



Министерство науки и высшего образования Российской Федерации
Федеральное государственное бюджетное образовательное учреждение
высшего образования
«Московский государственный технический университет
имени Н.Э. Баумана
(национальный исследовательский университет)»
(МГТУ им. Н.Э. Баумана)

ФАКУЛЬТЕТ Информатика и системы управления

КАФЕДРА Программное обеспечение ЭВМ и информационные технологии

РАСЧЕТНО-ПОЯСНИТЕЛЬНАЯ ЗАПИСКА
К ВЫПУСКНОЙ КВАЛИФИКАЦИОННОЙ РАБОТЕ
НА ТЕМУ:

***Разработка алгоритма для синтеза речи по заранее
неизвестному тексту на русском языке***

Студент ИУ7-85Б
(Группа)

(Подпись, дата) А. П. Овчинникова
(И.О.Фамилия)

Руководитель ВКР

(Подпись, дата) Т. Н. Романова
(И.О.Фамилия)

Консультант

(Подпись, дата) _____
(И.О.Фамилия)

Консультант

(Подпись, дата) _____
(И.О.Фамилия)

Нормоконтролер

(Подпись, дата) _____
(И.О.Фамилия)

2021 г.

РЕФЕРАТ

Дипломная работа на тему: Разработка алгоритма для синтеза речи по заранее неизвестному тексту на русском языке.

Работа состоит из введения, четырех разделов, заключения, списка использованной литературы и двух приложений.

Алгоритмы синтеза речи по тексту являются объектом и предметом исследования в данной работе. Целью работы является разработка алгоритма для синтеза речи по произвольному тексту на русском языке и создание ПО для озвучивания текста на основе этого алгоритма.

Во введении раскрывается актуальность выбранной темы, описывается цель и задачи работы.

В аналитическом разделе рассмотрены существующие системы преобразования текста в речь; проведен анализ существующих методов синтеза речи и выбран наиболее подходящий; выбраны исходные элементы для синтеза и описана модификация выбранного метода.

В конструкторском разделе приведены требования к программе; описаны входные, выходные данные и ограничения; приведена функциональная модель разработанного ПО; описано создание акустической базы данных.

В технологическом разделе описан используемый стек технологий; приведена структура хранилищ данных; описаны зависимости и структура разработанного программного обеспечения; описан процесс тестирования разработанного ПО; приведено руководство пользователя.

В исследовательском разделе проведена оценка качества синтезируемой с помощью разработанного ПО речи и проведено сравнение исходного и модифицированного метода.

В заключении обобщается проделанная работа и формулируются выводы по недостаткам и достоинствам разработанного программного обеспечения.

Расчетно-пояснительная записка содержит 73 страницы, 28 рисунков, 20 таблиц, 22 источника, 2 приложения.

СОДЕРЖАНИЕ

РЕФЕРАТ	2
ОБОЗНАЧЕНИЯ И СОКРАЩЕНИЯ	5
ВВЕДЕНИЕ	6
1 Аналитический раздел	8
1.1 Задача преобразования текста в речь	8
1.2 Обзор существующих систем преобразования текста в речь	8
1.2.1 Amazon Polly	8
1.2.2 IBM Watson Text to Speech	9
1.2.3 Azure Text to Speech API	10
1.2.4 Google Cloud Text-to-Speech	10
1.2.5 Обоснование необходимости разработки нового продукта	11
1.3 Методы синтеза речи	12
1.3.1 Параметрический синтез	12
1.3.2 Компилятивный синтез	12
1.3.3 Синтез речи по фонетическим правилам	13
1.3.4 Выбор метода синтеза речи	14
1.4 Выбор исходных элементов при создании базы данных для синтеза речи по правилам	15
1.5 Мини- и макси-наборы аллофонов для синтеза русской речи	17
1.6 Модификация метода аллофонного синтеза по фонетическим правилам	18
2 Конструкторский раздел	22
2.1 Требования к программе	22
2.2 Описание входных и выходных данных и ограничений	22
2.3 Система синтеза речи по тексту	25
2.3.1 Лингвистический текстовый процессор	25
2.3.2 Просодический процессор	34

2.3.3 Фонетический процессор	36
2.3.4 Акустический процессор	44
2.4 Создание акустической базы данных	46
3 Технологический раздел	49
3.1 Описание используемого стека технологий	49
3.2 Структура хранения данных	49
3.3 Описание зависимостей	51
3.4 Структура разработанного программного обеспечения	51
3.5 Тестирование	56
3.6 Сборка и запуск разработанного программного обеспечения	58
3.7 Руководство пользователя	59
3.8 Руководство пользователя	59
4 Исследовательский раздел	62
4.3 Оценка качества синтеза речи	62
4.4 Проведение исследования	64
4.5 Вывод	68
ЗАКЛЮЧЕНИЕ	69
СПИСОК ИСПОЛЬЗОВАННЫХ ИСТОЧНИКОВ	71
ПРИЛОЖЕНИЕ А	74
ПРИЛОЖЕНИЕ Б	88

ОБОЗНАЧЕНИЯ И СОКРАЩЕНИЯ

БД – База данных

ИК – Интонационная конструкция

ПО – Программное обеспечение

СРЛЯ – Современный русский литературный язык

TTS – Text to speech system, система преобразования текста в речь

UML – Unified Modeling Language, унифицированный язык моделирования

ВВЕДЕНИЕ

Речевое общение является доминирующим способом человеческих социальных связей и обмена информацией. Эта тенденция находит отражение и во взаимодействии человек-машина. В современном мире появилась необходимость создания систем разговорного языка. Такие системы сочетают в себе как распознавание речи, так и ее синтез. Создание подобных жизнеспособных систем уже давно привлекает внимание ученых и инженеров во всем мире. Успешное развитие этой технологии повысит доступность компьютеров и автоматизированных систем для широкого круга пользователей [13].

Под автоматическим синтезом речи понимается технология, позволяющая преобразовать текстовую информацию в звучащую речь. Первые синтезаторы речи, механические, появились еще в 18 веке. Электрические синтезаторы стали создаваться в 19 веке после того, как была разработана резонаторная теория Гельмгольца. С тех пор в технологии синтеза речи прошло три поколения, и современные синтезаторы третьего поколения позволяют получать синтезированную речь, наиболее приближенную к естественной.

В настоящее время технология автоматического синтеза находит широкое применение в таких отраслях, как телекоммуникации, мобильные устройства, автомобильная индустрия, компьютеризованные системы, образовательные системы и многих других. Голосовые роботы берут на себя большую часть рутинных задач, например, в колл-центрах. Использование систем синтеза речи может повысить качество жизни слабослышащих, глухих, немых людей.

Подавляющее большинство систем разговорного языка создаются как универсальные, поддерживающие широкий набор языков. Однако стремление к универсальности может нанести серьезный ущерб качеству синтезируемой речи, потому как каждая языковая система обладает уникальными чертами, которые стираются при попытке создать универсальную систему для всех (или многих) языков. Для создания качественного синтезатора речи необходимо, помимо

знаний в программировании и в обработке звука на компьютере, необходимо также обратиться к таким областям лингвистики, как фонетика и просодика.

Цель данной работы – разработка алгоритма для синтеза речи по произвольному тексту на русском языке и создание ПО для озвучивания текста на основе этого алгоритма – призвана решить обозначенную проблему.

Для достижения поставленной цели необходимо решить следующие задачи:

- провести аналитическое исследование существующих подходов к синтезу речи и выбрать наиболее подходящий;
- на основе проведенного исследования модифицировать выбранный метод для достижения поставленной цели;
- разработать алгоритм по модифицированному методу;
- создать программное обеспечение, реализующее данный алгоритм;
- разработать тесты для разработанного ПО;
- провести исследование разработанного ПО для оценки качества синтезируемой речи.

1 Аналитический раздел

В данном разделе рассматриваются существующие системы преобразования текста в речь (Text To Tpeech Systems), проводится обзор методов синтеза речи, а также дается краткое описание фонетических особенностей русского языка.

1.1 Задача преобразования текста в речь

Программы синтеза речи преобразуют входные текстовые сообщения в устный вывод, автоматически генерируя синтетическую речь. Синтез речи часто называют преобразованием текста в речь (TTS). Под системами автоматического синтеза речи понимают системы, преобразующие письменный орфографический текст и другую информацию в звучащую речь [8].

1.2 Обзор существующих систем преобразования текста в речь

Многие крупные компании, такие как Microsoft, Amazon и другие активно занимаются разработкой собственных text-to-speech систем, позволяющих преобразовать в речь неизвестные заранее входные данные. Каждая такая система обладает своими достоинствами и недостатками.

1.2.1 Amazon Polly

Amazon Polly - это облачный сервис Amazon Web Services, дочерней компании Amazon.com, который преобразует текст в речь, используя технологии глубокого обучения [11]. Сервис позволяет работать с различными языками и включает в себя несколько голосов для каждого языка. Основные достоинства и недостатки Amazon Polly приведены в таблице 1.

Таблица 1 – Сервис Amazon Polly

Достоинства	Недостатки
<ul style="list-style-type: none"> • Поддержка различных языков, в том числе русского. • Поддержка одного или более голосов для одного языка. • Предоставляет возможность управления такими аспектами речи, как произношение, громкость, высота голоса, скорость и т. д. • Подробная документация. • Наличие обучающих пособий от разработчика. • Бесплатная ограниченная версия 	<ul style="list-style-type: none"> • Является платным сервисом, стоимость использования зависит от количества символов в тексте. • Нельзя заранее проверить качество синтезированной речи на необходимом тексте. Можно только прослушать заранее заготовленные примеры на сайте. • Сервис имеет закрытый исходный код. • Необходимо обучение для использования продукта. • Обучающие пособия и документация на английском языке.

1.2.2 IBM Watson Text to Speech

IBM Watson Text to Speech – облачный сервис, который позволяет преобразовывать письменный текст в естественную речь на разных языках и разных голосах [22]. Сервис основан на технологии нейронного голоса. Основные достоинства и недостатки IBM Watson Text to Speech приведены в таблице 2.

Таблица 2 – Сервис IBM Watson Text to Speech

Достоинства	Недостатки
<ul style="list-style-type: none"> • Поддержка различных языков. • Поддержка различных голосов. Помимо стандартных голосов есть возможность создать уникальный пользовательский голос. • Бесплатная ограниченная версия. 	<ul style="list-style-type: none"> • Нет поддержки русского языка. • Является платным сервисом, стоимость использования зависит от количества символов в тексте. • Нельзя заранее проверить качество синтезированной речи на необходимом тексте. Можно только прослушать заранее заготовленные примеры на сайте. • Сервис имеет закрытый исходный код. • Необходимо обучение для использования продукта. • Обучающие пособия и документация на английском языке.

