, жен, од = им, ед} мыла {мыть = V, несов = прош, ед, изъяв, жен, перех | мыло = S, сред, неод = им, мн | = S, сред, неод = род, ед | = S, сред, неод = вин, мн} пол {пол = S, муж, неод = им, ед | = S, муж, неод = вин, ед | = A, кратк, муж, им, ед}

Для автоматического морфологического анализа применяются **парсеры** – специальные компьютерные программы для автоматического анализа слов. Кроме морфологических существуют и синтаксические парсеры, применяемые для автоматического анализа синтаксических структур предложений.

При **синтаксическом анализе** необходимо определить роли слов в предложении и их связи между собой. Результатом этого этапа автоматического анализа является представление синтаксических связей каждого предложения в виде моделей, например, в виде дерева зависимостей.

В модели непосредственно составляющих важно разбиение синтаксической структуры на пары ее элементов: предложение (S) разбивается на группу подлежащего (NP), представленную в данном случае одним существительным (N), и группу сказуемого (VP). Вторая делится на изменяемый глагол (V) и дополнение (O). В дереве зависимостей исходным пунктом анализа выступает сказуемое (V), находящееся в вершине графа, от которого зависят подлежащее (N) и дополнение (O). В итоге в обоих типах анализа выделяются одни и те же синтаксические единицы – N, V и O – но синтаксические отношения между ними оказываются разными. Примеры представлены на рис. 1.1 и 1.2.

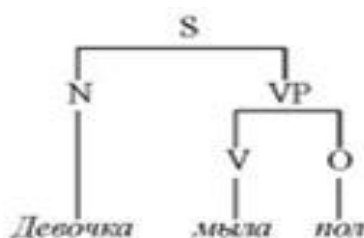


Рисунок 1.1. Дерево непосредственно составляющих

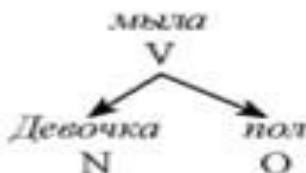


Рисунок 1.2. Дерево зависимостей

**Семантический анализ** представляет собой наиболее сложное направление автоматического анализа текста. В этом случае требуется установление семантических отношений между словами в тексте, объединение различных языковых выражений, относящихся к одному и тому же понятию, и т.п.

### ***Ход работы***

**Задание 1.** Посетите сайт <http://starling.rinet.ru/morph.htm> и протестируйте работу онлайн-анализатора морфологии.

**Задание 2.** Опишите результаты работы онлайн-анализатора в таблице 1.1 для различных частей речи русского языка и изучаемого иностранного языка.

Таблица 1.1– Анализ работы сервиса

Слово	Пример, рус. яз.	Пример, иностр. яз.	Ваш комментарий о возможностях морфологического анализатора
Существительное			
Глагол			
Прилагательное			
Наречие			

**Задание 3.** Посетите сайт <http://sz.ru/parser>. Введите три любых предложения на русском языке, имеющие разную синтаксическую структуру. Сравните результаты их синтаксического анализа в таблице 1.2, оценивая при этом полезность представленной в анализе лингвистической информации. Прокомментируйте возможности применения подобных систем анализа.

Таблица 1.2 Анализ работы сервиса для русского языка

Предложение, рус.	Лингвистическая информация, представленная в синтаксическом анализаторе
1.	
2.	
3.	
Общий комментарий:	

**Задание 4.** Посетите сайт <http://nlp.stanford.edu:8080/parser/index.jsp>. Введите три любых предложения на английском языке, имеющие разную синтаксическую структуру. Сравните результаты их синтаксического анализа в таблице 1.3, оценивая при этом полезность представленной в анализе лингвистической информации. Прокомментируйте возможности применения подобных систем анализа.

Таблица 1.3 Анализ работы сервиса для английского языка

Предложение, англ.	Лингвистическая информация, представленная в синтаксическом анализаторе
1.	
2.	
3.	
Общий комментарий	

**Задание 5.** Посетите сайт <http://teneta.rinet.ru/hudlomer>, помогающий определить функциональный стиль текста. Поместите в поле ввода любой отрывок текста объемом от 75 до 500 слов (примерно от трех абзацев до одной страницы):

- а) из вашей курсовой работы или реферата;
- б) из художественного произведения (используйте для этого, например, библиотеку М. Мошкова, режим доступа: <http://lib.ru>);
- в) газетный текст (используйте текст любого сетевого СМИ, например, [www.rg.ru](http://www.rg.ru)).

Оцените результаты автоматического определения стиля в таблице 1.4. Представьте комментарии о возможностях такой системы.

Таблица 1.4 Стиль текста

Текст	Исходный стиль	Стиль, определенный с помощью сервиса	Ваши комментарии
а			
б			
в			

**Задание 6.** Перейдите по ссылке <http://www.antiplagiat.ru/> и введите текст из вашей актуальной курсовой работы или реферата. Впишите результат получившейся статистики в таблицу 1.5 и ваш комментарий. Для каких целей можно использовать данную программу?

Таблица 1.5 Анализ работы сервиса «Антиплагиат»

Текст	
Результат	
Ваш комментарий	

**Задание 7.** Изучите работу онлайн-сервиса по поиску синонимов: <http://www.raskruty.ru/tools/synonymizer/>. Введите текст из предыдущего задания. После работы сервиса осуществите редакцию текста при необходимости. С помощью сервиса «Антиплагиат» повторно проверьте отредактированный текст. Как изменился процент оригинальности?

### **Контрольные вопросы**

1. Как представляется текст в памяти компьютера?
2. Каковы возможности компьютера по анализу текста?
3. В чем заключаются особенности автоматического анализа текста?
4. Перечислите этапы автоматического анализа текста?
5. Объясните особенности каждого этапа автоматического анализа текста?
6. Приведите примеры онлайн-сервисов, используемых для автоматического анализа текста.

### ***Задания для самостоятельной работы***

Осуществите поиск в сети Интернет других сервисов синонимайзеров. Сравните результаты их работы для одного и того же текста. Отчет представьте в таблице 1.6.

Таблица 1.6 Анализ работы синонимайзеров

Показатели	Первый сервис	Второй сервис
Уровень оригинальности текста до использования сервиса		
Наличие ошибок в согласовании слов		
Необходимость постредактирования		
Уровень оригинальности текста после использования сервиса		

### **Лабораторная работа № 3**

#### **Тема: АВТОМАТИЧЕСКИЙ СИНТЕЗ ДИАЛОГОВ**

***Цель:*** изучить особенности работы сервисов по автоматическому синтезу диалогов

#### ***Краткие теоретические сведения***

Первая компьютерная программа «Элиза», позволяющая вести диалог с пользователем на английском языке была выпущена в 1966 году Д. Вейцнбаумом. Программе использовала в основе простой алгоритм в технике активного слушания и моделирована речевое поведение, а не мышление.

Современные компьютерные программы и онлайн-сервисы, представляющие виртуальных собеседников, способны поддержать общение на различных языках мира. В их основе теперь лежат более сложные алгоритмы, нежели в «Элизе». Например, для создания ботов используется специальный язык разметки для искусственного интеллекта AIML – *Artificial Intelligence Markup Language* [1].

Автоматический синтез диалогов широко используется в компьютерных играх, на порталах предприятий для ответов на вопросы клиентов и т.д.

В настоящее время боты (сокр. от робот) используются для выполнения рутинных действий пользователя, в частности, в различных мессенджерах. Они помогают людям осуществить перевод сообщений, оставлять комментарии, тестировать работоспособность программ, развлекать, осуществлять поиск и др. Таким образом, сфера применения автоматического синтеза диалогов широка и активно востребована.



### ***Ход работы***

**Задание 1.** Побеседуйте на русском языке с виртуальным собеседником по адресу <http://chatbot.tw1.ru/>. Постарайтесь узнать, сколько лет вашему собеседнику. Получили ли вы ответ? Оцените качество синтезируемых реплик и возможности использования данной программы.

**Задание 2.** Побеседуйте на английском языке с тем же виртуальным собеседником. Оцените качество синтезируемых реплик на английском языке. Возможно ли вести с ним разговор на двух языках сразу?

**Задание 3.** Побеседуйте на английском языке с виртуальным собеседником по адресу: английский: [www-ai.ijs.si/eliza/eliza.html](http://www-ai.ijs.si/eliza/eliza.html). Оцените качество синтезируемых реплик собеседника. Воспринимает ли виртуальный собеседник текст без знаков препинания?

**Задание 4.** Организуйте беседу с виртуальным капитаном Кирком: <http://sheepridge.pandorabots.com/pandora/talk?botid=fef38cb4de345ab1&skin=iframe-voice>. Оцените дидактические возможности данной программы для обучения иностранному языку. Какой уровень знаний иностранного языка необходим для ее использования? Можно ли использовать сокращенное написание слов (например, *и* вместо *you*)?

**Задание 5.** Изучите бесплатный сервис <http://viu-piu.iii.ru/> по созданию, обучению и демонстрации другим пользователям инфов – виртуальных существ, с которыми можно поддерживать диалог посредством текстового чата. Каковы дидактические возможности данного сервиса?

### **Контрольные вопросы**

1. Перечислите особенности автоматического синтеза диалогов?
2. Приведите пример первой программы, позволяющей имитировать диалог?
3. Для чего используется автоматический синтез диалогов в настоящее время?
4. Что такое бот?
5. Какие возможности онлайн-сервисов по автоматическому синтезу диалогов?
6. Приведите примеры таких сервисов.

### ***Задания для самостоятельной работы***

**Задание 1.** С помощью сети Интернет осуществите поиск онлайн-сервисов по автоматическому синтезу диалогов на других языках, например, немецком и французском.

**Задание 2.** Сделайте доклад о применении ботов в современном интернете вещей (*Internet of Things*). Как они применяются в сфере бизнеса?

## Лабораторная работа № 4

### Тема: АВТОМАТИЧЕСКОЕ РЕФЕРИРОВАНИЕ И АННОТИРОВАНИЕ ТЕКСТОВ

**Цель:** изучить возможности программных продуктов и онлайн-средств по автоматическому реферированию и аннотированию текстов.

#### *Краткие теоретические сведения*

В связи с тем, что количество информации в мире с каждым годом все увеличивается, специалистам (переводчикам, лингвистам) необходимо быстро находить нужные документы и оперативно знакомиться с их содержанием. Решением такой проблемы может стать создание рефератов и аннотация для любых полнотекстовых документов.

Необходимо дать определение терминам «реферат» и «аннотация».

Под рефератом обычно понимают связанный текст, который кратко выражает центральную тему, предмет, цель, методы и результаты исследования, т.е. основные положения полнотекстового документа.

Аннотация в свою очередь меньше реферата по объему слов и содержит краткое изложение именно содержания документа. Их создают для того, чтобы получить общее представление об изложенной теме.

Рефераты бывают трех видов:

1. Связный текст – новый текст, образованный на основе логико-смыслового анализа первоначального текста.
2. Реферат-клише – создание реферата по заранее заданной структуре.
3. Квазиреферат – перечень наиболее информативных предложений текста.

Есть два способа создания аннотации и рефератов:

- 1) вручную. В этом случае работа проводится либо самим автором текста, либо библиографическим работником;
- 2) автоматически – эти действия выполняются с использованием компьютерных программ специального назначения.

В каждом способе есть свои достоинства и недостатки. В первом случае можно получить аннотации и рефераты высокого качества. Во втором – за короткий промежуток времени можно обработать большое количество текстов. Первый вид рефератов создается с помощью сложных компьютерных программ. Такие программы выделяют из текста наиболее важную информацию с помощью методов искусственного интеллекта и специальных информационных языков [4]. Оставшиеся два вида рефератов – широко распространены в онлайн-сервисах. В основном все программы автоматического составления тестов могут создавать как рефераты, так и аннотации, все зависит только от степени компрессии (сжатия) текста.

Примерами таких программ может служить *Microsoft Word* версии ниже 2007. Далее такая функция была исключена из меню. Также программа компании *IBM «Intelligent Text Miner»*, «*Oracle Context*» от *Oracle*, разработка компании Информатик «ОРФО» и др.

Одной из сложных задач остается создание сводных рефератов (из нескольких документов), т.к. тексты могут быть разнородными и их еще необходимо упорядочить.

Автоматическое сводное реферирование в настоящий момент реализуется следующими сервисами:

- разработка Колумбийского университета *NewsBlaster* для аннотации новостей из сети Интернет;
- автоматическая обработка и систематизации новостей *Google News* и Яндекс.Новости;
- *Ultimate Research Assistant* и др.