1.2.3 Azure Text to Speech API

Azure Text to Speech API – служба, позволяющая приложениям, средствам или устройствам преобразовывать текст с помощью языка разметки синтеза речи (SSML) [14]. Можно использовать речевой пакет SDK или REST API. Основные достоинства и недостатки Azure Text to Speech API приведены в таблице 3.

Таблица 3 – Служба Azure Text to Speech API

Достоинства	Недостатки
<ul style="list-style-type: none">• Поддержка различных языков, в том числе русского.• Поддержка различных голосов. Помимо стандартных голосов есть возможность создать уникальный пользовательский голос.• Подробная документация на русском языке.• Бесплатная пробная подписка.• Можно заранее проверить качество синтезированной речи на необходимом тексте.	<ul style="list-style-type: none">• Является платной службой, стоимость использования зависит от количества символов в тексте.• Имеет закрытый исходный код.• Необходимо обучение для использования продукта.

1.2.4 Google Cloud Text-to-Speech

Google Cloud Text-to-Speech позволяет разработчикам синтезировать естественную речь с 30 голосами, доступными на нескольких языках и в разных вариантах [12]. Google Cloud Text-to-Speech основан на исследованиях DeepMind. Сервис позволяет преобразовывать текст с помощью языка разметки синтеза речи (SSML). Можно использовать REST API. Основные достоинства и недостатки Google Cloud Text-to-Speech представлены в таблице 4.

Таблица 4 – Google Cloud Text-to-Speech

Достоинства	Недостатки
<ul style="list-style-type: none"> • Можно заранее проверить качество синтезированной речи на необходимом тексте. • Подробная документация. • Можно выбрать формат выходного аудиофайла. • Оптимизация под тип динамика. • Бесплатный пробный период. • Поддержка различных голосов. Помимо стандартных голосов есть возможность создать уникальный пользовательский голос. Для стандартных голосов можно изменять тон, громкость и скорость. 	<ul style="list-style-type: none"> • Документация на английском языке. • Является платным сервисом, стоимость использования зависит от количества символов в тексте. • Имеет закрытый исходный код. • Необходимо обучение для использования продукта.

1.2.5 Обоснование необходимости разработки нового продукта

Нами были рассмотрены лишь некоторые из наиболее известных text-to-speech систем. На сегодняшний день можно найти большое количество подобных систем, но все они в основном обладают схожими недостатками. Кроме того, ни одна из приведенных систем не осуществляет просодическое оформление текста, так как эта задача является сложной. В связи с этим в данной работе также не ставится задача придания тексту интонационного окраса. Проанализировав существующие на рынке решения для преобразования текста в речь, можно сделать вывод, что разработка нового продукта имеет смысл.

Таким образом, разрабатываемый программный продукт должен сочетать в себе следующие особенности:

- поддержка русского языка;
- учет фонетических особенностей русского языка;
- открытый исходный код;
- бесплатное распространение;
- простота использования, документация на русском языке.

Главной чертой программы является поддержка русского языка для ее распространения в русскоязычном сообществе.

1.3 Методы синтеза речи

Существует несколько алгоритмов генерации речи. Выбор алгоритма зависит, в первую очередь, от задачи, для которой он будет использоваться.

В настоящее время выделяют три группы методов синтеза речи:

- параметрический синтез;
- компилятивный синтез;
- синтез речи по фонетическим правилам [8].

1.3.1 Параметрический синтез

Параметрический синтез применяется в вокодерных системах. Вокодеры синтезируют речь на основе сигнала, представленного набором непрерывно изменяющихся во времени параметров. Эти параметры и управляют синтезатором речи.

Параметрический метод используется, когда набор текстовых сообщений, которые необходимо озвучить, ограничен и заранее известен. В этом случае качество синтезируемой речи может быть очень высоким. Недостатком метода является невозможность его применения к произвольному набору заранее неизвестных текстовых сообщений.

1.3.2 Компилятивный синтез

Суть данного метода заключается в компиляции речи из предварительно заготовленных и записанных исходных элементов синтеза – слов, фраз, предложений. Минимальным размером элементов при этом является слово.

Таким образом, возможности компилятивного метода ограничены объемом словаря.

Главная проблема метода – необходимость в большом объеме памяти для хранения предварительно записанного словаря, при этом содержание и вариативность синтезируемой речи ограничена этим словарем. Несмотря на это компилятивный синтез получил большое распространение и широкое практическое применение в узконаправленных системах с ограниченным набором озвучиваемых сообщений. Например, объявления в метро: «Следующая станция ...», «Поезд следует до станции ...». Названия станций записаны заранее и просто соединяются с объявлением. Другим примером являются справочные службы сотовых операторов при озвучивании информации об остатке средств на счету абонента.

Этот метод является самым простым и качество выходной речи зависит только от способа записи минимальных элементов синтеза.

1.3.3 Синтез речи по фонетическим правилам

Более сложными, но худшими по качеству являются алгоритмы, разделяющие речь на более мелкие части. Метод синтеза речи по фонетическим правилам, основываясь на размере минимальных исходных элементов синтеза, делят на следующие виды:

- микросегментный;
- аллофонный;
- дифонный;
- полуслоговой;
- слоговой [8].

Синтез речи по фонетическим правилам позволяет синтезировать речь по заранее неизвестному тексту. Однако синтезированная таким образом речь может отличаться от естественной, потому что при склейке элементов на их

границах возникают искажения. Чем длиннее становятся минимальные блоки, тем больше таких элементов в словаре и качество синтезируемой речи увеличивается вместе с требуемой памятью.

Кроме того, при таком подходе сложно управлять просодическими характеристиками синтезированной речи, поскольку интонационные и акустические характеристики слова зависят от его места в предложении.

1.3.4 Выбор метода синтеза речи

В таблице 5 приведен краткий перечень достоинств и недостатков всех описанных выше методов.

Таблица 5 – Сравнение методов синтеза речи

Название метода	Достоинства	Недостатки
Параметрический	<ul style="list-style-type: none"> Очень высокое качество речи. 	<ul style="list-style-type: none"> Набор текстовых сообщений, которые необходимо озвучить, должен быть заранее известен.
Компилятивный	<ul style="list-style-type: none"> Простота. Качество синтезируемой речи зависит только качества записи элементов синтеза. 	<ul style="list-style-type: none"> Большой объем памяти для хранения предварительно записанного словаря. Содержание и вариативность синтезируемой речи ограничена этим словарем.
По фонетическим правилам	<ul style="list-style-type: none"> Самый сложный из всех трех методов. Синтезированная речь имеет худшее качество, чем в предыдущих двух методах. 	<ul style="list-style-type: none"> Позволяет синтезировать речь по заранее неизвестному тексту.

Исходя из приведенной в таблице информации, в качестве способа генерации речи был выбран синтез по фонетическим правилам, потому как это единственный метод, позволяющий озвучить произвольный заранее неизвестный текст. Для реализации данного подхода необходимо создать фонетико-акустическую базу данных, содержащую большое количество минимальных исходных элементов синтеза.

1.4 Выбор исходных элементов при создании базы данных для синтеза речи по правилам

Выбор исходных единиц синтеза существенно влияет на качество синтезируемой речи, отсюда возникает задача оптимального выбора единиц, с которыми будет работать синтезатор речи.

Необходимо учитывать, что фонетические единицы могут изменяться под действием языковых правил (ассимиляция, диссимиляция, оглушение и изменение гласных звукотипов в зависимости от ударения), а также под действием моторной программы (аккомодация). Аккомодацией (коартикуляцией) называется одновременное осуществление артикуляционных движений, относящихся к разным артикуляционным жестам, что облегчает артикуляционный переход от одного звука к другому [3]. При произнесении последовательности звуков возникают переходные участки, так как для произнесения следующего звука необходимо перестроить речевой аппарат. Переходные участки могут быть короткими по длительности, однако несут в себе важную для восприятия речи информацию. Например, информация о твердости или мягкости согласного часто содержится в следующем за ним гласным. Твердость/мягкость согласных в русском языке является смысловозначительной характеристикой (*гроза – грозя*). Очевидно, что для качественного синтеза речи необходимо учитывать переходные участки. Для этого необходимо в качестве исходных элементов выбирать более крупные чем отдельные звуки элементы. Такими элементами являются аллофоны.

Кроме аллофонного, существует дифонный синтез речи. Дифоном называют сегмент речи между серединами соседних аллофонов. Однако при дифонном синтезе иногда бывает трудно выделить границу деления на дифоны, что негативно сказывается на качестве синтезируемой речи. Чтобы устранить потерю качества, для каждого дифона необходимо более детально учитывать контекст, что приводит к значительному увеличению размеров исходной базы

данных. Использование полуслогов и более крупных элементов в качестве исходных имеет те же недостатки, что и использование дифонов.

В таблице 6 приведен краткий перечень достоинств и недостатков использования различных исходных элементов в синтезе.

Таблица 6 – Сравнение исходных элементов синтеза речи

Исходные элементы	Достоинства	Недостатки
Микросегменты	<ul style="list-style-type: none"> • Наиболее «детально» из всех вариантов учитывается контекст. 	<ul style="list-style-type: none"> • Большой объем БД. • Трудно выделить границу микросегментов. • Много «склеек».
Аллофоны	<ul style="list-style-type: none"> • При использовании мини-набора объем БД будет небольшим. 	<ul style="list-style-type: none"> • Трудно выделить границу аллофонов. • Для улучшения качества синтеза необходимо сильно увеличить объем БД.
Дифоны	<ul style="list-style-type: none"> • Контекст учитывается более детально. 	<ul style="list-style-type: none"> • Трудно выделить границу дифонов. • Для улучшения качества синтеза необходимо сильно увеличить объем БД.
Полуслоги	<ul style="list-style-type: none"> • Контекст учитывается более детально. 	<ul style="list-style-type: none"> • Трудно выделить границу полуслогов. •
Слоги	<ul style="list-style-type: none"> • Меньше проблем с выделением границ. • 	<ul style="list-style-type: none"> • Большой объем БД.

Исходя из представленной выше информации можно сделать вывод, что аллофоны являются наилучшим компромиссом между возможным качеством синтезируемой речи и объемом БД исходных элементов.

1.5 Мини- и макси-наборы аллофонов для синтеза русской речи

Всего в русском языке насчитывается 42 фонемы, из них 6 гласных и 36 согласных. В потоке речи акустико-артикуляционные характеристики фонем изменяются, что и приводит к появлению аллофонов (оттенков фонем). Аллофоны делятся на два типа: позиционные и комбинаторные.

Позиционные аллофоны определяются положением фонемы по отношению к ударному слогу в слове [5]. Основными фонетическими коррелятами ударения в современном русском литературном языке служат длительность и спектральные характеристики гласных. По этим параметрам в СРЛЯ выделяется двухкомпонентное просодическое ядро слова, состоящее из ударного и 1-го предударного слогов [3]. Ударный и первый предударный слоги противопоставлены другим слогам. Русский лингвист А. А. Потебня описал ритмическую схему слова с просодическим ядром при помощи так называемой формулы Потебни: 112311. Здесь цифрой 3 обозначен ударный слог, цифрой 2 – первый предударный, цифрой 1 – остальные безударные слоги. Эта ритмическая схема является яркой особенностью русского языка и отличает его от всех других языков мира. Типичной схемой индоевропейских языков является схема 1213121.

Комбинаторные аллофоны определяются ближайшим контекстом фонемы [5], то есть являются результатом действия аккомодации (коартикуляции) и ассимиляции. Эффекты аккомодации и ассимиляции в большей степени заметны внутри слоговых комплексов, в меньшей степени на стыках слогов и в еще меньшей, но по-прежнему заметной, на стыках слов внутри синтагмы. На стыках синтагм эти эффекты практически не действуют. Таким образом, описать фонемы в потоке речи (аллофоны) можно с различной степенью детализации.

Речевой корпус должен удовлетворять следующим требованиям:

- быть фонетически полным, то есть содержать все основные варианты фонем;
- объем корпуса должен быть по возможности минимизирован;

- корпус должен быть фонетически сбалансированным.

Обычно выделяют и используют два практически обоснованных варианта набора аллофонов: мини- и макси-наборы. В макси-наборах выделяют пять позиционных аллофонов гласных (ударный, частично ударный, первый предударный, не первый предударный, заударный) и двадцать два комбинаторных аллофона гласных, итого для шести гласных русского языка выделяют 2700 аллофонов. Для согласных всего выделяют 1440 аллофонов [9]. В мини-наборе создается 240 аллофонов гласных и 180 аллофонов согласных [9]. Эти оценки являются теоретическими и сильно завышены, поскольку на практике многие аллофоны либо не используются, либо не являются смыслоразличающими, поэтому ими можно пренебречь. В итоге в макси-наборе практически используются 1550 аллофонов гласных и 209 аллофонов согласных, а в мини-наборе – 175 аллофонов гласных и 81 аллофон согласных.

Очевидно, что использование макси-набора представляет собой объемную и сложную задачу, поэтому в данной работе будет использоваться мини-набор. В связи с тем, что для записи базы диктора необходимо производить в студийных условиях, в качестве речевой базы будет использоваться корпус RUSLAN (Russian Spoken Language Corpus For Speech Synthesis) [21]. RUSLAN является открытым корпусом русского разговорного языка, содержит 22200 аудиозаписей с текстовыми аннотациями и является крупнейшим с точки зрения продолжительности речи одного диктора корпусом. Корпус распространяется под лицензией CC BY-NC-SA 4.0.

1.6 Модификация метода аллофонного синтеза по фонетическим правилам

Очень часто в потоке речи взаимовлияние соседних аллофонов друг на друга оказывается настолько сильным, что при создании БД аллофонов не всегда удается провести между ними четкую границу. Так, например, очень часто невозможно отделить сонорный согласный от соседнего гласного.

При работе с речевым сигналом часто используются спектрограммы. Спектрограмма – это трехмерная диаграмма. На горизонтальной оси представлено время, по вертикальной оси – частота. Третье измерение с указанием амплитуды на определенной частоте в конкретный момент времени представлено интенсивностью или цветом каждой точки изображения. Рассмотрим спектрограмму слова «загремела» на рисунке 1. Спектрограмма получена с помощью аудиоредактора Adobe Audition [10]. На представленной спектрограмме видно, что границы первых трех звуков прослеживаются достаточно отчетливо, в то время как границы последующих звуков провести очень трудно. Такая ситуация приведет к тому, что при «нарезке» аллофонов из-за невозможности корректно провести границу качество «нарезанных» аллофонов будет низким. При их «склейке» в процессе синтеза речи в местах соединения будут возникать резкие переходы (скачки) и обрывы, что в свою очередь сильно снизит качество синтезируемой речи.

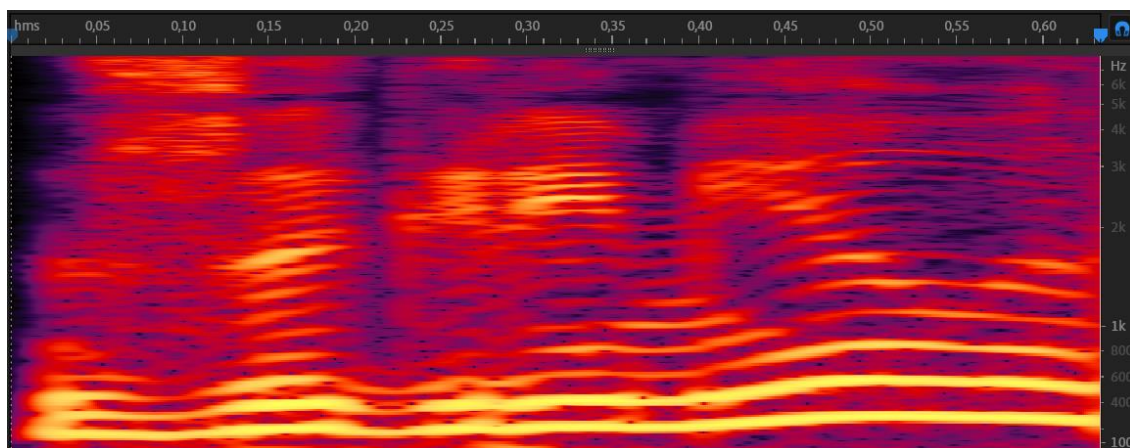


Рисунок 1 – Спектрограмма слова «загремела»

Для того, чтобы преодолеть этот недостаток, в данной работе предлагается в качестве исходных элементов синтеза использовать не только аллофоны, но и более протяжённые фонетические сегменты (аллослоги) [7].

Определяются три типа слоговых комплексов [7]:

- трудно сегментируемые (1-й тип);
- умеренной трудности сегментации (2-й тип);
- легко сегментируемые (3-й тип).

Слоговой комплекс первого типа (трудно сегментируемый) – это открытый слог, при этом:

- если за гласным, определяющим конец слога, находится гласный, он присоединяется к текущему слогу;
- если за гласным, определяющим конец слога, находится последовательность “J’ – безударный гласный”, вся последовательность присоединяется к текущему слогу;
- если слог состоит из одного гласного, он присоединяется к последующему слогу, формируя слоговой комплекс.

Слоговой комплекс второго типа (умеренной трудности сегментации) – это открытый слог, для которого выполняются правила слогового комплекса первого типа, и:

- если за гласным, определяющим конец слога, следует не менее двух согласных, первый из которых – сонант или J’, V, V’, а последующий – нет, то граница определяется после первого из них.

Слоговой комплекс третьего типа (легко сегментируемый) – это открытый слог, для которого выполняются правила слогового комплекса второго типа, и:

- если за концом слога находится последовательность “сонант – безударный гласный”, она присоединяется к текущему слогу.

В процессе синтеза речи аллофонная последовательность, полученная на этапе преобразования фонем в аллофоны, размечается на слоговые комплексы трех типов. После этого в БД исходных элементов синтеза осуществляется поиск в соответствии со следующим приоритетом: слоговые комплексы 3-го типа, слоговые комплексы 2-го типа, слоговые комплексы 1-го типа. В случае, если в БД не найден слоговой комплекс 3-го типа, проводится поиск составляющих его слоговых комплексов других типов в соответствии с приоритетом. Если же в БД не найден ни один из сформированных типов аллослогов, осуществляется поиск составляющих его аллофонов.

Таким образом, синтагма может быть размечена четырьмя способами (три типа слоговых комплексов и просто аллофоны). При описанном выше алгоритме

выбора исходных элементов синтеза из БД аллофоны используются только в тех крайних случаях, когда в БД отсутствуют элементы синтеза верхних уровней: аллослоги трех типов.

Таким образом, в качестве модификации исходного метода аллофонного синтеза речи по фонетическим правилам в данной работе предлагается использование комбинированного метода синтеза по фонетическим правилам с использованием аллофонов и аллослогов. Для реализации описанного подхода необходимо, помимо БД аллофонов, создать БД аллослогов.

2 Конструкторский раздел

2.1 Требования к программе

Программа должна синтезировать речь по произвольному тексту на русском языке. Для синтеза используется комбинированный метод синтеза речи по фонетическим правилам с использованием аллофонов и аллослогов. Входные данные подаются программе в виде файла. Синтезированная речь должна быть сохранена в виде аудио файла.

2.2 Описание входных и выходных данных и ограничений

Входными данными для программы является текстовый файл. Файл может содержать в себе:

- собственно текст на русском языке, состоящий из предложений или отдельных слов;
- аббревиатуры;
- знаки пунктуации;
- слова с буквой «е» вместо «ё».

Следует обратить внимание, что текст должен иметь правильную пунктуацию, так как от этого зависит правильность его разбиения на предложения и слова, что, в свою очередь, влияет на качество синтезируемой речи.

Файл не может содержать в себе (ограничения):

- иностранные слова;
- специальные символы;
- многоразрядные и дробные числа;
- телефонные номера;
- разметка словесного ударения;

- сокращения;
- обозначения времени и даты;
- интернет-адреса;
- какую-либо разметку текста (например, просодическую);
- математические выражения.

Важным ограничением разрабатываемой программы является отсутствие полноценного просодического оформления текста. Программа делит текст на синтагмы опираясь только на знаки пунктуации (пунктуационные синтагмы), членение на лексические и синтаксические синтагмы не производится, потому что эта задача является объемной и сложной и выходит за рамки данной работы.

Функциональная модель (IDEF0-диаграмма) разрабатываемой системы представлена на рисунках 2 – 5.