Программы автореферирования в основном работают по тем же принципам и этапам создания реферата, что и люди:

1. Подготовительный этап. Здесь необходимо определить тему текста и понять, о чем он.
2. Аналитический этап. Необходимо разделить текст на фрагменты или абзацы. В каждом таком фрагменте выделить ключевые слова. Составить план.
3. Непосредственное составление реферата и/или аннотации. Теперь необходимо соединить выделенные ранее смысловые единицы в связный текст.

К смысловым единицам относят ключевые слова, словосочетания и предложения. Ключевое слово относится к основному содержанию текста, встречается в тексте несколько раз. Ключевые словосочетания включают одно или несколько ключевых слов. Ключевые предложения содержат более двух ключевых слов.

### ***Ход работы***

**Задание 1.** Создание автореферата с помощью *MS Word 2007*.

1. Откройте текстовый документ на русском языке в *MS Word 2007*.
2. В верхней строке командного меню нажмите вкладку «Настройка панели быстрого доступа» (треугольник, обращенный вниз). Выберите вкладку «Другие команды».
3. В открывшемся окне найдите указатель «Выбрать команды из...», в котором нужно установить параметр «Команды не на ленте».
4. Выделите в списке команд «Автосуммирование» и нажмите «Добавить».
5. Далее нажмите «ОК» и убедитесь, что значок «Автосуммирование» появился в верхней строке командного меню.

6. Нажмите значок «Автосуммирование», выберите вкладку «Автореферат». Выберите тип реферата. Установите степень детализации в области «Процент от оригинала», нажмите кнопку «ОК».

7. Проанализируйте получившийся реферат и отразите результаты анализа в таблице 3.1.

Таблица 3.1 Результаты анализа текста

Параметр анализа	Ваш комментарий
Связный текст или набор словосочетаний/предложений	
Функциональная нагруженность элементов реферата	
Отражены ли необходимые структурные компоненты реферата (тема, цель, методы, результаты работы)	
Общий вывод	

Составьте вручную реферат того же самого исходного текста. Ориентируйтесь при этом на этапы составления реферата, указанные выше.

8. Сохраните оба реферата (автоматический и созданный вами) и таблицу с комментарием об автоматическом реферате в *Word*.

9. Создайте с помощью функции «Автосуммирование» реферат на другом языке (английском, французском или немецком).

### **Задание 2.** Реферирование онлайн.

1. Перейдите по ссылке на сервис автоматического реферирования: <http://visualworld.ru/referat.jsp>.

2. Перейдите на вкладку «Рефератор».

3. Введите в область текст или загрузите файл с текстом на русском языке из предыдущего задания.

4. Нажмите кнопку «Сделать реферат!».

5. Сравните реферат, полученный с помощью *MS Word* и данного сервиса.

### **Задание 3.** Реферирование статей из «Википедии».

1. Перейдите по ссылке в проект *WikiSummarizer*. Он предлагает автоматическое сокращение статей из «Википедии»: <http://www.wikisummarizer.com/Pages/Default.aspx>.

2. Введите ключевое слово (на английском языке), по которому необходимо сделать автореферат и нажмите кнопку «*Summarize*».

3. Изучите все виды рефератов, полученных с помощью сервиса.

### **Задание 4.** Создание облака слов/тэгов.

Изучите следующие сервисы для создания облака слов. Продемонстрируйте работу с ними:

– <https://wordart.com/> – сервис для генерация облака слов с действующими ссылками из текста или по заданному адресу в виде графических изображений;

- <https://www.text2mindmap.com/> – сервис создания из слов интеллект-карты;
- <https://wordcloud.pagemon.net/> – создание облака слов для конкретной страницы сайта.

**Задание 5.** Работа с сервисом Яндекс.Новости.

1. Изучите алгоритм работы сервиса Яндекс.Новости: <https://yandex.ru/support/news/index.html>.
2. Ознакомьтесь с процессом создания таких новостных сводных рефератов в презентации: <http://cache-mskstoredata10.cdn.yandex.net/download.cdn.yandex.net/from/yandex.ru/support/ru/news/files/About-News.pdf>.

**Задание 6.** Знакомство с онлайн-сервисами автореферирования.

1. Изучите возможности сервиса: <http://ultimate-researchassistant.com/GenerateResearchReport.asp>.
2. Изучите возможности сервиса: <http://iresearch-reporter.com/#/irr>.
3. Сравните их возможности.

**Контрольные вопросы**

1. Дайте определение термину «реферат».
2. Дайте определение термину «аннотация».
3. В чем принципиальная разница между этими понятиями?
4. Перечислите виды рефератов.
5. Назовите способы создания рефератов и аннотаций. В чем состоит преимущество каждого способа?
6. Приведите примеры программных продуктов и онлайнсервисов для создания авторефератов и аннотаций.

**Задания для самостоятельной работы**

Изучите возможности других программ для автоматической обработки текстов:

- режим доступа: <http://tpl-it.wikispaces.com/%D0%9F%D0%A1+%D0%B4%D0%BB%D1%8F+%D0%BE%D0%B1%D1%80%D0%B0%D0%B1%D0%BE%D1%82%D0%BA%D0%B8+%D1%82%D0%B5%D0%BA%D1%81%D1%82%D0%B0>;
- режим доступа: <http://www.analyst.ru/index.php?lang=rus&dir=content/products/&id=ta>; режим доступа: <http://www.legroom.net/software/uniextract>.

## Лабораторная работа №5

### Тема: РАБОТА С СЕРВИСОМ YouTube

**Цель:** изучить возможности онлайн-сервиса *youtube.com*.

#### **Краткие теоретические сведения**

Сервис *http://YouTube.com* – это платформа для распространения видео. Сервис позволяет пользователям просматривать видео, выкладывать видео, комментировать ролики, оценивать их и делиться ими друг с другом. Следовательно, данный сервис можно воспринимать не только как платформу для хранения видео, но и как социальную сеть.

Изначально *http://YouTube.com* задумывался как сайт для знакомств. От существующих сайтов знакомств он отличался возможностью загрузки собственного видео. Идея принадлежала Чаду Херли (*Chad Hurley*), Стиву Чену (*Steve Chen*) и Джавиду Кариму (*Jawed Karim*). Первое видео было выложено в 2005 г.

Однако пользователи сети Интернет определили судьбу сервиса сами: они выкладывали видео не о себе, а о том, что происходит вокруг. К концу 2005 года сервис получает большие инвестиции и осуществляет ребрендинг. В 2006 году платформу выкупает компания *Google*.

#### **Ход работы**




**Задание 1.** Загрузка видео.

1. Откройте сайт *http://www.youtube.com/*.
2. Нажмите на кнопку «Добавить видео».
3. Перед вами появится окно входа в аккаунт *Google*. Если вы не зарегистрированы, то перейдите по ссылке «Создать аккаунт» и зарегистрируйтесь. Если вы зарегистрированы, то войдите в свой аккаунт на *Google*.
4. В открывшемся окне вы можете создавать, загружать и редактировать видео. Загрузите видео (выберите файл небольшой длительностью около минуты) с помощью ссылки «Выберите файлы для загрузки» (см. примеры видеофайлов в папке «Видео»).
5. По окончании загрузки заполните поля в разделе «Основная информация»: Название, Описание, Тэги. В списке настроек выберите доступ к видео: Открытый доступ (видео доступно всем пользователям), Доступ по ссылке (видео будет доступно только тем, у кого есть ссылка) и Ограниченный доступ.
6. Перейдите в раздел «Расширенные настройки» и заполните поля по усмотрению.
7. По окончании нажмите «Готово».
8. Обменяйтесь ссылками на видео, просмотрите фрагмент и оставьте комментарии, если они разрешены.

## **Задание 2.** Создание слайд-шоу.

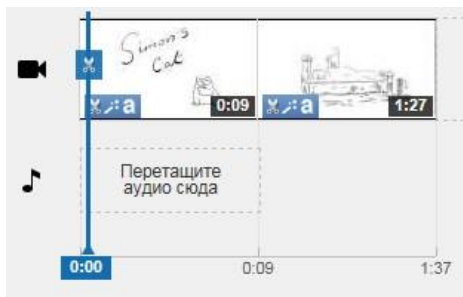
1. Подготовьте несколько фотографий для создания слайдшоу. Сохраните их в отдельную папку.
2. Снова перейдите на сайт <http://www.youtube.com/> и нажмите кнопку «Добавить видео».
3. В области слайд-шоу нажмите кнопку «Создать».
4. Перейдите в раздел Загрузка фотографий.
5. Укажите название альбома и загрузите подготовленные фотографии, нажав на кнопку «Выбрать...».
6. После загрузки фото можно поменять их порядок, перетаскивая мышью. Затем нажмите на кнопку «Далее».
7. Установите настройки: Время показа, Эффекты слайда, Переход.
8. Затем выберите музыкальное сопровождение в разделе Аудио. Можно выбрать направление музыки в списке «Композиции».
9. По окончании нажмите на кнопку «Загрузить».
10. По окончании загрузки заполните поля в разделе «Основная информация»: Название, Описание, Тэги. В списке настроек выберите доступ к видео: Открытый доступ (видео доступно всем пользователям), Доступ по ссылке (видео будет доступно только тем, у кого есть ссылка) и Ограниченный доступ.
11. Перейдите в раздел «Расширенные настройки» и заполните поля по усмотрению.
12. По окончании нажмите «Готово».
13. Обменяйтесь ссылками на видео, просмотрите фрагмент и оставьте комментарии, если они разрешены.

## **Задание 3.** Редактирование видео.

1. Снова перейдите на сайт <http://www.youtube.com/> и нажмите кнопку «Добавить видео».
2. В разделе «Видеоредактор» нажмите на кнопку «Изменить».
3. Выберите одно из загруженных вами ранее видео и перетащите его в область «Перетащите видео сюда».
4. Затем добавьте таким же образом второе видео.
5. В разделе «Быстрое исправление» измените настройки. Например, установите автокоррекцию, яркость и контраст, можете замедлить видео или поворачивать и т.д.
6. Просмотрите фрагменты с помощью ползунка в правой части экрана . Верните ползунок в исходное положение.
7. Добавьте фото к видео с помощью кнопки .
8. Выберите в области аудио  файл, установите настройки (например, громкость) и перетащите его под видео.
9. Изучите и продемонстрируйте остальные возможности редактора.

10. Для того чтобы редактировать видео, необходимо щелкнуть по нему левой кнопкой мыши и выбрать фильтры в области «Фильтры», в области «Текст» введите текст и установите настройки к нему.

11. Можно изменять длительность видео с помощью кнопки «Ножницы»:



12. Дождитесь обработки видео.

#### **Задание 4.** Работа с каналами.

1. Перейдите в «Менеджер видео».
2. Убедитесь, что все ваши видео загружены.
3. Перейдите в «Панель управления», изучите статистику просмотров ваших файлов.
4. Перейдите в «Сообщество» в разделе «Комментарии» и изучите их.
5. Перейдите в «Канал», изучите возможности продвижения видео в сети Интернет. В разделе «Фирменный стиль» добавьте логотип канала. Его можно предварительно создать в графическом редакторе в соответствии с требованиями.

#### **Задание 5.** Подписка на каналы.

1. Подпишитесь на канал ЮУрГГПУ. Для этого на странице <http://www.csru.ru/> выберите значок YouTube. Перейдите по ссылке. Выберите пункт «Подписаться».

2. Подпишитесь на канал DisneyPixar (рис. 4.1).

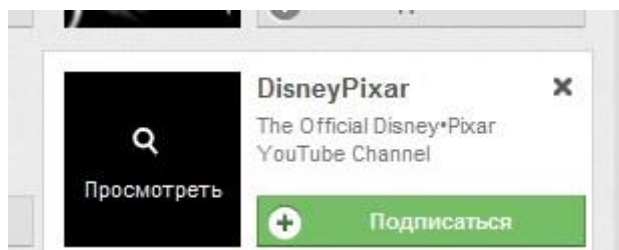



Рисунок 4.1. Оформление подписки на канал

#### **Задание 6.** Скачивание видео.

1. Найдите видео с рекламой любого фильма (например, Соловей-разбойник) в строке поиска.
2. Скачайте его. Это можно сделать двумя способами:



2.1.К адресу ролика добавьте «ss», например:  
<http://ssyoutube.com/watch?v=Y0cmSsBfafg>.

2.2.Перейдите на сайт <http://ru.savefrom.net/>, вставьте ссылку в строку поиска и нажмите на кнопку .

3. Выберите нужный формат, например, *flv*, и нажмите «Скачать».