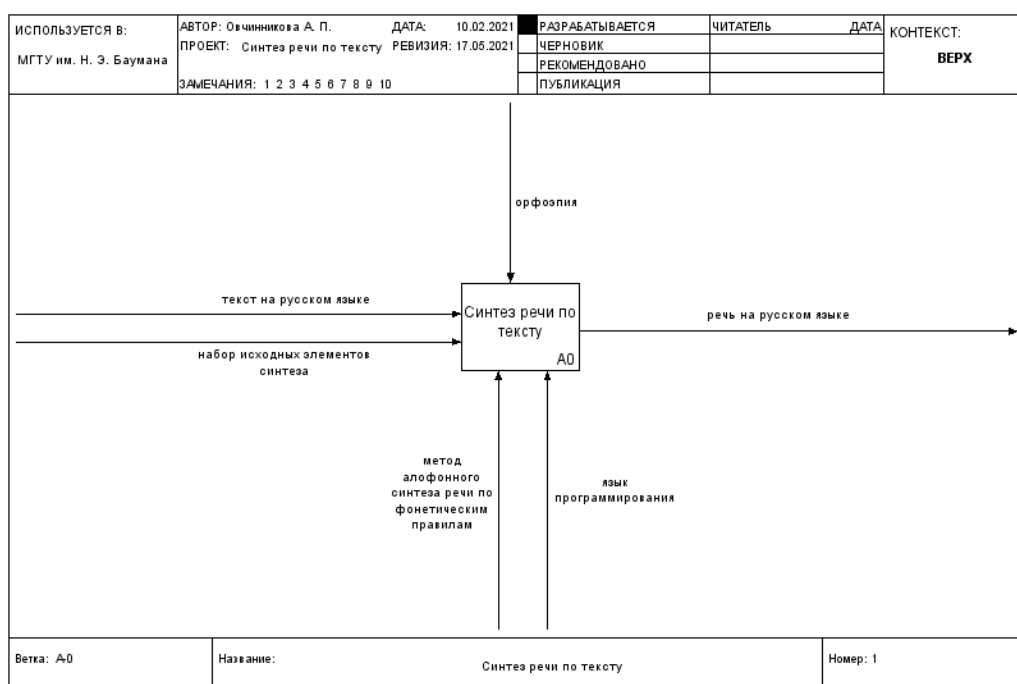


Рисунок 2 – Контекстная диаграмма системы

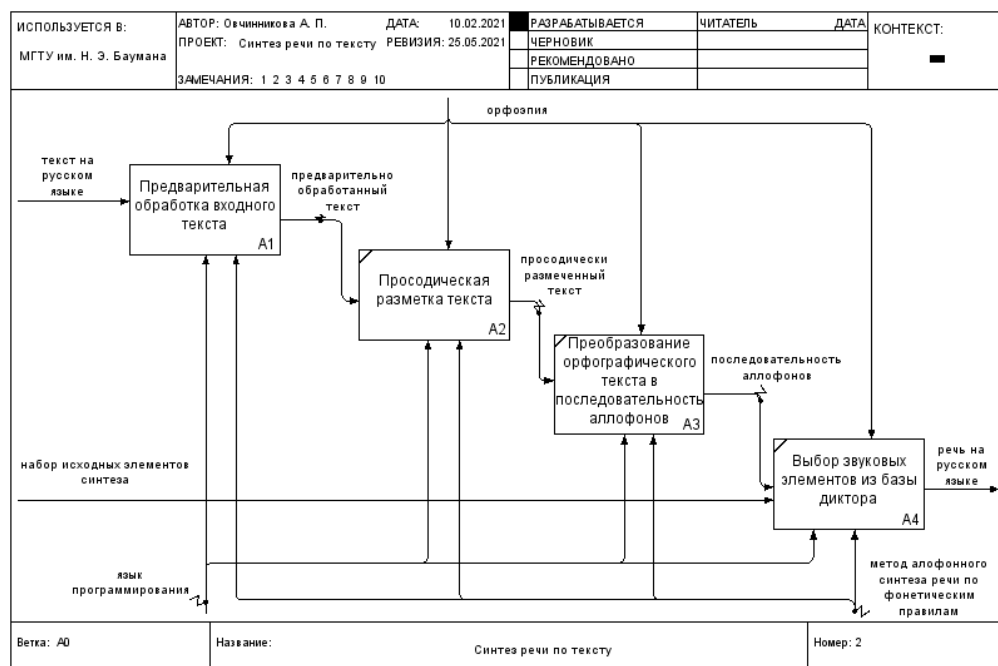


Рисунок 3 – Декомпозиция работы «Синтез речи по тексту»

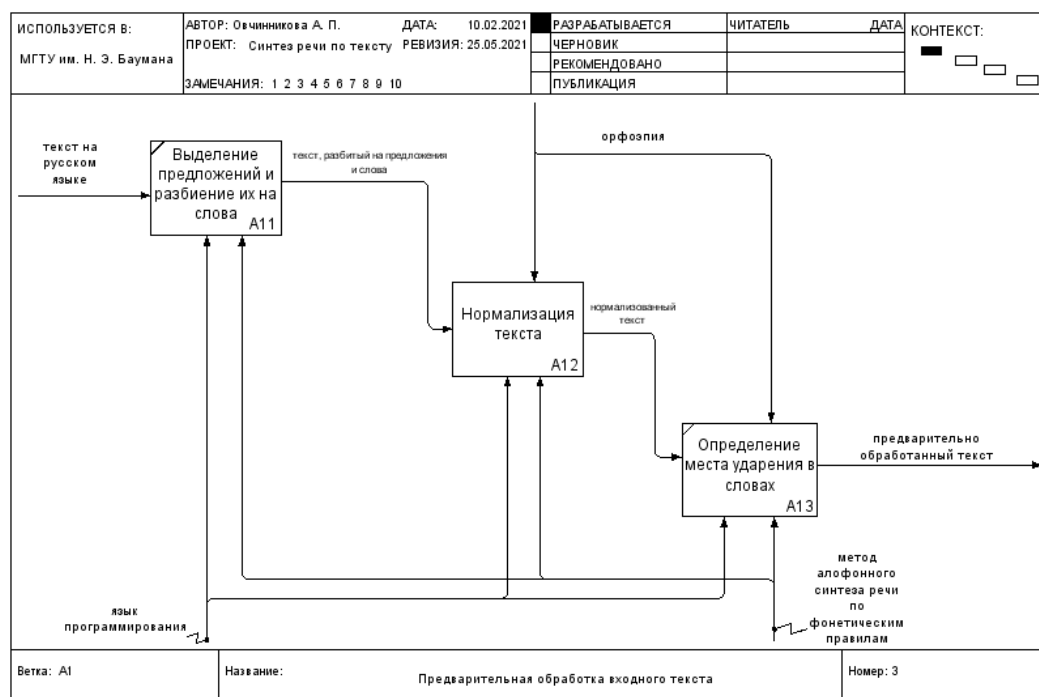


Рисунок 4 – Декомпозиция работы «Предварительная обработка входного текста»

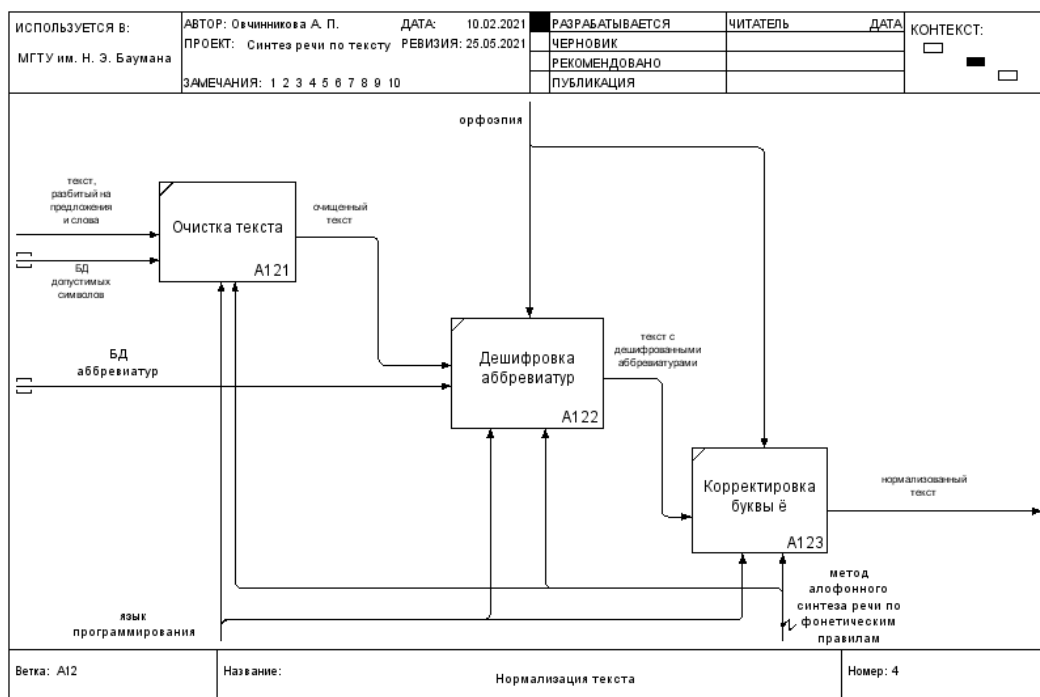


Рисунок 5 – Декомпозиция работы «Нормализация текста»

2.3 Система синтеза речи по тексту

Исходя из приведенной IDEF0-диаграммы, система аллофонного синтеза речи по тексту должна состоять из четырех подсистем: лингвистического текстового процессора, просодического процессора, фонетического процессора и акустического процессора.

2.3.1 Лингвистический текстовый процессор

Задачей лингвистического текстового процессора является предварительная обработка текста, а именно выделение предложений и разбиение их на слова, нормализация текста (очистка текста, дешифровка аббревиатур и корректировка буквы «ё»), определение места ударения в словах, преобразование слов-фонетических исключений и формирование фонетических слов. Схема алгоритма

работы лингвистического текстового процессора представлена на рисунке 6.



Рисунок 6 – Схема алгоритма работы лингвистического текстового процессора

Выделение предложений

В первом приближении предложениями можно считать отрезки исходного текста между точками или другими конечными знаками препинания. Однако у такого подхода есть серьезный недостаток: в русском языке точка не всегда является маркером конца предложения. Она может быть элементом сокращения, инициалов, даты, дроби, интернет ссылки или email адреса. Большинство из этих случаев невозможно отличить от конца предложения без семантического анализа текста. Например, если в тексте встретилась последовательность «рис.», то без понимания смысла текста невозможно понять, является ли «рис.» сокращением слова «рисунок», или обозначает зерна одноименного растения, а точка свидетельствует о конце предложения.

Выделение слов

В лингвистическом энциклопедическом словаре дано следующее определение слова. «Слово – основная структурно-семантическая единица языка, служащая для именования предметов и их свойств, явлений, отношений действительности, обладающая совокупностью семантических, фонетических и грамматических признаков, специфичных для каждого языка. Характерные признаки слова — цельность, выделяемость и свободная воспроизводимость в речи. В слове различаются следующие структуры: фонетическая (организованная совокупность звуковых явлений, образующих звуковую оболочку слова), морфологическая (совокупность морфем), семантическая (совокупность значений слова)» [4]. В данной работе словом можно считать отрезки текста между пробелами (т.е. будем говорить о графическом слове), но со следующими оговорками:

- знаки препинания обычно не отделяются пробелами от слов, а после слова может следовать несколько знаков препинания;

- специальные символы могут быть не отделены пробелами от предыдущего слова (например сочетания цифры и процента).

Исходя из этих соображений корректная пунктуация входящего текста является крайне важной.

Нормализация текста

Нормализация текста состоит из нескольких этапов: очистка текста, дешифровка аббревиатур, корректировка буквы «ё».

Целью очистки текста является удаление неинформативных для синтеза речи символов. Исходя из описанных выше в данном разделе ограничений на входной текст, такими символами являются иностранные буквы, специальные символы (проценты, доллары и т. д.), многоразрядные и дробные числа, телефонные номера, обозначения времени и даты, интернет-адреса (т. е. снова иностранные буквы), математические выражения. От алгоритма работы блока нормализации зависит полнота озвучивания исходного текста, однако обработка всех этих случаев является очень объемной задачей. Так, например, для озвучивания символа «%» необходимо предварительно провести синтаксический разбор предложения и построить дерево зависимостей, чтобы правильно согласовать слово «процент» с тем словом, от которого оно зависит.

Под дешифровкой аббревиатур здесь понимается установление правил их чтения, а не расшифровка слов, сокращением которых она является. Дешифровка аббревиатур необходима в связи с тем, что правила чтения аббревиатур в русском языке отличаются от правил чтения обычных слов. Для обнаружения аббревиатур в тексте можно использовать правило, что в русском языке они всегда пишутся заглавными буквами. При произношении инициальных аббревиатур возникают некоторые трудности. Чтобы чтения аббревиатуры было правильным, необходимо завести возможно более полную базу данных аббревиатур и их произношений. Если же в тексте встретилась аббревиатура, не

присутствующая в базе данных, необходимо руководствоваться следующими правилами:

- если в состав аббревиатуры входят только буквы, обозначающие согласные звуки, она читается по названиям составляющих ее букв (*СССР* [эс-эс-эс-эр]);
- если в составе аббревиатуры есть буквы, обозначающие гласные звуки, то она читается по слогам как обыкновенное слово (*МИД* [мид]); однако если гласный звук конечный или начальный, то она читается по названиям составляющих ее букв (*МГУ* [эм-гэ-у]);
- некоторые аббревиатуры читаются частично «по слогам», частично – по названиям букв: *ГИБДД* [ги-бэ-дэ-дэ];
- при произношении некоторых аббревиатур могут использоваться разговорные названия букв (*США* читается как [сэ-шэ-а]);
- буква «л» всегда читается как «эл», а не «эль»;
- словесное ударение при чтении аббревиатур обычно падает на последний слог, однако это не всегда так, а ударение в иноязычных аббревиатурах лучше проверять по словарю.

Все эти правила необходимо учитывать при расшифровке. Схема алгоритма дешифровки неизвестной аббревиатуры представлена на рисунке 7. На рисунке 8 представлена схема алгоритма поиска и дешифровки аббревиатур в синтагме. Алгоритм проверяет каждое слово, являются ли все буквы в нем заглавными. Если это так, то слово считается аббревиатурой. Аббревиатуру необходимо найти в словаре аббревиатур, чтобы определить её произносительный вариант. В случае, если аббревиатура не найдена в словаре, запускается алгоритм дешифровки неизвестной аббревиатуры, который работает по описанным выше правилам.

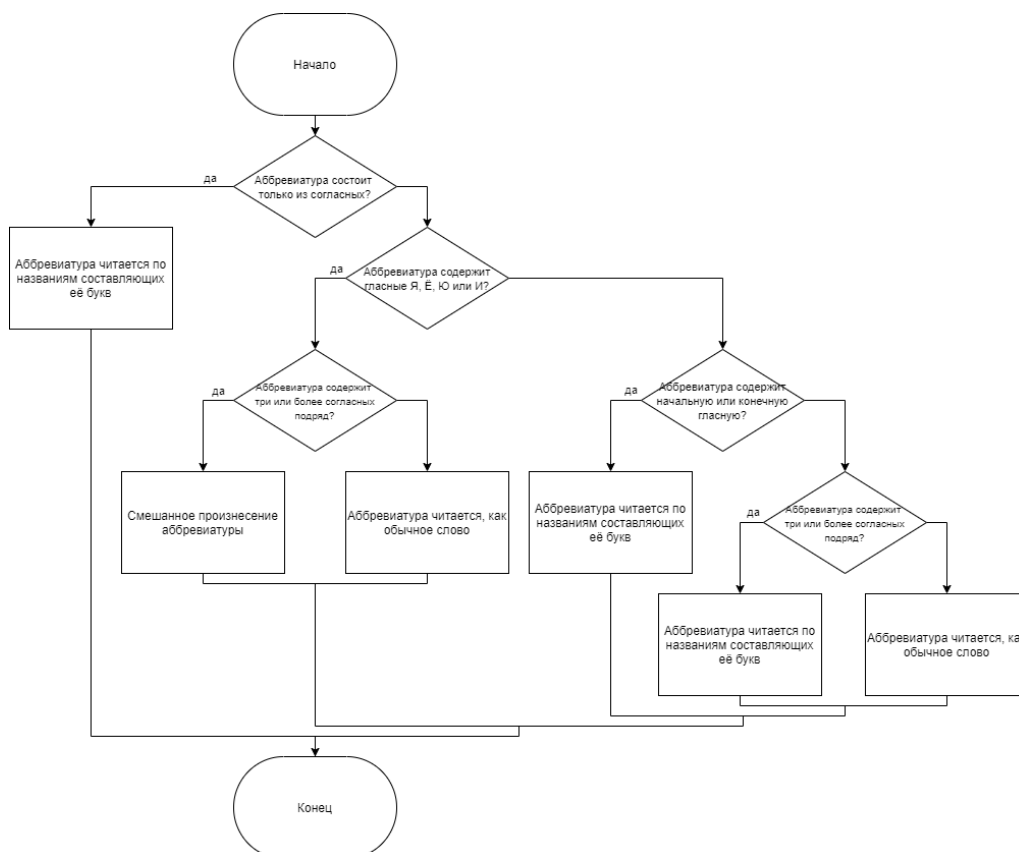


Рисунок 7 – Схема алгоритма дешифровки неизвестной аббревиатуры

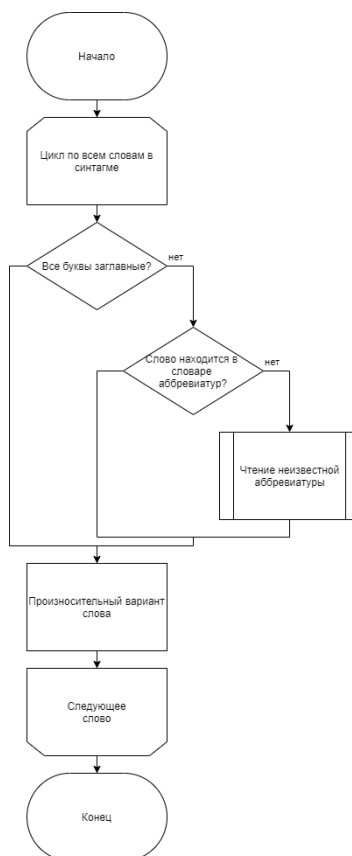


Рисунок 8 – Схема алгоритма поиска аббревиатур

В современном русском языке буква «ё» все более выходит из употребления. Правила СРЛЯ не запрещают писать букву «е» вместо «ё» и русскоговорящий человек при чтении не задумывается о том, с какой же буквой правильно прочитав слово. Однако в речи при произнесении «е» вместо «ё» слушатель сразу заметит неточность. Поэтому крайне важно при синтезе речи проводить корректировку буквы «ё». Для этого необходимо завести базу данных с наиболее полным перечнем слов с буквой «ё». Слова, где содержится одна или более букв «е» необходимо искать в этой БД и в случае, если слово будет найдено, необходимо произвести замену буквы «е» на «ё».

Схема алгоритма определения наличия буквы «ё» в слове представлена на рисунке 9.

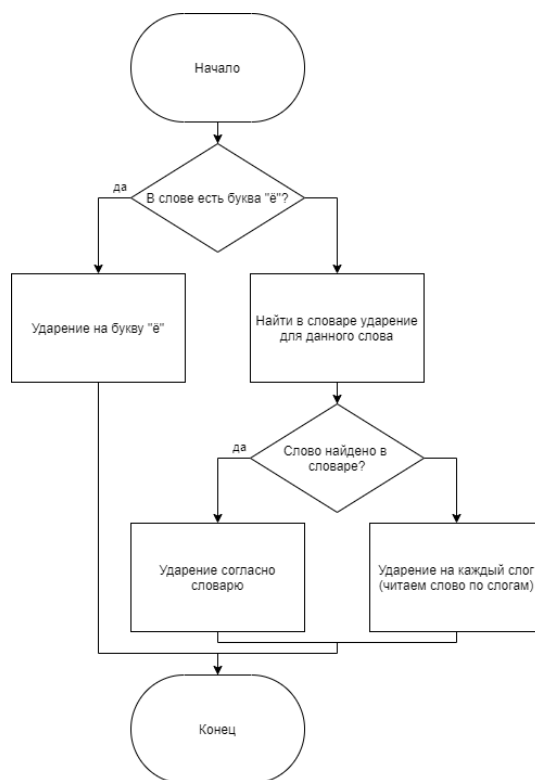


Рисунок 9 – Схема алгоритма определения наличия буквы «ё» в слове

Определение места ударения в словах

Для каждого слова необходимо определить положение словесного ударения. Для этого используется полная акцентуированная парадигма по А. А. Зализняку, которая содержит наиболее полный перечень словоформ русского

языка с пометами об основном и побочном ударении. Следует учитывать омографы, которые при одинаковом буквенном составе имеют разное ударение. Для разрешения таких ситуаций в некоторых случаях достаточно морфологического анализа текста (слова «ру+ки» и «руки+» имеют разные падежи), а в некоторых случаях морфологического анализа недостаточно и необходимо проводить семантический анализ. Например, слова «за+мок» и «замо+к» стоят в одинаковых падежах, и чтобы понять, где ставить ударение, необходимо знать, о чем идет речь в тексте. В данной работе обрабатывается только первый случай.