#### **Задание 7. Работа с субтитрами.**

1. Просмотрите видео «Как добавить субтитры к ролику»:  
<https://www.youtube.com/watch?v=LCZ-cxfxzvk>.
2. Откройте Менеджер видео.
3. Выберите видео, для которого нужно создать субтитры. Нажмите на раскрывающееся меню справа от кнопки «Изменить».
4. Выберите пункт «Субтитры».
5. Нажмите «Добавить новые субтитры».
6. Выберите способ добавления или редактирования субтитров.

#### **Задание 8. Работа с автоматическими субтитрами.**

1. Откройте Менеджер видео.
2. Выберите видео, для которого нужно создать автоматические субтитры. Нажмите на раскрывающееся меню справа от кнопки «Изменить».
3. Выберите пункт «Субтитры».
4. Выберите субтитры, которые были сгенерированы автоматически (справа от видео в списке «Опубликованные» рядом с языком указано «автоматически»).
5. Прочитайте субтитры и при необходимости измените их.

#### **Контрольные вопросы**

1. Что такое *YouTube*?
2. Пишите историю сервиса *YouTube*.
3. Перечислите основные возможности сервиса по работе с видео.
4. Перечислите возможности *YouTube* как социальной сети.

#### **Задания для самостоятельной работы**

1. Изучите справку по работе с видеоредактором:  
<https://support.google.com/youtube/answer/183851?hl=ru>. Всеми ли функциями вы воспользовались? Если нет, то продемонстрируйте работу с ними.
2. Изучите и продемонстрируйте остальные возможности работы с каналами.

## Лабораторная работа №6

### Тема: КОМПЬЮТЕРНОЕ ОБУЧЕНИЕ ЯЗЫКАМ

**Цель:** изучить особенности обучения языкам с помощью интернет-сервисов.

#### ***Краткие теоретические сведения***

Лингвистика – это наука о естественном человеческом языке. Традиционно ее делят на теоретическую лингвистику или фундаментальную и прикладную. Теоретическая лингвистика направлена на объективное установление состояния отдельного языка, его истории и закономерностей. Прикладная лингвистика развивается только с конца двадцатых годов прошлого века. Ее вопросы связаны с разработкой решения практических задач использования языка. И точка зрения на содержание практической лингвистики и подходы к ней различаются. За рубежом под прикладной лингвистикой понимается процесс обучения иностранному языку, методика преподавания, описание грамматики для учебных целей. В России же прикладная лингвистика связана с информационными и компьютерными технологиями и автоматическими системами обработки информации [3].

Рассмотрим подходы к обучению иностранному языку с помощью компьютера (*Computer Assisted Language Learning, CALL*). Сначала выделим общие подходы:

1. Грамматико-переводной метод. Целью такого подхода является чтение текстов и их перевод на родной язык, а также изучение грамматики.
2. Прямой метод обучения иностранным языкам. Используется при полном погружении в языковую среду. При этом освоение грамматики происходит не в системе, а от примеров к правилу.
3. Коммуникативно-ориентированный подход. При этом подходе происходит изучение в рамках речевых ситуаций из практики носителей языка, а не конкретных тем.

В компьютерной лингвистике сложилась новая технология

*CALL*, объединяющая методы обучения иностранным языкам. Использование возможностей компьютера позволяет моделировать языковую среду, имитировать лингвистическую и коммуникативную реальность. Для поддержки такого обучения используется широкий класс программного обеспечения. Это могут быть и электронные словари, справочники, компьютерные курсы, системы машинного перевода, гипертекстовые документы и т.п. [2].

#### ***Ход работы***

**Задание 1.** Перейдите по ссылке <http://www.cambridge.org/us/cambridgeenglish/catalog> и изучите каталог продуктов Кембриджского университета для изучения иностранных языков. Выберите раздел, например, *Listening & Speaking* (рис. 5.1).

Нажмите на кнопку «View product» и изучите особенности курса. Ознакомьтесь с доступными ресурсами на вкладке «Resources» и скачайте их, нажав на кнопку «Download».

Оцените возможности использования такого каталога при изучении иностранных языков.



Рисунок 5.1. Раздел *Listening & Speaking*

**Задание 2.** Перейдите по ссылке <http://www.arcademics.com/> и откройте раздел *Language Arts* (рис. 5.2).

Ознакомьтесь с приложениями для изучения английского языка *Coconut Vowels*, *Elephant Feed*, *Furious Frogs* и другие.



Рисунок 5.2. Раздел *Language Arts*

Определите вид мультимедийной программы и теоретический подход, использованный при ее создании.

Результаты анализа представьте в таблице 5.1.

Таблица 5.1 Анализ игр

Программа	Вид	Подход	Обоснование
Coconut Vowels			
Coconut Vowels			
Elephant Feed			
Furious Frogs			
Geraffe Karts			
И др.			

**Задание 3.** Перейдите на сайт Интернет-университета информационных технологий по ссылке [www.intuit.ru](http://www.intuit.ru).

Зарегистрируйтесь на сайте, выберите один из бесплатных дистанционных курсов и запишитесь на него.

Изучите информацию о курсе, результаты представьте в таблице 5.2.

Таблица 5.2 Анализ курса

Название курса	Автор курса	Цель	Уровень	Кол-во часов

Просмотрите выборочно несколько модулей курса. Результаты анализа представьте в таблице 5.3.

Таблица 5.3 Анализ модуля

Параметр	Описание
Составляющие курса (вид: урок, тема или др., количество, примеры)	
Возможные трудности	
Подтверждение (сертификат, свидетельство, диплом)	

**Задание 4.** Найдите в интернете определение «портала» и требования к нему. Определите, какие из перечисленных веб-ресурсов являются порталами. Результаты анализа представьте в таблице 5.4.

Таблица 5.4 Анализ веб-ресурсов

Ресурс	Да/нет	Обоснование
<a href="http://www.all-abc.ru">www.all-abc.ru</a>		
<a href="http://www.gramota.ru">www.gramota.ru</a>		
<a href="http://pearsonpte.com">http://pearsonpte.com</a>		
<a href="http://deutsche-sprache.ru">http://deutsche-sprache.ru</a>		
<a href="http://www.english.language.ru">www.english.language.ru</a>		

**Задание 5.** Найдите с помощью различных поисковых систем и укажите в таблице по два примера русскоязычных и иноязычных интернет-ресурсов (на английском, русском, немецком или французском языке). Результаты представьте в таблице 5.5.

Таблица 5.5 Анализ ресурсов

Вид ресурса	Русский язык	Иностранный язык
Электронная библиотека		
Электронный журнал		

**Контрольные вопросы**

1. Чем занимается наука лингвистика?
2. Назовите вопросы теоретической лингвистики.
3. Назовите вопросы прикладной лингвистики.
4. Перечислите общие подходы к обучению иностранным языкам.
5. Каковы особенности каждого подхода?
6. В чем особенность компьютерного обучения иностранным языкам.
7. Приведите примеры программных продуктов, используемых при обучении иностранным языкам с помощью компьютера.

**Задания для самостоятельной работы**

Приведите примеры других интернет-ресурсов, используемых при обучении иностранному языку с помощью компьютера. Результат представьте в таблице (табл. 5.6).

Таблица 5.6 Примеры интернет-ресурсов

	Пример 1	Пример 2	Пример 3
Название			
Ссылка (URL)			
Возможности ресурса			
Ваш комментарий			

**Лабораторная работа №7**

**Тема: СКАНИРОВАНИЕ ТЕКСТОВОГО ДОКУМЕНТА**

**Цель:** изучить особенности сканирования документов и распознавания текстов.

**Краткие теоретические сведения**

Автоматический анализ текста должен начинаться с его ввода в компьютер. И в этом случае есть два варианта: набор текста на компьютере или его сканирование. Для первого варианта необходимы устройство ввода информации – клавиатура, во втором – сканер.

Сканер – это устройство ввода информации в компьютер, которое работает по принципу фотоаппарата и таким образом позволяет компьютеру «увидеть» текст в виде фотографии.

В настоящее время существует несколько оснований для классификации сканеров:

1. По типу обрабатываемых изображений сканеры делятся на цветные сканеры и черно-белые. Черно-белые делятся на штриховые и полутоновые (их еще называют «серыми»).

2. По способу считывания сканеры делятся на ручные (для сканирования штрих-кодов, лазерные и фотосканеры) (рис. 6.1) и настольные (бывают планшетные (рис. 6.2), рулонные (рис. 6.3) и проекционные). В настоящее время активно применяются 3D-сканеры (рис. 6.4).

3. По типу используемого интерфейса (способу подключения): USB, SCSI и параллельный порт.

4. По области применения: SOHO (для домашнего и офисного использования), CORPORATE (сканеры бизнес-класса), GRAPHIC ART (для профессионального использования в работе дизайнеров и художников), PREPRESS (для издательских компаний), CAD (для использования в инженерии) и DOCUMENT SCANNER (для компаний с большим документооборотом с целью сканирования большого количества информации).



Рисунок 6.1. Пример ручного сканера



Рисунок 6.2– Пример планшетного сканера



Рисунок 6.3– Пример рулонного сканера



Рисунок 6.4. Пример ручного 3D-сканера

Однако после сканирования необходимо перевести фото в текст. Для того чтобы компьютер смог «понять» текст, т.е. перевести графическое (растровое) изображение символов в текстовую форму, при которой у каждого символа имеется свой двоичный код (например, в системе кодировок *ASCII*), требуется специальная программа автоматического распознавания символов (*OCR, Optical Character Recognition*).

Такая программа (*OCR-программа*) – это компьютерная программа, позволяющая преобразовать отсканированный текст в электронный текстовый файл, который можно затем обработать в текстовом редакторе.

В качестве примеров можно привести следующие:

- *ABBYY FineReader*.
- *OmniPage* от компании *Nuance Communications*.
- *CuneiForm* от российской компании *Cognitive Technologies*.
- Онлайн-ресурсы: *CVisionTech* <http://www.cvisiontech.com/online-tool/ocr-convert-compress-pdf.html>, *OnlineOCR* <https://www.onlineocr.net/>, *FreeOCR* <http://www.free-ocr.com/>, *OCRTerminal* <http://www.abbyyonline.com/en/home/ocrterminal>.

Как показывает практика, результат работы большинства *OCR-программ* достаточно точен. Однако все возникающие при сканировании трудности впоследствии необходимо исправлять вручную.

Рассмотрим трудности, которые могут привести к ошибкам при распознавании текстов:

- использование шрифта разной гарнитуры и размера;
- использование текста на нескольких языках;
- размещение текста в несколько колонок;
- включение в текст таблиц и рисунков;
- искажения символов (разрывы, слипания букв и т.п.);
- посторонние включения в изображение и т.д.

Для решения возникающих проблем разработчики предлагают использовать программы распознавания текста. Ориентирующиеся на смысл текста, т.е. программные продукты выводятся на более высокий **уровень искусственного интеллекта**. Такой класс программ называют **ICR-программами**.

На сегодняшний день проблема массового ввода текста в компьютер считается практически решенной, так как точность распознавания компьютером текстов хорошего и среднего качества достигает 99%.

### ***Ход работы***

**Задание 1.** Просмотрите фрагмент видео для получения представления о том, как работает сканер: <https://www.youtube.com/watch?v=QoC1ilZx73Q>.

**Задание 2.** Просмотрите фрагмент передачи о работе и использовании 3D-сканеров: <https://www.youtube.com/watch?v=ZhMFDuBAmna>.

**Задание 3.** Сканирование и распознавание документов с помощью онлайн-сервисов.

3.1. Убедитесь, что сканер подключен к сети питания и к компьютеру.

3.2. Выберите три страницы печатного текста, содержащих только текста, текст и таблицу, текст и изображение.

3.2. Поместите первую выбранную страницу в сканер, совместив левый нижний угол сканируемого текста с уголком на разметке сканера.

3.3. Нажмите на кнопку меню «Пуск», выберите пункт «Сканеры и факсы».

3.4. В открывшемся окне программы на панели инструментов нажмите на кнопку «Новое сканирование».

3.5. Установите необходимые настройки сканирования (например, тип документа, разрешение и др.). Нажмите на кнопку «Сканировать».

3.6. Повторите операцию для оставшихся страниц.

3.7. Убедитесь, что отсканированные изображения появились в папке «Документы» – «Отсканированные изображения». Отредактируйте их в графическом редакторе: обрежьте лишние поля, отрегулируйте контраст, яркость.

3.8. Ознакомьтесь с сервисом <http://www.free-ocr.com/>





Рисунок 6.5. Онлайн сервис Free OCR

3.9. Нажмите на кнопку «Выберите файл» и укажите путь к файлу с изображением первой страницы на вашем компьютере.

3.10. В поле «Language» укажите язык документа, например, Russian (русский). Языков можно выбрать несколько, если текст многоязычный.

3.11. Нажмите на кнопку «Start».

3.12. В области Result появится результат перевода. Щелкните в этой области, выделите весь текст сочетанием клавиш Ctrl+A, скопируйте и вставьте его в Word. Тест сохраните.

3.13. Также можно скачать файл в формате \*.txt, нажав по ссылке с именем файла после «Download as text file».

3.14. Повторите действие по распознаванию текста с двумя оставшимися изображениями.

3.15. Ознакомьтесь с сервисом <http://www.newocr.com/>



Рисунок 6.6– Онлайн-сервис Free Online OCR

3.16. Нажмите кнопку «Обзор» в области «Select your file» и укажите путь к файлу с рисунком на компьютере.

3.17. Затем нажмите на кнопку «Preview».

3.18. В открывшемся окне программа определит язык распознавания. Можно удалить язык, если он определен неправильно, нажав на крестик. Если все правильно, то нажмите на кнопку «OCR».

3.19. При необходимости измените размер области для распознавания.

3.20. Сохраните файл в формате doc, выбрав его по нажатию на кнопку «Download».

3.21. Повторите операцию распознавания оставшихся изображений с помощью данного сервиса.

3.22. В итоге должно получиться шесть документов: три распознанных с помощью первого сервиса и три с помощью второго. Сравните результаты распознавания одного и того же изображения разными сервисами. Для этого можно воспользоваться *MS WORD* (в меню «Рецензирование» выберите пункт «Сравнить»).

### ***Контрольные вопросы***

1. Что такое сканер?
2. Перечислите принципы работы сканера.
3. Как классифицируются сканеры?
4. Какой сканер лучше выбрать для домашнего использования?
5. Приведите примеры программ для распознавания текста.
6. Какой из перечисленных онлайн-сервисов для распознавания лучше справился со своей функцией?

### ***Задания для самостоятельной работы***

1. Из предложенного списка онлайн-сервисов выберите тот, с которым еще не осуществлялась работа. Распознайте отсканированные изображения. Проанализируйте результаты его работы в сравнении с *Free OCR* и *Free Online OCR*.

2. Распознайте изображения с помощью демонстрационного портала программы *FineReader* <https://finereaderonline.com/ruru/Tasks/Create>. Сравните результаты.

## **Лабораторная работа №8 СОЗДАНИЕ САЙТА С ПОМОЩЬЮ UML.CMS**

***Цель занятия:*** изучить основные требования к сайтам и освоить приемы управления сайтом в программах класса *CMS*.

### ***Краткие теоретические сведения***

Сайт организации – это ее «лицо», представляющее организацию в сети Интернет. Такую же роль играет и персональный сайт.

Создавать сайты организаций можно различными способами. Один из распространенных – это системы управления контентом (*Content Management System, CMS*). Такие системы позволяют создать сайт, наполнить его содержимым (контентом), администрировать, увеличивать функциональность и т.д.

*CMS* делятся на три вида: коробочные, свободные, индивидуальные (студийные). Первый вид – это коммерческие программные продукты. Второй вид можно использовать свободно в любых целях. Третий вид – *CMS*, разработка сайтов, которые может использовать только их разработчик.

По данным аналитического портала рынка веб-разработок «*CMS Magazine*» (<http://www.ratingruneta.ru/cms/>) самыми популярными являются следующие виды *CMS*:

1. Коробочные:  
1С-Битрикс. UMI.CMS.  
NetCat.
2. Свободные: Drupal.  
Joomla!  
MODX.  
WordPress.
3. Индивидуальные:  
Текарт CMS/CMF. UralCMS.  
RBC Contents.

### ***Ход работы***

**Задание.** Изучите возможности системы UMI.CMS. Для этого:

1. Перейдите по ссылке к каталогу шаблонов сайтов: <https://demo.umi-cms.ru/create/category:all:commerce>.
2. Перейдите в отрасль «Образование», раздел «Персональные сайта». Выберите шаблон «Сайт переводчика». И нажмите «Создать демо».
3. Введите защитный код и нажмите «Создать сайт» (рис. 7.1).

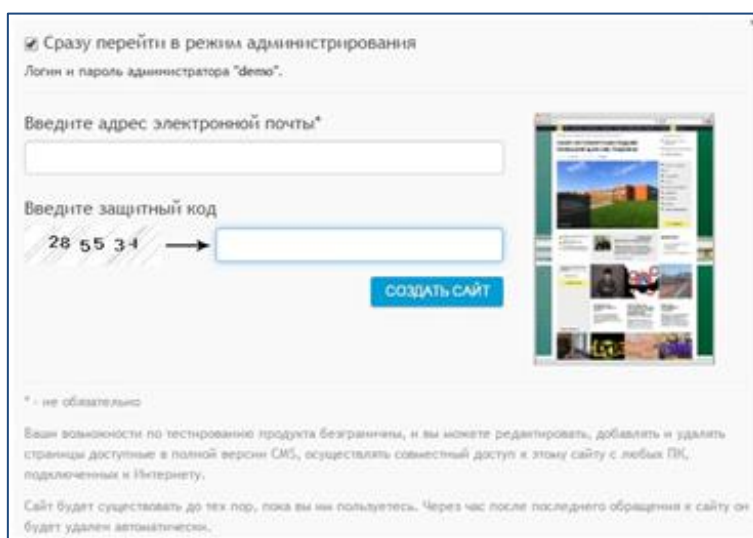


Рисунок 7.1. Окно создания сайта

4. Ознакомьтесь с полученным сайтом.
5. Изучите справку по работе с системой UMI.CMS <http://help.docs.umi-cms.ru/>.
6. Вернитесь на вкладку с созданным сайтом. Перейдите в режим редактирования. Для этого на панели инструментов нажмите на кнопку «Редактирование».
7. Внесите изменения: удалите один из разделов, измените названия оставшихся разделов, загрузите новые фотографии. При внесении изменений устанавливайте заметки. Для этого щелкните в меню «Заметка», а затем щелчком мыши установите ее в необходимое место.
8. Внесите свое содержание: три новости, контакты.
9. Перейдите в раздел «Модули» – «Структура» (в панели меню слева сверху). Изучите структуру сайта, внесите изменения. Добавьте новый раздел. Заполните его содержимым.
10. Перейдите в раздел «Модули» – «Меню». Добавьте пункты меню.
11. Сохраните изменения. Перейдите в режим просмотра, убедитесь, что сайт обновился. Покажите изменения преподавателю.

### ***Контрольные вопросы***

1. Что такое CMS?
2. Приведите примеры CMS.
3. Укажите основные принципы работы по созданию сайта в CMS.
4. Перечислите возможности UMI.CMS по созданию и администрированию персональных сайтов.

### ***Задания для самостоятельной работы***

Изучите особенности работы системы 1С-Битрикс: <https://www.1c-bitrix.ru/>. Каковы ее достоинства, недостатки по сравнению с другими системами?

## **Лабораторная работа № 9**

### **Тема: СОЗДАНИЕ ПЕРСОНАЛЬНОГО САЙТА ПЕРЕВОДЧИКА С ПОМОЩЬЮ КОНСТРУКТОРА WIX.COM**

**Цель:** разработать персональный сайт переводчика с помощью онлайн-конструктора *Wix.com*.

### ***Краткие теоретические сведения***

Создание веб-сайта – это серьезный процесс, который требует планирования и организации большого объема работ. Для решения данной проблемы конструктор сайтов Wix предлагает большое количество шаблонов. Именно шаблоны могут быть положены в основу сайта, а дальнейшая их настройка с помощью большого количества инструментов позволяет сделать сайт индивидуальным.

Рассмотрим процесс планирования. На первом этапе необходимо определить цель разработки сайта. Самый главный вопрос в этом случае: «Для чего нужен сайт?». Здесь нужно учитывать специфику сайта, его предназначение. Например, хотите ли получать прибыль от посещения клиентами сайта или комментарии от клиентов, организовывать регистрацию клиентов, либо чтобы они оставляли вопросы и/или делали заказы. В этом случае важно четко определить цели и понимать разницу между средством и целью.

В качестве примера можно привести следующие цели: демонстрация образцов работ; получение звонков от клиентов для заказа работ; подписка клиентов на рассылку.

Тогда демонстрация образцов работ будет не целью, а лишь средством привлечения клиентов.

Далее сформулированные цели необходимо расставить в порядке уменьшения приоритетов. От этого зависит содержание и компоновка сайта. Например, кнопка оформления заказа должна быть на главной странице сайта, а подписка на рассылку на второстепенной.

На следующем этапе планирования определяется целевая аудитория сайта: кому он адресован? Обычно определяются интересы целевой аудитории, примерный возрастной диапазон. После этого можно указать, чем их будет привлекать сайт, что может клиентов заинтересовать на нём.

На следующем шаге планирования необходимо распределить содержание сайта по предполагаемым страницам. Основные акценты расставляются на главной странице или ее еще называют домашней. Там можно расположить текст, применить визуальные графические или видеоэффекты. Обязательно сайт должен содержать страницы с контактной информацией (Контактная информация, Контакты или Для связи), а также информацию о переводчике/компании (О нас и др.).

Для упрощения работы по созданию дизайна сайта конструктор Wix предлагает большое количество шаблонов. Любой шаблон можно изменить: поменять цвет фона, удалить или добавить страницы, добавить или изменить меню, приложения и прочее.

Следующим шагом в разработке сайта является наполнение его контентом. Для этого необходимо осуществить подготовительную работу:

1. Подготовить тексты для размещения на всех страницах сайта.
2. Подготовить изображения: осуществить редактирование, оптимизацию размеров и объемов и др.

После этого можно переходить к редактированию каждой страницы отдельно. Дизайнеры рекомендуют не приступать к созданию следующей страницы, пока не закончена работа с предыдущей.

### ***Ход работы***

**Задание 1.** Просмотрите ролик о сервисе [https://www.youtube.com/watch?time\\_continue=53&v=DgBJ2\\_gNfow](https://www.youtube.com/watch?time_continue=53&v=DgBJ2_gNfow).

**Задание 2.** Сформулируйте цели создания сайта для переводчика.

**Задание 3.** Расставьте сформулированные в предыдущем задании цели в порядке уменьшения приоритета.

**Задание 4.** Определите круг потенциальных посетителей сайта переводчика. Результат представьте в таблице (табл. 8.1).

Таблица 8.1 Анализ посетителей сайта

Кому адресован сайт	
Примерный возраст посетителей	
Каковы их предпочтения	
Чем можно заинтересовать	

**Задание 5.** Перейдите по ссылке <http://ru.wix.com/sample/website>, чтобы изучить образцы сайтов, созданных помощью конструктора Wix (рис. 8.1).



Рисунок 8.1. Примеры сайтов, созданных с помощью Wix

**Задание 6.** Изучите и выберите подходящий шаблон для вашего будущего сайта по ссылке: [http://ru.wix.com/website/templates/?utm\\_campaign=bloglinks\\_ru&experiment\\_id=templates\\_ru](http://ru.wix.com/website/templates/?utm_campaign=bloglinks_ru&experiment_id=templates_ru).

**Задание 7.** Нарисуйте в графическом редакторе или на бумаге примерное содержимое нескольких страниц вашего сайта в соответствии с выбранным шаблоном.

**Задание 8.** Приступите к редактированию вашего сайта.

Для этого:

1. После выбора шаблона нажмите на кнопку «Редактировать».
2. Перед вами откроется окно Авторизация/Регистрация. Если у вас еще нет аккаунта на данном сайте-конструкторе, тогда зарегистрируйтесь (рис. 8.2).