Необходимо также учитывать, что в тексте может встретиться слово, которого нет в словаре. В этом случае можно поступить несколькими способами:

- прочитать слово по слогам, то есть сделать ударным каждый слог;
- попробовать установить позицию словесного ударения исходя из статистической информации.

Первый способ является более предпочтительным, потому что слово, прочитанное по слогам, будет легче воспринято слушателем, чем слово с неправильным словесным ударением. В данной работе будет использоваться первый способ.

Схема алгоритма определения места ударения в слове представлена на рисунке 10.

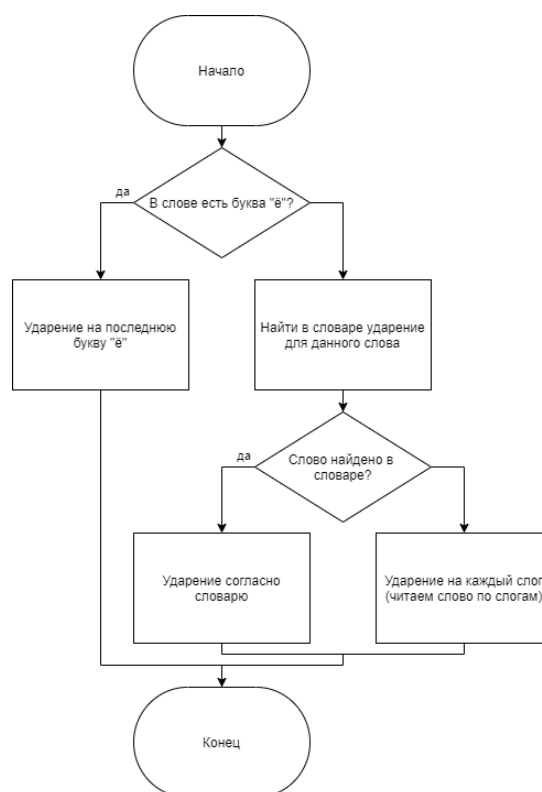


Рисунок 10 – Схема алгоритма определения места ударения в слове

Преобразование слов-фонетических исключений

Фонетическими исключениями являются большое количество иностранных слов или слов с иностранным корнем, например «ателье». Эти слова не подчиняются стандартным правилам преобразования буква-фонема, поэтому их необходимо заменить на эквиваленты, которые этим правилам подчиняются. Например, для слова «ателье» таким эквивалентом будет «атэлье». Список слов – фонетических исключений – приведен, например в [5]. Для хранения этого списка также понадобится база данных.

Формирование фонетических слов

«Фонетическим словом или тактом называется отрезок речевой цепи, объединяемый одним (словесным) ударением» [3]. Главным показателем единства фонетического слова является схема редукции гласных, а также другие

правила, действующие только в пределах фонетического слова (ассимиляция согласных по мягкости). Слова, являющиеся обычно безударными, называются клитиками. В русском языке существует ряд служебных слов, которые не имеют ударения и являются частью предыдущего (энклитики) или следующего (проклитики) фонетического слова. Клитики иногда могут перетягивать на себя ударение со знаменательных слов (энклиноменов), но этот случай является довольно редким и поэтому не будет учитываться в данной работе.

К безударным словоформам относятся:

- неслоговые предлоги и частицы (*в, к, с, б, ж, ль*);
- односложные предлоги и союзы (*без, во, для, за, и* и др.);
- односложные частицы (*ка, бы* и др.).

Энклитиками являются только частицы «бы», «де», «дескать», «ли», «же», «мол», «-то», «-ка», «-либо», «-нибудь».

2.3.2 Просодический процессор

Фонетические средства языка включают в себя фонетические единицы, фонетические признаки и фонетические правила [3]. Задачей этого компонента языка является создание фонетической характеристики речевого сообщения, которая состоит не только из сегментов-фонем, но и из просодических признаков.

Просодические параметры подразделяются на тональные (мелодика, изменения частоты основного тона), количественно-динамические (паузы, длительность, темп и интенсивность), артикуляционные (раствор рта, назализация, смещение язычной артикуляции) и фонационные (различные типы голоса). Все параметры могут быть реализованы как интегрально (на всем высказывании) или локально (обычно на одном или нескольких слогах).

Слова внутри синтагмы объединяются с помощью интонации. «Синтагмой называется речевой фрагмент, на границах которого происходит смена просодического оформления» [3]. Границы синтагмы формируются путем смены

просодического оформления на ее стыках. Центром фонетического слова является ударный слог, а центром синтагмы является слог, на котором реализуется синтагматический акцент [3].

Синтагматический (интонационный) акцент передает слушателю значительную часть информации, поэтому носители языка оперируют словарем минимальных интонационных единиц – интоном. Кроме них словарь содержит их часто используемые устойчивые сочетания, интонационные конструкции (ИК). В русском языке для описания интонации чаще всего используется система Е. А. Брызгуновой [3]. В ней выделяется несколько различных типов ИК, которые используются для передачи специфических значений.

Просодический процессор выделяет в каждом предложении последовательности слов, связанные синтаксической связью, которые, скорее всего, будут представлять из себя цельные просодические единицы (синтагмы). В речи синтагмы отделяются друг от друга паузами.

Считается, что пунктуационные синтагмы ограничены следующими знаками:

- точка с запятой;
- двоеточие;
- запятая;
- тире;
- открывающая и закрывающая скобки;
- комбинация запятой тире;
- открытие и закрытие кавычек [5].

Однако если знак препинания стоит после сочинительного союза (и, да, но и, так и, а и др.), то границу синтагмы в этом месте лучше не проводить.

В данной работе производится выделение только пунктуационных синтагм. Длинные участки текста, расположенные между знаками препинания могут звучать монотонно и осложнять восприятие речи, однако более глубокий анализ требует разработки серьезной системы семантического анализа. При этом

некорректное членение текста на синтагмы может сильно нарушить смысловое восприятие речи и понимание слушателем его смысла.

2.3.3 Фонетический процессор

Задачей фонетического процессора является преобразование орфографического текста в последовательность аллофонов, а также деление последовательности аллофонов на аллослоги. Фонетический процессор состоит из трех подсистем – блока преобразования текста в последовательность фонем (фонемный транскриптор), блока преобразования фонемной транскрипции в последовательность аллофонов и блока деления последовательности аллофонов на аллослоги. Схема алгоритма работы фонетического процессора представлена на рисунке 11.



Рисунок 11 – Схема алгоритма работы фонетического процессора

«Фонема — это абстрактная единица языка, служащая для формирования означающего языковых знаков» [3]. Фонема является абстрактной единицей языка, а ее конкретной реализацией в речи является звук. В отличие от

орфографической записи, которая отражает произносительную норму языка в период зарождения письменности, фонемная транскрипция отражает современные правила произношения. Норма произношения не всегда совпадает с орфографической записью, существуют правила преобразования орфографического текста в фонемную последовательность, а также целый ряд сложившихся исключений. Например, в словах «что» и «чтобы» происходит замена «ч» на «ш», а в таких словах, как «солнце» и «лестный» есть произносимые согласные и т. д.

В русской орфографии в два раза больше гласных букв, чем гласных фонем, а согласных букв, наоборот, в два раза меньше, чем согласных фонем. Всего в русской речи 42 фонемы. Правила преобразования буква-фонема подробно описаны, например, в [3]. Блок преобразования буква-фонема работает на этих правилах.

Как уже было сказано, фонема является абстрактной единицей языка, то есть описывает его в статическом виде. В потоке речи в различных условиях акустико-артикуляционные характеристики фонем изменяются. Таким образом появляются оттенки фонем, аллофоны. Аллофоны разделяют на позиционные, определяющиеся положением фонемы по отношению к словесному ударению, и комбинаторные, определяющиеся ближайшим контекстом фонемы и обусловленные ассимиляцией и коартикуляцией.

Преобразованием фонем в аллофоны занимается блок преобразования фонема-аллофон. Он определяет комбинаторные и позиционные варианты аллофонов в зависимости от контекста, в котором они находятся. Для гласных особенно важным является позиционный индекс аллофона, так как в русском языке гласные подвержены количественной и качественной редукции. Выделяют полноударные, частичноударные, первые предударные и безударные гласные. Самой большой длительностью и силой звука обладают полноударные гласные, а самые слабые и самые короткие – безударные.

В мини-наборе аллофонов, который используется в данной работе, выделяют всего два позиционных аллофона гласных – ударные (0) и безударные

(1). Схема алгоритма определения индекса позиционного аллофона гласных представлена на рисунке 12.

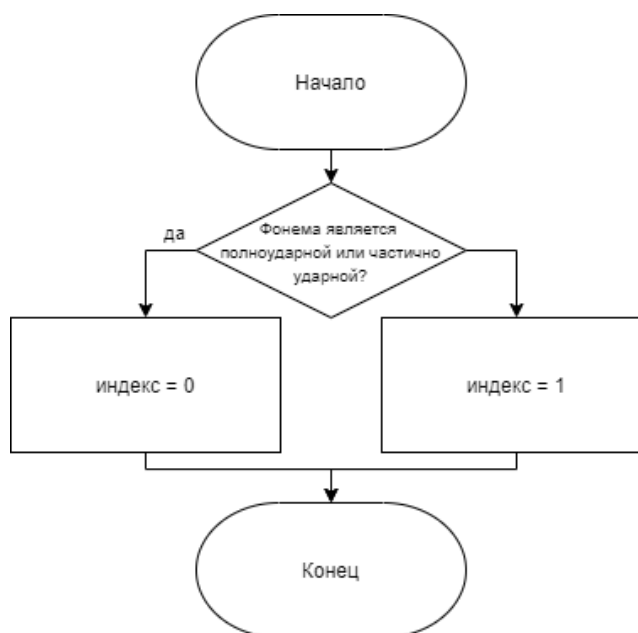


Рисунок 12 – Схема алгоритма определения индекса позиционного аллофона гласных

Для согласных в мини-наборе позиционный индекс не вводится, так как редукция согласных в русском языке далеко не так значительна, как редукция гласных.

С учётом левого контекста создаются следующие комбинаторные аллофоны гласных:

- после синтагматической паузы - (0);
- после твёрдых губных - (1),
- после передне- и среднеязычных - (2);
- после твёрдых заднеязычных и гласных - (3);
- после мягких - (4).

С учётом правого контекста создаются следующие комбинаторные аллофоны гласных:

- перед синтагматической паузой - (0);
- перед переднеязычными и заднеязычными твёрдыми согласными и гласными /У/, /О/, /А/, /Э/, /Ы/ - (1);

- перед губными согласными - (2);
- перед мягкими согласными и гласной /И/ - (3).

Схема алгоритма определения левого комбинаторного индекса гласных представлена на рисунке 13.

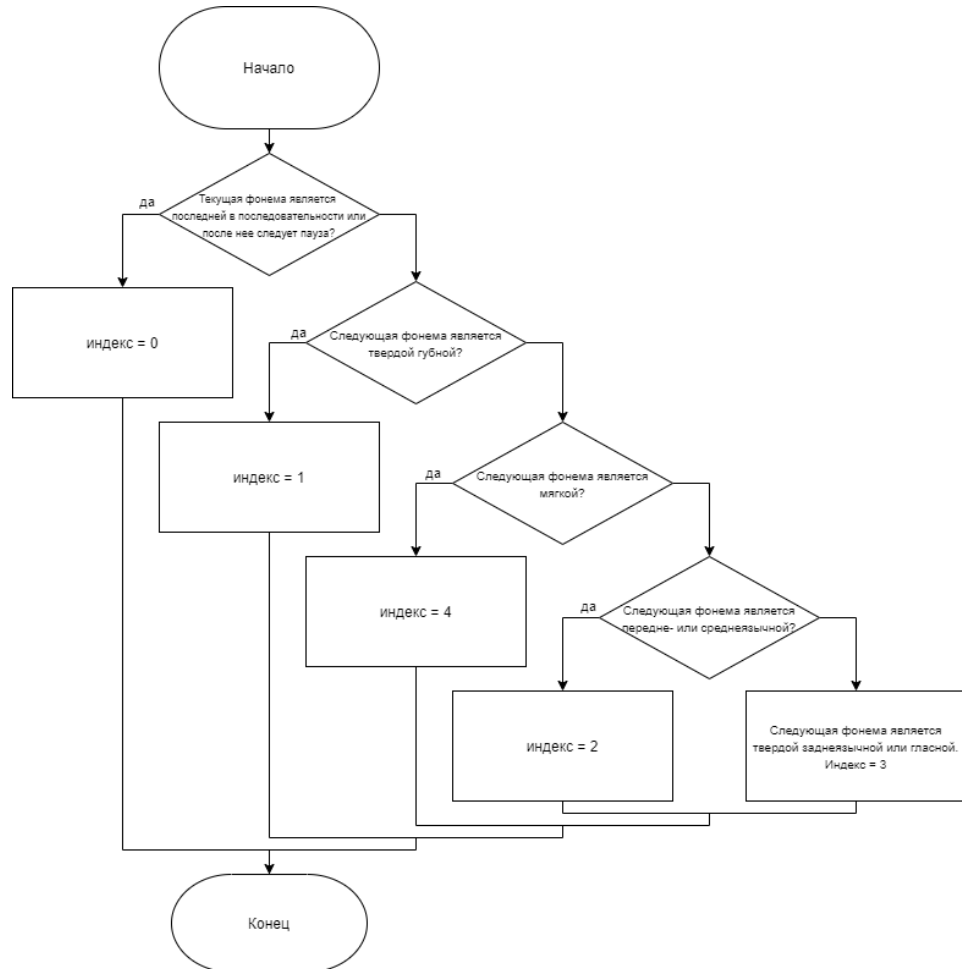


Рисунок 13 – Схема алгоритма определения левого комбинаторного индекса гласных

Схема алгоритма определения правого комбинаторного индекса гласных представлена на рисунке 14.

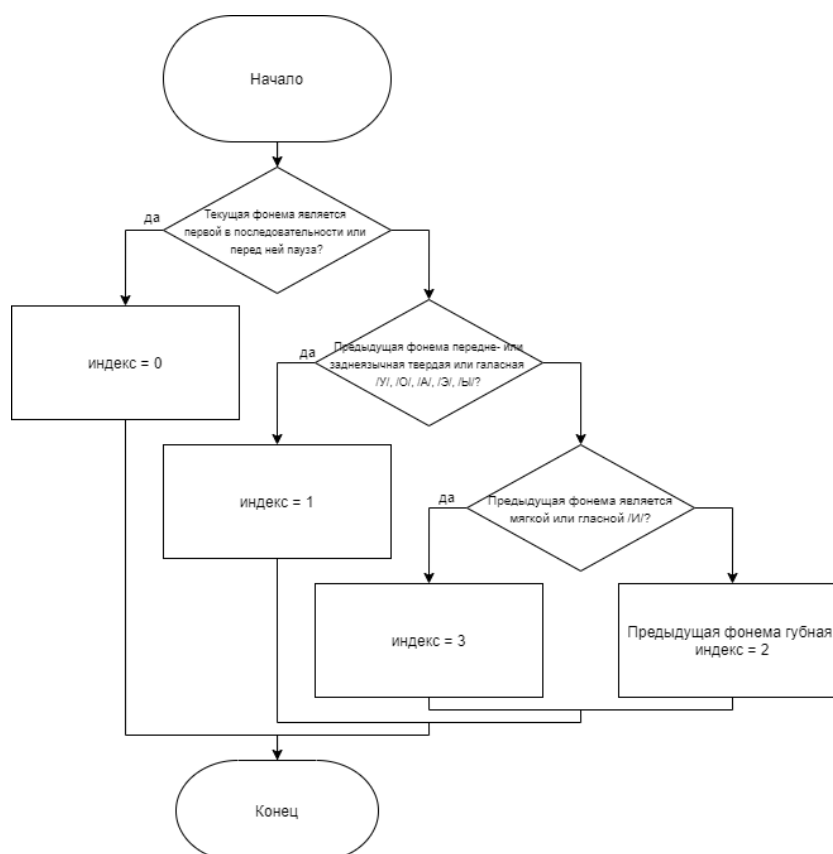


Рисунок 14 – Схема алгоритма определения правого комбинаторного индекса гласных

Аллофоны согласных в мини-наборе создаются только с учётом правого контекста:

- перед паузой - (0);
- перед глухими согласными - (1);
- перед звонкими согласными - (2);
- перед безударными гласными - (3);
- перед ударными гласными - (4).

Схема алгоритма определения комбинаторного индекса согласных представлена на рисунке 15.

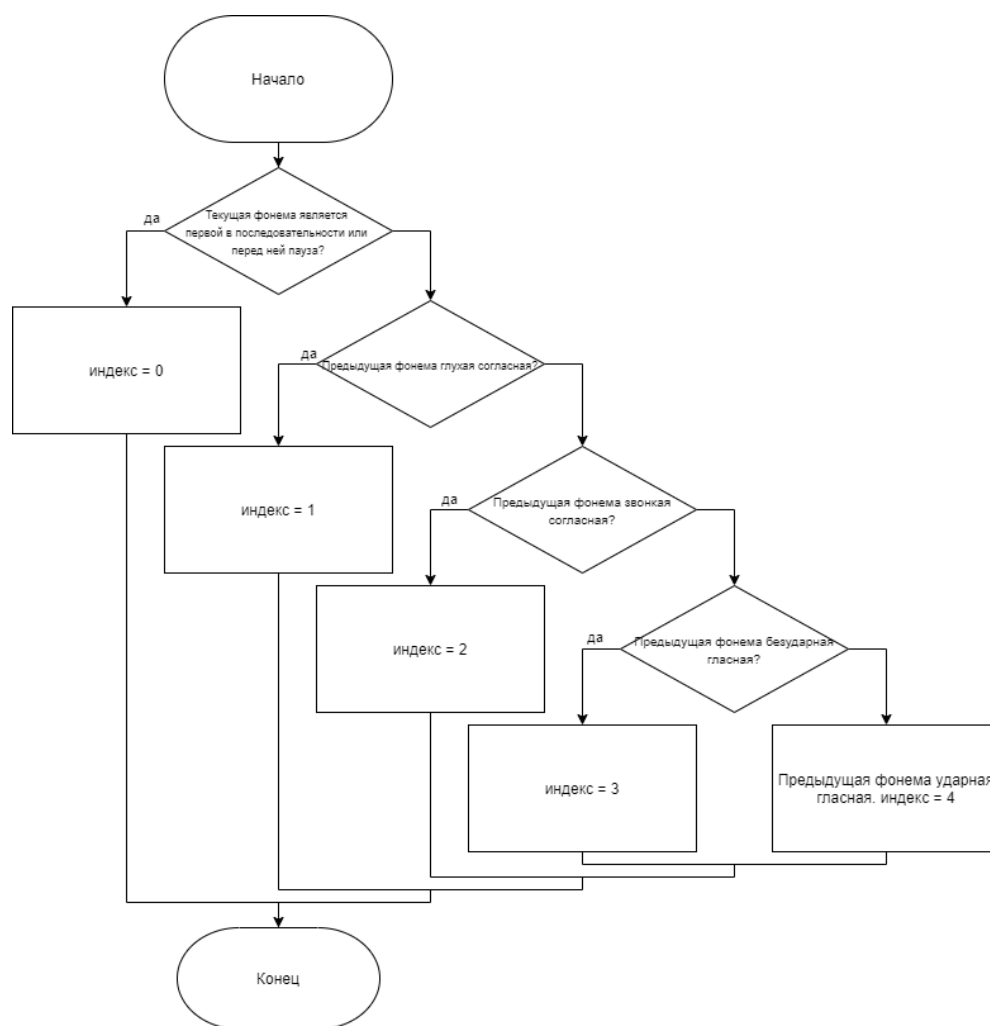


Рисунок 15 – Схема алгоритма определения комбинаторного индекса согласных

На рисунках 16 – 18 представлены схемы алгоритмов деления последовательности аллофонов на слоговые комплексы первого, второго и третьего типов соответственно.