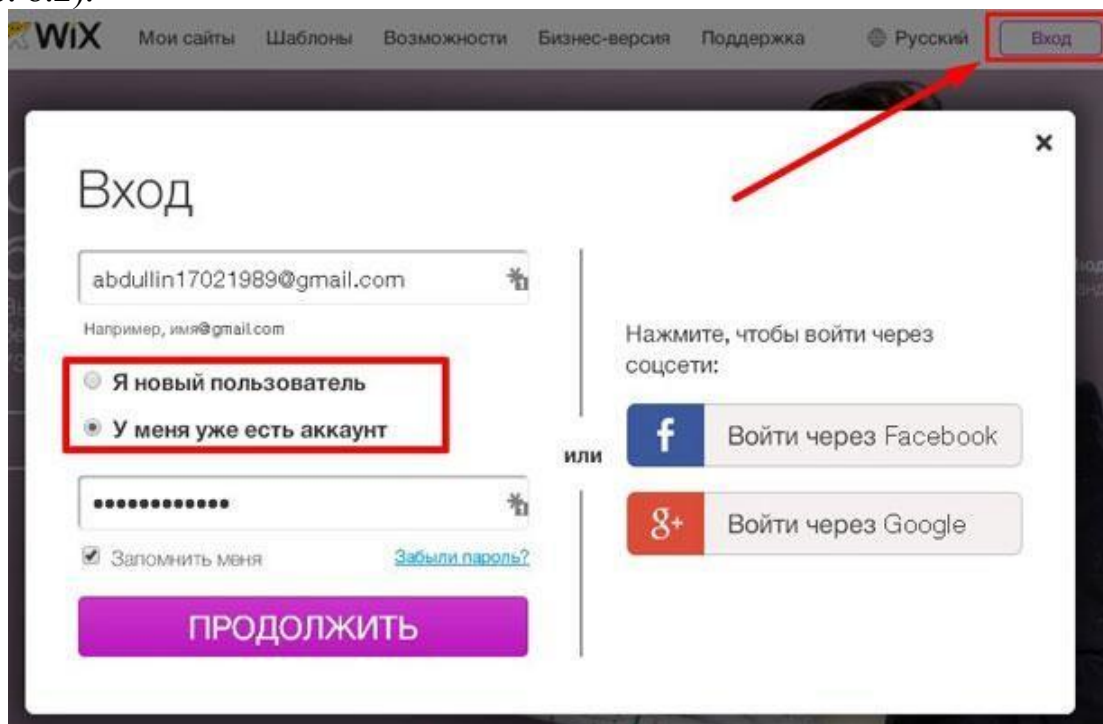


Рисунок 8.2– Авторизация/регистрация на сайте

3. После регистрации откроется окно, в котором будет расположена кнопка «Начните создание сайта!». Нажмите на неё.

4. Если вы уже выбрали шаблон, то он откроется для редактирования. Если нет, то вам предложат выбрать шаблон для своего веб-ресурса. После того, как вы подтвердили свой выбор шаблона, введите название своего проекта и сохраните.

5. Отредактируйте будущий сайт в соответствии с заготовками (см. задание 7).

Слева имеется панель инструментов со следующими функциями:

«Фон» (выберите подходящий фон для страниц сайта).

«Добавить» (различные объекты: текст, фото, видео, меню, кнопки, контакты и прочее) (рис. 8.3).



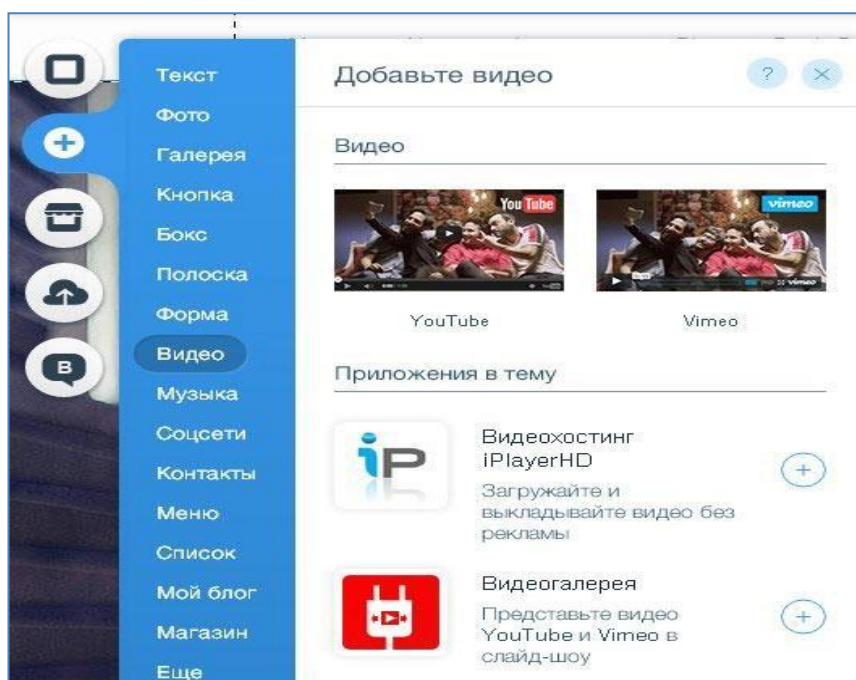


Рисунок 8.3– Добавление различных объектов на страницы

«WIX App Market» – специальный раздел в настройках, который позволяет устанавливать любые приложения на страницах (рис. 8.4). Выберите бесплатные и разместите их на страницах своего сайта.

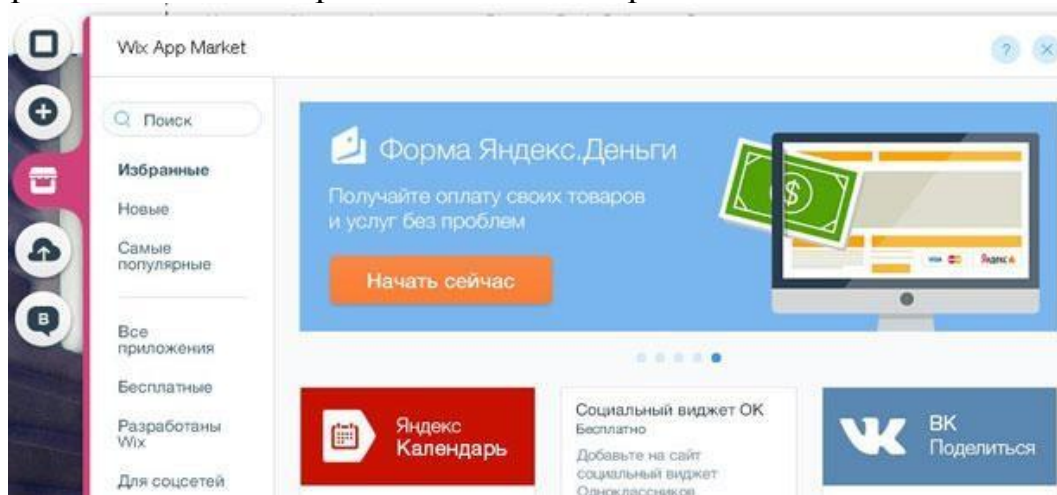


Рисунок 8.4– Функция «Wix App Market»

«Мои загрузки». В этом разделе имеется возможность загружать свои файлы (фото, видео, документы) для их последующего размещения на сайте.

«Блог» – функция предназначена для размещения блога на сайте с целью создания постов, публикации в социальных сетях и т.д.

6. В редакторе имеется также верхнее меню. Изучите две его возможности: Панель управления и Настройки (рис. 8.5).

7. В пункте «Настройки сайта» сделайте сайт видимым для всех пользователей. Это позволит редактировать сайт без показа его будущим пользователям.



8. В меню «Настройки» в разделе *Seo* включите функцию «Позволить поисковикам найти мой сайт», это предоставляет возможность раскручивать свой сайт в поисковых системах.

9. В любой момент редактирования можно просмотреть получившийся сайт – кнопка «Предпросмотр».

10. Если редактирование закончено, то итоговый вариант сайта можно опубликовать – кнопка «Опубликовать».

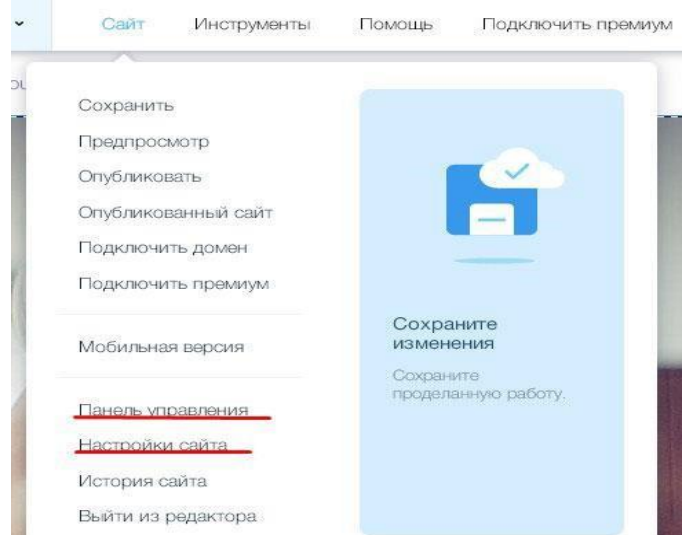


Рисунок 8.5 Пункты верхнего меню

Далее вам необходимо выбрать домен. Есть два варианта (рис. 8.6): бесплатный домен от *wix.com* или ваш собственный домен.

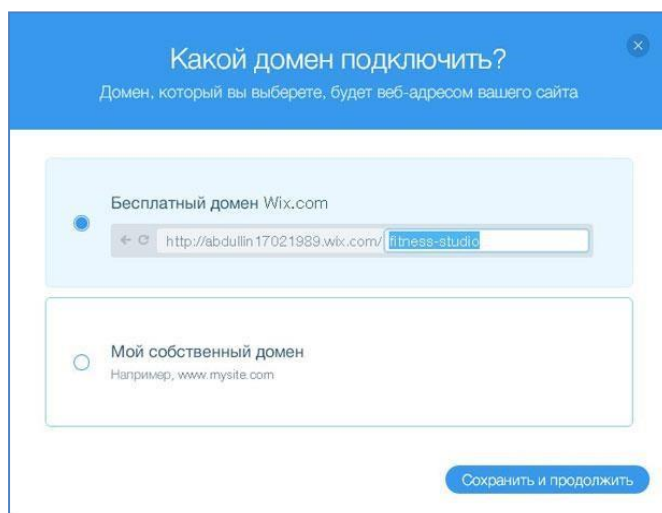


Рисунок 8.6– Выбор домена

Выберите бесплатный вариант. После этого появляется окно, в котором можно опубликовать сайт, то есть разместить его в сети Интернет (рис. 8.7).

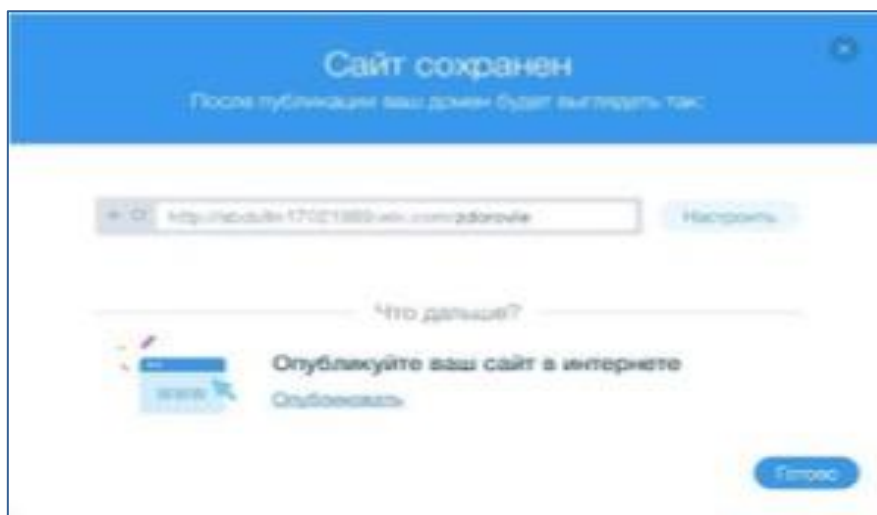


Рисунок 8.7– Публикация сайта

11. Если редактирование сайта еще не окончено, то опубликовать его можно и позже в пункте меню «Настройки сайта».

### ***Контрольные вопросы***

1. Что такое сайт?
2. Каковы цели создания сайта?
3. Какая информация обязательна для размещения на сайте?
4. Что такое домен?
5. Какие средства создания сайтов вы знаете?
6. Чем перечисленные средства создания сайтов отличаются друг от друга?
7. Перечислите возможности конструктора сайтов Wix.
8. Каковы особенности сайта переводчика?

### ***Задания для самостоятельной работы***

1. Найдите в сети Интернет сайты переводчиков. Изучите их особенности. По результатам анализа дополните свой сайт необходимой информацией.
2. Опубликуйте итоговую версию сайта и покажите ее преподавателю.

## **Лабораторная работа № 10** **Тема: СЕРВИСЫ РАБОТЫ СО ЗВУКОМ**

**Цель:** изучить особенности обработки аудиоинформации с помощью облачных сервисов.

### ***Краткие теоретические сведения***

Рассмотрим основные понятия, связанные с физикой звука, его оцифровкой и обработкой.

Физика звука такова, что он представляет собой волну, распространяющуюся в атмосфере (воздухе, воде или др. среде), и воспринимаемую человеком с помощью органов слуха. Кажущаяся сила звука для нас – это его громкость. Громкость измеряется в децибелах (дБ).

Звуковая волна непрерывно изменяется и характеризуется интенсивностью и частотой. Существует зависимость громкости и высоты тона звука от этих двух показателей (рис. 9.1).

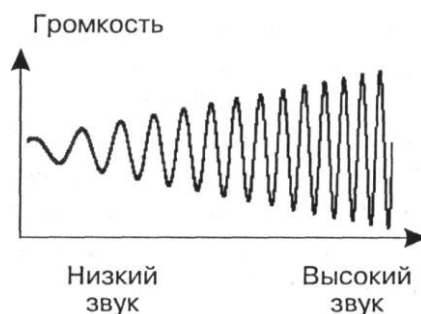


Рисунок 9.1. Зависимость громкости от интенсивности и частоты

С одной стороны, звук представляет собой непрерывную волну, а с другой – компьютер не может хранить и обрабатывать непрерывную форму. Именно поэтому для работы со звуковой информацией в компьютере необходимо звук преобразовать в дискретную форму – цифровую. Осуществляется это с помощью временной дискретизации. Непрерывная волна звука разбивается на отдельные временные участки, а для каждого такого участка устанавливается определенная величина интенсивности звука. В процессе получается, что непрерывная зависимость громкости звука от времени  $A(t)$  заменяется на дискретную последовательность уровней громкости (рис. 9.2).

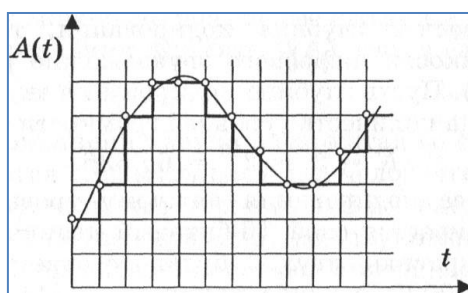


Рисунок 9.2– Дискретизация звуковой волны по времени

Для записи аналогового звука (непрерывного) и его преобразования в цифровую форму (дискретную) используется техническое средство – микрофон, подключенный к звуковой плате компьютера. При этом качество полученного цифрового звука зависит от количества измерений уровня его громкости в единицу времени, т.е. частоты дискретизации.

Можно выделить зависимость: чем большее количество измерений производится за 1 секунду, тем точнее «лесенка» цифрового звукового сигнала повторяет кривую диалогового сигнала.

Дадим определение. *Частота дискретизации звука – это количество измерений громкости звука за одну секунду.* Данный показатель изменяется в диапазоне от 8 000 до 48 000 измерений громкости звука за одну секунду. Получаем следующий термин – *глубина кодирования звука – это количество информации, которое необходимо для кодирования дискретных уровней громкости цифрового звука.* Следовательно, чем больше частота и глубина дискретизации звука, тем более качественным будет звучание оцифрованного звука.

Самое низкое качество оцифрованного звука у телефонной связи. Частота дискретизации телефонных разговоров 8000 раз в секунду, глубина дискретизации 8 битов, запись происходит на одной звуковой дорожке, т.е. в режиме «моно».

Самое высокое качество оцифрованного звука у аудиозаписей в формате CD. Частота дискретизации такого звука 48 000 раз в секунду, глубина дискретизации 16 битов, запись осуществляется на двух звуковых дорожках, т.е. в режиме «стерео».

С учетом таких значений можно вычислить примерное значение объема аудиофайла. И здесь мы видим зависимость: чем выше качество цифрового звука, тем больше информационный объем звукового файла.

Человечество всегда стремилось к обработке звука компьютером. И даже первые компьютеры имели небольшие динамики, издававшие ограниченный набор звуков. С расширением технических возможностей компьютера стало развиваться и программное обеспечение. Появились и первые музыкальные редакторы для создания звуковых файлов. В 1989 г. уже создали звуковые карты.

Современные компьютеры и устройства уже содержат целые звуковые подсистемы. Они представляют собой комплекс программно-аппаратных средств и предназначены для:

- записи звуковых сигналов, поступающих от внешних источников (микрофона или магнитофона);
- воспроизведения записанных ранее звуковых данных с
- помощью внешней акустической системы или головных телефонов (наушников);
- микширования (смешивания) при записи или воспроизведении сигналов от внешних источников;
- одновременной записи и воспроизведения звуковых сигналов;
- обработки звуковых сигналов: редактирования, объединения или разъединения фрагментов сигнала, фильтрации, изменения уровня и т.п.;
- управления панорамой стереофонического звукового сигнала и уровнем сигнала в каждом канале при записи и воспроизведении;
- обработки звукового сигнала в соответствии с алгоритмами объемного звучания;
- генерирования с помощью синтезатора звучания музыкальных инструментов через специальный интерфейс MIDI;

- воспроизведения звуковых компакт-дисков;
- управления компьютером и ввода текста с помощью микрофона.

Звуковые системы выполняются как в виде самостоятельных звуковых карт, устанавливаемых на материнской плате (интегрированные) (рис. 9.3), так и могут быть на другой карте расширения (рис. 9.4).

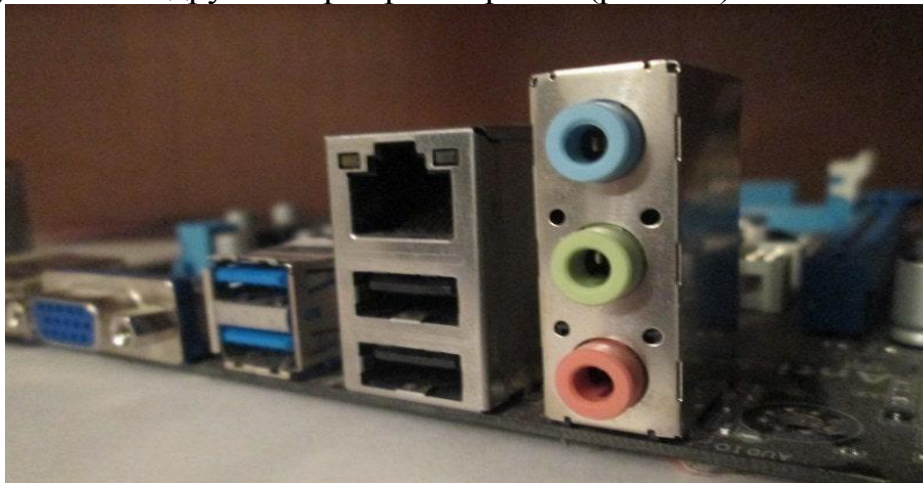


Рисунок 9.3. Интегрированная звуковая карта

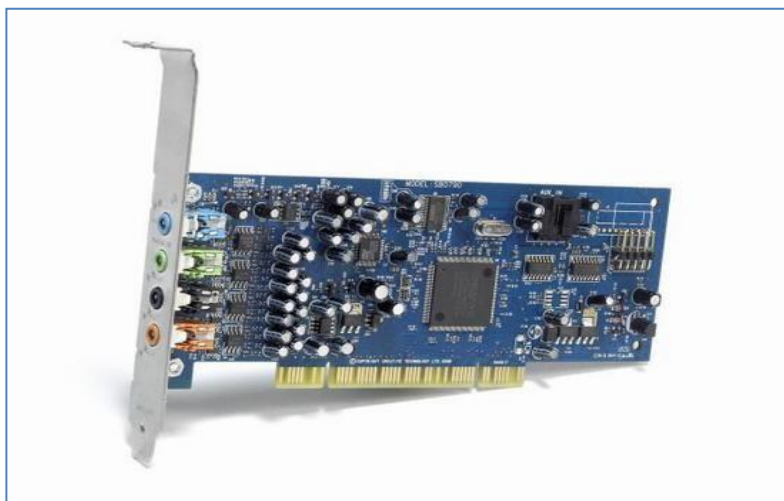


Рисунок 9.4. Звуковая карта на отдельной плате

Звуковые карты также могут быть и внешними (рис. 9.5).



Рисунок 9.5. Внешняя звуковая карта

## ***Ход работы***

***Внимание!*** Для работы понадобятся наушники и микрофон.

**Задание 1.** Перейдите на сайт <https://123apps.com/ru/> и изучите спектр программ, предоставляемых в режиме онлайн для работы со звуком (рис. 9.6):

Обрезать песню.

Аудиоконвертор. Соединить песни. Запись звука.

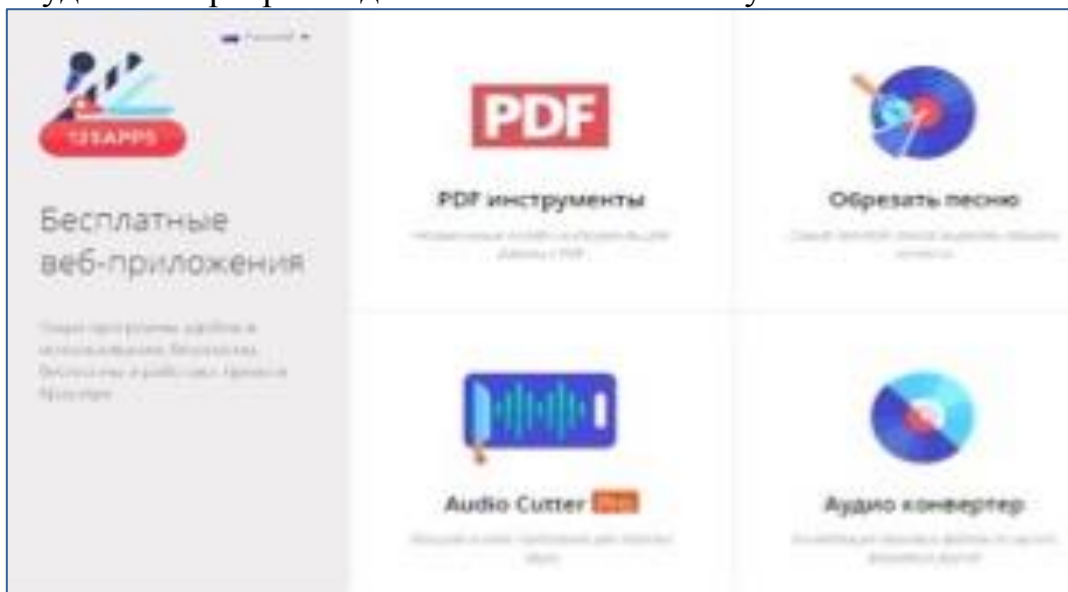


Рисунок 9.6. Главное окно сайта

**Задание 2.** Скачайте любую песню и, используя сервис «Обрезать песню», вырежьте из нее припев или любимую часть. Для этого нажмите по кнопке «Открыть файл» (рис. 9.7) и укажите место хранения песни.

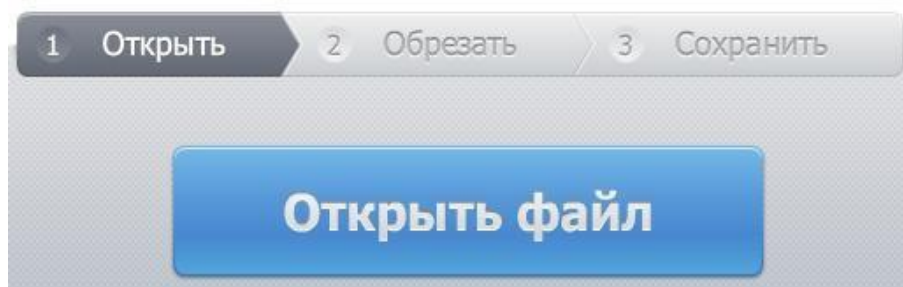


Рисунок 9.7. Кнопка «Открыть файл»

Когда файл станет доступен для редактирования, то звуковая дорожка окрасится в голубой цвет, а кнопка воспроизведения станет зеленого цвета. Два бегунка внизу дорожки указывают на начало и конец нужного отрывка композиции. Выделите необходимый отрезок звуковой дорожки и нажмите внизу на кнопку «Обрезать» (рис. 9.8).



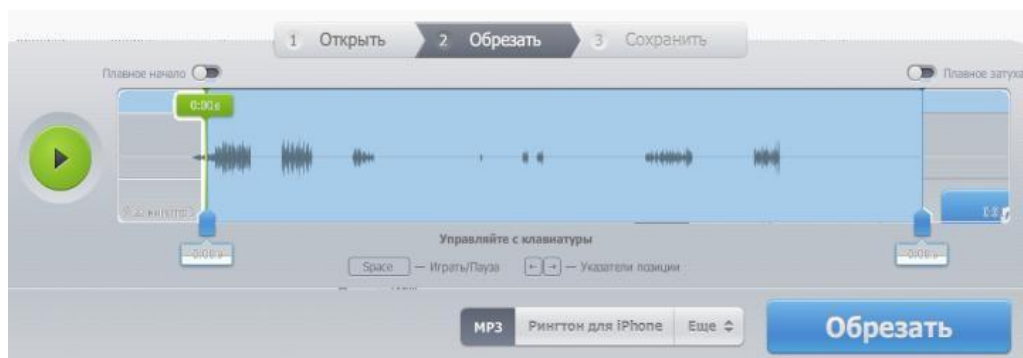


Рисунок 9.8. Обрезка звукового файла

Затем перейдите по ссылке «Скачать» и сохраните файл. Результат покажите преподавателю.

**Задание 3.** Используя сервис «Соединить песни», трижды соедините ранее вырезанный фрагмент в новую композицию.

Для этого добавьте трек три раза, воспользовавшись кнопкой «Добавить треки» (рис. 9.9).



Рисунок 9.9. Добавление трека

Настройте интервалы с помощью бегунков, перемещая их указателем мыши (рис. 9.10).



Рисунок 9.10. Настройка интервалов

Изучите значки до и после композиции:

- Кроссфейд: громкость первого трека плавно уменьшается, в то время как громкость следующего плавно нарастает.
  - Фейдин: плавное нарастание громкости первого трека.
  - Фейдаут: плавное затухание громкости последнего трека.
- По окончании нажмите на кнопку «Соединить».

Программе потребуется некоторое время, чтобы склеить песни между собой. После того как склейка завершена, можно нажать на кнопку «Скачать» для сохранения полученной композиции у себя в компьютере. Результат покажите преподавателю.

**Задание 4.** Изучите возможности сервиса «Запись звука». Запишите несколько фраз, нажав на кнопку (рис. 9.11).

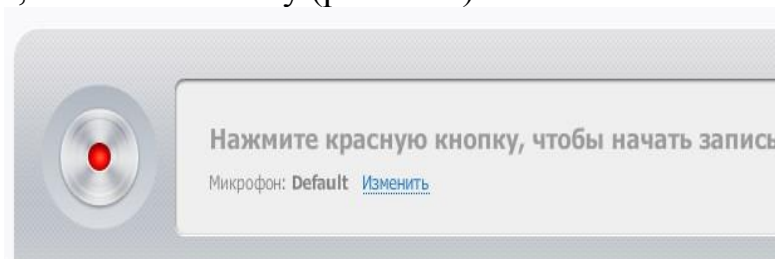


Рисунок 9.11. Сервис записи звука

По окончании записи нажмите на зеленую кнопку. Затем нажмите на кнопку «Сохранить» и скачайте полученный файл.

Результат покажите преподавателю.

**Задание 5.** Изучите возможности бесплатного сервиса «Аудиоконвертер». Для этого конвертируйте файл, полученный в задании 3 в различные форматы: из *mp3* в *wav*, из *mp3* в *ogg* (рис. 9.12).

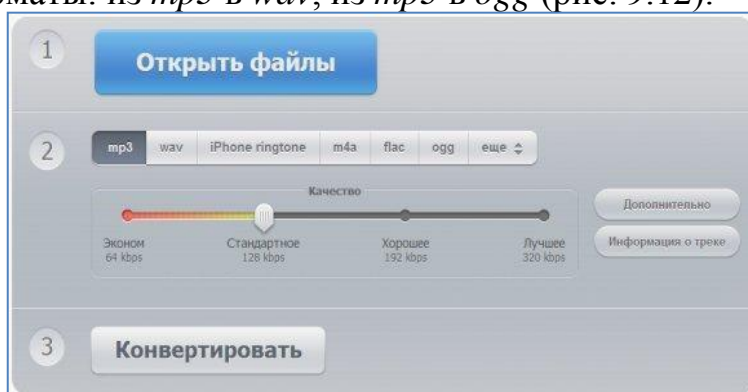


Рисунок 9.12. Конвертирование файлов

Нажмите на кнопку «Открыть файлы». В появившемся окне выберите файл, который будете конвертировать.

Пока файл загружается, выберите формат, в котором вы хотите получить итоговую композицию.

С помощью бегунка в разделе «Качество» установите необходимые настройки, установите настройки в разделе «Дополнительно».

Нажмите кнопку «Конвертировать». Когда файл будет обработан, нажмите кнопку «Скачать».

### ***Контрольные вопросы***

1. Что такое звук с точки зрения физики?
2. Чем характеризуется звуковая волна?



3. Может ли компьютер хранить непрерывную звуковую волну?
4. Как осуществляется процесс дискретизации звуковой волны?
5. Что такое частота дискретизации звука?
6. Что такое глубина кодирования звука?
7. Что такое звуковая подсистема компьютера?
8. Опишите классификацию звуковых карт.

***Задания для самостоятельной работы***

1. Изучите основные характеристики форматов звуковых файлов (например, *mp3*, *wav*, *ogg*, *flac* и др.).
2. Составьте сравнительную таблицу характеристик основных форматов звуковых файлов.

**Комплект тестовых вопросов и заданий**  
**по дисциплине: «Информационные технологии в лингвистике»**

**Тесты 1**

Тесты к разделу 2

1. Схема перевода как коммуникативного акта предусматривает
  1. отправителя и получателя информации
  2. отправителя, получателя, канал связи.
  3. наличие перекодирующего звена
  4. канал связи
  
2. Закономерные соответствия подразделяются на
  1. эквиваленты, вариантные соответствия и адекватные замены
  2. адекватные замены и аналоги
  3. аналоги и вариантные соответствия
  4. эквиваленты и синонимичные замены
  
3. Функциональная модель перевода исходит из того, что
  1. главное в переводе – обеспечить заданную ИТ функцию
  2. сделать перевод функциональным
  3. текст представляет собой вербализованную часть социокультуры
  
4. Эквиваленты В основе классификации текстов А.Федорова лежат различия
  1. в системах языков
  2. жанрово-стилистической принадлежности текстов
  3. между языком и речью
  4. функциональных особенностей текстов в переводе – это
  5. все словарные соответствия
  6. все термины
  7. постоянные равнозначные соответствия, не зависящие от контекста
  8. соответствия, эквивалентные во всех значениях
  
5. Концепция порождающей грамматики принадлежит
  1. В. Коллеру
  2. Н. Хомскому
  3. Ю. Найде
  4. Дж. Кэтфорду
  
6. Ядерные предложения при переводе
  1. преобразуются в поверхностные структуры
  2. трансформируются в простейшие высказывания
  3. элиминируются
  4. определяют доминанту информативного запаса

7. Что общего между папирусом, берестяной грамотой, книгой и дискетой?

1. материал, из которого они изготовлены
2. размер
3. хранение информации
4. стоимость

8. За единицу измерения информации принят...

1. 1 бод
2. 1 бит
3. 1 байт
4. 1 Кбайт

9. Какое количество цифр используется в шестнадцатеричной системе счисления для представления чисел?

1. десять
2. восемь
3. шестнадцать
4. две
1. 5

10. В каком устройстве компьютера производится обработка информации?

1. внешняя память
2. процессор
3. дисплей
4. клавиатура

11. Какое устройство компьютера может оказывать вредное воздействие на человека?

1. гибкий диск
2. дисплей
3. системный блок
4. клавиатура

12. Программы обслуживания устройств компьютера называются:

1. загрузчиками;
2. драйверами;
3. трансляторами;
4. интерпретаторами;
5. компиляторами.

13. Целью информационной технологии является:

1. обработка информации
2. выпуск информации

3. принятие информации
4. производство информации
5. воспроизведение информации

14. К свойствам алгоритма относятся следующие (укажите наиболее правильный ответ):

1. дискретность, массовость, абсолютность
2. результативность, дискретность, массовость,
3. дискретность, массовость, абсолютность, информативность, формальность
4. массовость, дискретность, результативность, точность, эффективность

#### Тесты к разделу 4

15. Для того чтобы подготовиться к копированию слова, необходимо...

1. выделить слово
2. установить на слово курсор
3. выделить абзац
4. установить курсор на начало абзаца

16. Гипертекст – это:

1. способ организации текстовой информации, внутри которой установлены смысловые связи между ее различными фрагментами;
2. обычный, но очень большой по объему текст;
3. текст, буквы которого набраны шрифтом большого размера;
4. распределенная совокупность баз данных, содержащих тексты.

1. Имена файлам, в которых хранятся на диске созданные документы (тексты или рисунки), задаются...

1. автоматически программой (текстом или графическим редактором)
2. создателем документа
3. операционной системой
4. документы не имеют имен

18. Электронные таблицы – это:

1. таблицы электронных схем компьютера;
2. программное средство для автоматизации вычислений;
3. программное средство для автоматизации хранения и поиска информации.

19. Какие из перечисленных программ являются электронными таблицами:

1. Лексикон;
2. Supercalc;

3. Windows;
4. MsWord;
5. Excel;
6. FoxPro.

20. Электронные таблицы позволяют выполнять следующие операции:

1. вычислять значения арифметических выражений;
2. осуществлять циклические вычисления;
3. вычислять производные;
4. вычислять интегралы;
5. строить подвижные изображения.

21. База данных – это:

1. специальным образом организованная и хранящаяся на внешнем носителе совокупность взаимосвязанных данных о некотором объекте;
2. совокупность программ для хранения и обработки больших массивов информации;
3. интерфейс, поддерживающий наполнение и манипулирование данными;
4. определенная совокупность информации.

22. Области применения национального корпуса русского языка:

1. статистическая обработка,
2. фоносемантический анализ,
3. морфологический анализ.

23. Сколько слов будет найдено (выделено, указано) в процессе автоматического поиска в тексте: “Далеко за отмелью, в ельнике, раздалась птичья трель”, если в качестве образца задать слово “ель”:

1. 1 раз;
2. 0 раз;
3. 3 раза;
4. 2 раза.

24. Объединение сетей и компьютеров, расположенных на больших расстояниях в единую систему, называют сетью следующего вида:

1. локальной;
2. корпоративной;
3. региональной;
4. глобальной.

25. FTP-серверы Интернета обеспечивают:

1. работу телеконференций;
2. отправление электронных сообщений;

3. прием и передачу файлов;
4. размещение сайтов.

26. Основным объектом для хранения информации в реляционных базах данных является:

1. таблица;
2. запрос;
3. форма;
4. отчет.

28. Строка, описывающая свойства элемента таблицы базы данных, называется:

1. полем;
2. бланком;
3. записью;
4. ключом.

### **Критерии оценивания тестирования**

При тестировании все верные ответы берутся за 100%.

90%-100% отлично

75%-90% хорошо

60%-75% удовлетворительно

менее 60% неудовлетворительно

## *Тесты 2*

1. Информация – это:

- а) сведения, передаваемые от одного человека к другому устно, письменно, посредством каких-либо условных сигналов или с использованием каких-либо технических средств;
- б) сообщение, уменьшающее неопределенность в той области, к которой оно относится;
- в) данные, переработанные особым образом;
- г) объем данных, предназначенных для обработки пользователем;
- д) весь окружающий мир.

2. Данные – это:

- а) представление фактов и идей в формализованном виде, пригодном для передачи и обработки в некотором информационном процессе;
- б) сведения, передаваемые от одного человека к другому устно, письменно, посредством каких-либо условных сигналов или с использованием каких-либо технических средств;
- в) сообщение, уменьшающее неопределенность в той области, к которой оно относится;
- г) информация, переработанная особым образом;
- д) объем знаний, предназначенных для обработки пользователем.

3. Знания – это:

- а) закономерности источника информации (понятия, сведения, принципы, связи, законы), полученные или приобретенные в результате обучения, практической деятельности;
- б) представление фактов и идей в формализованном виде, пригодном для передачи и обработки в некотором информационном процессе;
- в) сведения, передаваемые от одного человека к другому устно, письменно, посредством каких-либо условных сигналов или с использованием каких-либо технических средств;
- г) сообщение, уменьшающее неопределенность в той области, к которой оно относится;
- д) информация, переработанная особым образом.

4. Технология – это:

- а) совокупность знаний о способах и средствах осуществления процессов, при которых происходит качественное изменение объекта;
- б) методы обработки продукции;
- в) совокупность информации о способах и средствах осуществления процессов, обработки материалов, приводящей к изменениям окружающей среды;

d) совокупность данных и фактов о способах и средствах осуществления процессов, при которых происходит качественное изменение объекта.

5. Установите соответствие между определением информационных технологий:

В узком смысле слова	Совокупность средств и методов сбора, обработки и передачи данных (первичной информации) для получения информации нового качества о состоянии объекта или явления (информационного продукта)
В широком смысле слова	Широкий класс дисциплин и областей деятельности, относящихся к технологиям управления и обработки данных, в том числе с применением вычислительной техники

6. К мировому рынку информационных технологий относятся:

A. Компьютерное и офисное аппаратное обеспечение; программное обеспечение;

B. ИТ-услуги.

C. Компании, оказывающие ИТ-услуги.

D. Вся компьютерная техника и персонал, обслуживающий ее.

7. Информационные технологии (ИТ) делятся на:

a) глобальные, базовые и конкретные ИТ;

b) используемые, неиспользуемые, разрабатываемые ИТ;

c) общедоступные, частные, образовательные ИТ;

d) бесплатные и платные ИТ.

8. Проблема использования ИТ:

a) устаревают и заменяются новыми;

b) сначала дорогие в разработке, а потом становятся дешевле;

c) разрабатываются либо только для настольных компьютеров, либо только для мобильных устройств;

d) подвергаются взлому и заражению вирусами.

9. Лингвистика – это:

a) наука о естественном человеческом языке;

b) наука об обработке информации;

c) практическая наука о строения языка (выделение в нем фонетического, лексического, грамматического уровня и уровня текста);

d) наука о принципах функционирования языка в обществах разных типов.

10. Установите соответствие между определениями:

Теоретическая лингвистика	Область языкознания, направленная на объективное установление состояния отдельного языка, его истории и закономерностей
Прикладная лингвистика	Область языкознания, связанная с разработкой методов решения практических задач использования языка



11. Информационные технологии в лингвистике – это:

- a) совокупность законов, методов и средств получения, хранения, передачи, распространения, преобразования информации о языке и законах его функционирования с помощью компьютеров;
- b) процесс обучения иностранному языку, методика его преподавания, особенности описания грамматики для учебных целей с использованием ИТ;
- c) компьютерные технологии и автоматические системы обработки информации для языкознания;
- d) область языкознания, связанная с разработкой методов решения практических задач использования языка на компьютере.

12. Язык – это:

- a) знаковая система, используемая для общения в некотором социуме;
- b) системы обозначений объектов окружающего мира;
- c) знаковая система, используемая для общения в некотором социуме, появившаяся естественным путем;
- d) все искусственные системы обозначений объектов окружающего мира.

13. Установите соответствие между видом языка и примером:

Естественный язык	Английский
Искусственный неспециализированный язык	Эсперанто
Искусственный специализированный язык	Ноты

14. Установите соответствие между понятием и определением:

Автоматический анализ звучащей речи	Речь преобразуется в печатный текст, над которым можно производить дальнейшие операции
Автоматический синтез звучащей речи	Процесс преобразования печатного текста, существующего в цифровой форме, в звучащий текст на естественном человеческом языке

15. Установите правильную последовательность этапов автоматического анализа речи:

- a) ввод звучащей речи в компьютер с помощью микрофона;
- b) выделение компьютерной программой в звуковом потоке отдельных знаков;
- c) идентификация выделенных знаков звучащей речи со знаками языка.

16. Что такое звук?

- a) волна, распространяющаяся в атмосфере и воспринимаемая человеком с помощью органов слуха;
- b) волна, отраженная от предметов и воспринимаемая человеком с помощью органов слуха;
- c) кривая, задающая зависимость амплитуды от времени;
- d) волна с меняющейся интенсивностью и частотой.

17. Частота дискретизации звука – это:

- a) количество измерений громкости звука за одну секунду;
- b) количество измерений интенсивности звука за одну секунду;
- c) количество измерений частоты звука за одну секунду;
- d) количество измерений амплитуды звука за одну секунду.

18. Глубина кодирования звука – это количество информации, которое необходимо для кодирования дискретных уровней громкости цифрового звука:

- a) количество информации, которое необходимо для кодирования дискретных уровней громкости цифрового звука;
- b) объем информации, который необходим для кодирования дискретных уровней громкости цифрового звука;
- c) количество информации, которое необходимо для кодирования дискретных уровней амплитуды цифрового звука;
- d) объем информации, который необходим для кодирования дискретных уровней амплитуды цифрового звука.

19. Цель создания звуков:

- a) имитация различных естественных звуков, получение принципиально новых звуков;
- b) получение новых звуков из уже существующих;
- c) придание звукам дополнительных качеств или устранение существующих.

20. Цель обработки звуков:

- a) получение новых звуков из уже существующих;
- b) имитация различных естественных звуков;
- c) получение принципиально новых звуков.

21. Сжатие звука без потерь:

- a) в ходе такого сжатия звук, записанный в цифровом виде, сохраняется полностью, без потерь;
- b) в ходе такого сжатия звук, записанный в цифровом виде, сохраняется почти полностью;
- c) в ходе такого сжатия звук, записанный в аналоговом виде, сохраняется полностью, без потерь;
- d) в ходе такого сжатия звук, записанный в аналоговом виде, сохраняется почти полностью.

22. Установите соответствие между методами автоматического синтеза речи и их пояснением:

Метод компилятивного синтеза	Используется для синтеза относительно небольшого и заранее известного набора фраз; обеспечивает высокое качество звучания
------------------------------	---

Форматноголосовой метод	Более сложный метод, требует точного моделирования особенностей речевого тракта человека, а также учет интонационной модуляции речи; дает относительно низкую точность синтезируемых звуков речи
-------------------------	--

23. Сканер – это:

- а) устройство ввода, работающее по принципу фотоаппарата, т.е. позволяющее компьютеру «увидеть» текст в виде фотографии;
- б) программа автоматического распознавания символов;
- с) устройство вывода, работающее по принципу фотоаппарата, т.е. позволяющее компьютеру «увидеть» текст в виде фотографии.

24. Программа автоматического распознавания текста (OCR-программа) – это:

- а) компьютерная программа, позволяющая преобразовать текст с бумажного носителя в электронный текстовый файл, который в дальнейшем может обрабатываться человеком в любом текстовом редакторе;
- б) компьютерная программа, позволяющая преобразовать речь диктора в электронный текстовый файл, который в дальнейшем может обрабатываться человеком в любом текстовом редакторе;
- с) это компьютерная программа, позволяющая преобразовать фото с бумажного носителя в электронный текстовый файл, который в дальнейшем может обрабатываться человеком в любом графическом редакторе.

25. Реферат определяется как связный текст, который кратко выражает:

- а) центральную тему, предмет, цель, методы, результаты исследования;
- б) центральную тему, цель, объект, предмет исследования;
- с) введение, краткое содержание и заключение;
- д) выводы конкретного исследования с постановкой цели и методами исследования.

26. Установите соответствие между видами рефератов и их пояснением:

Связный текст	Новое текстовое образование, порождаемое на основе логико-смыслового анализа исходного текста
Реферат-клише	Модификация заданной клишированной структуры, пустые ячейки которой заполняются после анализа заданного текста
Квализреферат	Перечень наиболее информативных предложений текста

27. Установите правильную последовательность этапов реферирования текстов:

- а) подготовительный: определение темы текста, его понимание;
- б) аналитический: деление текста на фрагменты (абзацы и т.п.) и выделение в каждом фрагменте главных смысловых слов;

с) непосредственное составление реферата или аннотации:  
соединение выделенных смысловых единиц в связный текст.

28. Ключевое слово – это:

а) знаменательное слово, относящееся к основному содержанию текста и повторяющееся в нем несколько раз;

б) профилирующее слово, относящееся к основному содержанию текста и повторяющееся в нем несколько раз;

с) знаменательное слово, относящееся к основному содержанию текста и несущее важную смысловую нагрузку;

д) профилирующее слово, относящееся к основному содержанию текста и несущее важную смысловую нагрузку.

29. Установите соответствие между методами автоматического реферирования и аннотирования текстов:

Статистические	Принадлежность слова к категории ключевых определяется его статистическими характеристиками, например, ключевое слово согласно этому методу встречается среди знаменательных слов текста наибольшее количество раз
Позиционные	Отнесение предложения к ключевым опирается на его местонахождение в тексте, например, ключевые предложения входят в заголовок, подзаголовок, находятся в начале и конце текста
Логикосемантические	Учитывается структура и семантика текста, выделяются предложения с наибольшим функциональным весом

30. Что такое облачные вычисления?

а) программно-аппаратное обеспечение, доступное пользователю через Интернет или локальную сеть в виде сервиса, позволяющего использовать удобный интерфейс для удаленного доступа к выделенным ресурсам;

б) все программы и ресурсы, доступные через глобальную сеть Интернет;

с) динамически масштабируемый способ доступа к внешним вычислительным ресурсам в виде сайта;

д) все порталы, предоставляющие услуги пользователя сети Интернет.

31. Какие решения НЕ относятся к облачным вычислениям?

а) сетевые операционные системы;

б) *blade*-сервера;

с) система хранения данных;

д) сети хранения данных;

е) технология виртуализации.

32. Виртуальная машина – это:

а) программная или аппаратная среда, которая скрывает настоящую реализацию какого-либо процесса или объекта от его видимого представления;

б) программа, установленная на компьютере, позволяющая запускать приложения для другой операционной системы;

с) аппаратная среда, установленная на компьютере, позволяющая запускать приложения для другой операционной системы;

д) программная среда, которая скрывает настоящую реализацию какого-либо процесса или объекта от его видимого представления.

33. Установите соответствие между видами облачных вычислений и их определением:

Инфраструктура как сервис	Предоставление компьютерной инфраструктуры как услуги на основе концепции облачных вычислений
Платформа как сервис	Предоставление интегрированной платформы для разработки, тестирования, развертывания и поддержки веб-приложений как услуги
Программное обеспечение как сервис	Модель развертывания приложения, которая подразумевает предоставление приложения конечному пользователю как услуги по требованию

## **Вопросы для коллоквиума**

### *Раздел 1. Основные вопросы информационных технологий*

1. Что такое информация?
2. Что такое данные?
3. Что такое знания?
4. Что такое технология?
5. Что такое информационные технологии в широком и узком смысле?
6. Что включает в себя мировой ИТ-рынок?
7. Виды ИТ.
8. Проблемы использования информационных технологий.

### *Раздел 2. Основные вопросы лингвистики*

1. Что такое лингвистика?
2. Что такое теоретическая лингвистика?
3. Что такое прикладная лингвистика?
4. Какова роль информационных технологий в лингвистике?
5. Что такое язык?
6. Виды языков, примеры.

### *Раздел 3. Автоматический анализ и синтез речи*

1. Что такое автоматический анализ звучащей речи?
2. Что такое автоматический синтез звучащей речи?
3. Этапы анализа речи.
4. Что такое звук?
5. Что такое частота дискретизации?
6. Что такое глубина кодирования звука?
7. Цели создания звука.
8. Цели обработки звука.
9. Что такое сжатие без потерь?
10. Что такое сжатие с потерями?
11. Методы автоматического синтеза речи.

### *Раздел 4. Автоматический анализ текстов*

1. Что такое сканер?
2. Программы автоматического распознавания символов.
3. Что такое реферат, что он выражает?
4. Виды рефератов.
5. Этапы реферирования текстов.
6. Что такое ключевое слово, ключевое словосочетание, ключевое предложение?
7. Методы автоматического реферирования текстов.

*Раздел 5. Облачные вычисления*

1. Что такое облачные вычисления?
2. Современные инфраструктурные решения для облачных вычислений.
3. Преимущества *Blade*-серверов.
4. Что такое система хранения данных, каковы ее преимущества?
5. Что такое сети хранения данных?
6. Что такое технология виртуализации?
7. Виды облачных вычислений.

АЛИЕВА Аза Башировна  
КОЧКАРОВ Расул Сагитович

## **ИНФОРМАЦИОННЫЕ ТЕХНОЛОГИИ В ЛИНГВИСТИКЕ**

Лабораторный практикум для обучающихся  
по направлению подготовки 45.03.02 Лингвистика

Корректор Чагова О.Х.  
Редактор Чагова О.Х.

Сдано в набор 19.08.2024 г.  
Формат 60x84/16  
Бумага офсетная  
Печать офсетная  
Усл. печ. л. 3,72  
Заказ № 4945  
Тираж 100 экз.

Оригинал-макет подготовлен  
в Библиотечно-издательском центре СКГА  
369000, г. Черкесск, ул. Ставропольская, 36