Рисунок 16 – Схема алгоритма деления последовательности аллофонов на слоговые комплексы первого типа

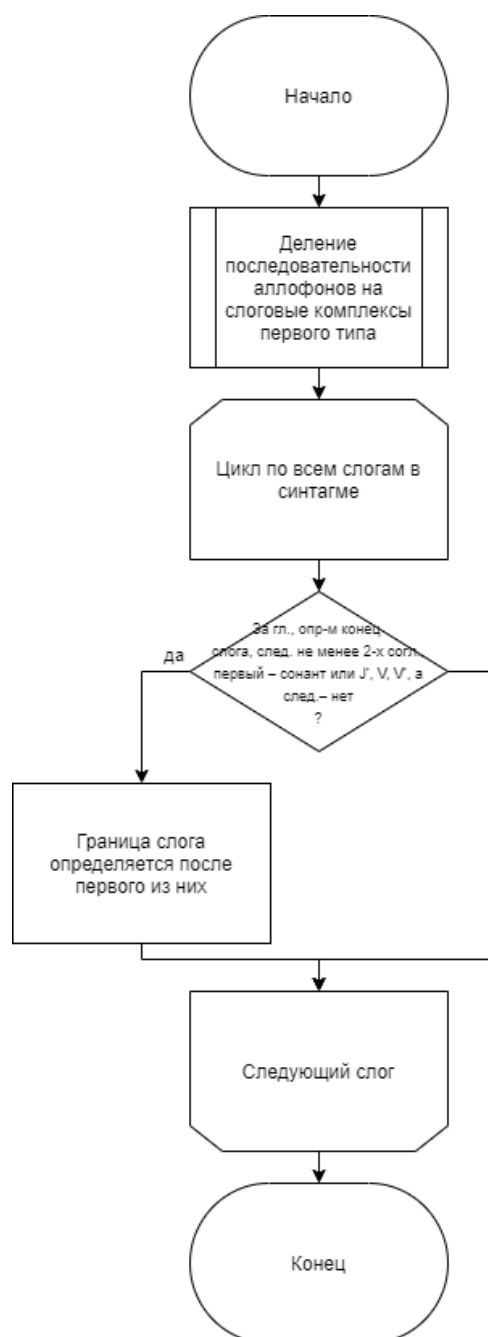


Рисунок 17 – Схема алгоритма деления последовательности аллофонов на слоговые комплексы второго типа

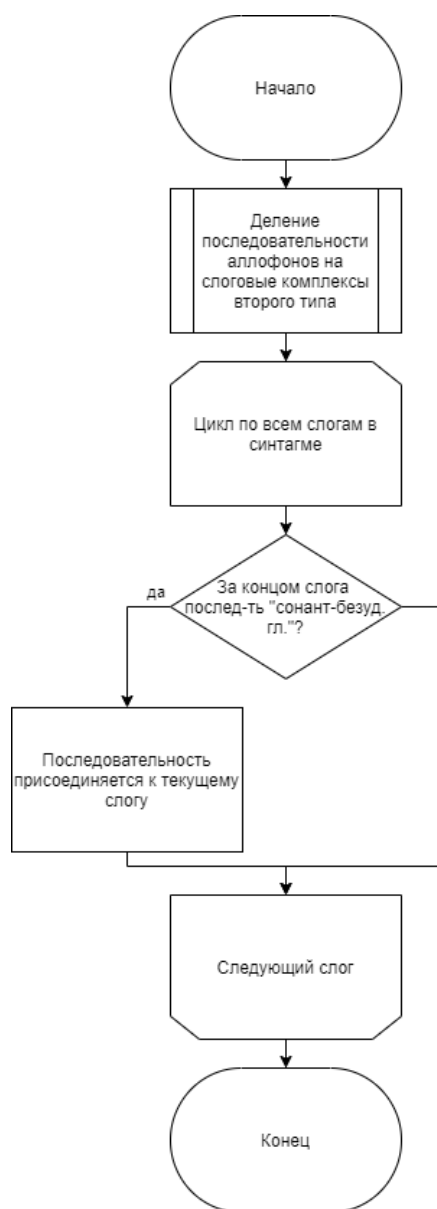


Рисунок 18 – Схема алгоритма деления последовательности аллофонов на слоговые комплексы третьего типа

2.3.4 Акустический процессор

Задачей акустического процессора является выбор необходимых элементов из базы данных исходных элементов и их склейка в общий выходной аудиофайл. Схема разработанного алгоритма поиска необходимых исходных элементов в БД представлена на рисунке 19. Между фонетическими словами добавляется короткая пауза, а между синтагмами добавляется более длинная пауза. После склейки всех аллофонов на синтагму накладывается эффект fade in,

чтобы синтезированная речь звучала чуть более эмоционально. Fade in – это постепенное увеличение громкости аудио-дорожки к ее концу.

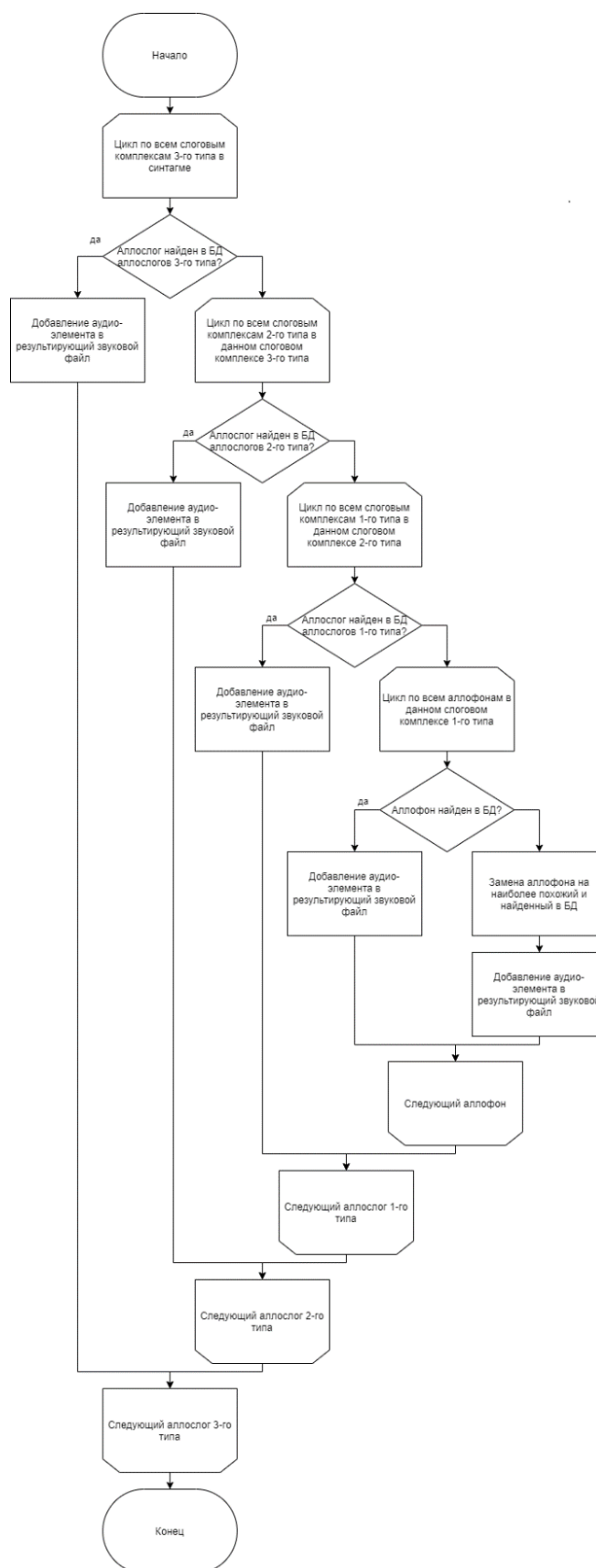


Рисунок 19 – Схема алгоритма поиска исходных элементов в БД

При склейке аллофонов могут возникать щелчки или резкие переходы из-за того, что при нарезке аллофонов бывает сложно разделить их друг от друга. Для того, чтобы речь звучала более равномерно, необходимо сгладить места склейки. Для этого используют алгоритм интерполяции звуковых сигналов – увеличения частоты дискретизации сигнала в N раз. Новые отсчеты сигнала вычисляются по уже имеющимся. Такая процедура называется наложением интерполирующего цифрового фильтра [2].

2.4 Создание акустической базы данных

Для создания акустической базы данных был использован корпус RUSLAN (Russian Spoken Language Corpus For Speech Synthesis) [21]. Корпус содержит много часов озвученных текстов, поэтому для создания БД необходимо «вырезать» аллофоны и аллослоги из текстов. В первую очередь необходим полный мини-набор аллофонов. Количество возможных аллослогов является очень большим, поэтому в данной работе ограничимся лишь необходимым набором аллослогов, необходимых для озвучивания нескольких контрольных примеров (о примерах подробнее будет рассказано в исследовательском разделе).

При использовании мини-набора аллофонов создаётся 2 типа позиционных аллофонов гласных, 5 левых контекстов и 4 правых контекста.

Всего для 6-ти гласных создаются $N_v = 2 * 5 * 4 * 6 = 240$ аллофонов. В мини-наборе создаются только 5 аллофонов согласных учётом правого контекста. Всего для 36-ти согласных создаются $N_c = 5 * 36 = 180$ аллофонов. Всего создаётся $240 + 180 = 420$ аллофонов для синтеза русской речи.

Для нарезки аллофонов из представленного в корпусе RUSLAN фонетического материала были выбраны слова или сочетания слов, где встречаются необходимые аллофоны. Пример таблицы слов, содержащих аллофоны ударной гласной «А» представлен в таблице 8. Пример таблицы слов, содержащих аллофоны согласной «Р», представлен в таблице 7. Таблицы для

остальных аллофонов представлены в приложении А. В таблицах для каждого слова приведена транскрипция и обозначен соответствующий аллофон, который присутствует в данном слове.

Таблица 7 – Мини-набор аллофонов для согласной R

Правый контекст (индекс аллофона)	0	1	2	3	4
	Пауза	Глухой согласный	Звонкий согласный	Безударный гласный	Ударный гласный
Язык					
Русский	<i>Вор</i> [в'ор]	<i>Марка</i> [м'арк'ъ]	<i>Кордон</i> [кар'дон]	<i>Караван</i> [к'ърав'ан]	<i>Парад</i> [пар'ат]

Таблица 8 – Мини-набор аллофонов для ударной гласной А

Третий индекс, правый контекст		0		1		2		3	
		пауза		п, ф, б, в, м		т, ц, с, д, з, н, л, ш, ж, р, к, х, г, у, о, а, э, ы		к', х', г', й, ч', ш', р', т', с', д', з', н', л', п', ф', б', в', м', и	
Второй индекс, левый контекст									
0	пауза	A000	<i>А</i> [а]	A001	<i>Автор</i> [а'ф'т'ър]	A002	<i>Ада</i> [а'д'ъ]	A003	<i>Ася</i> [а'с'ъ]
1	п, ф, б, в, м	A010	<i>Судьба</i> [суд'ба]	A011	<i>Баба</i> [ба'б'ъ]	A012	<i>Хваткой</i> [х'ват'к'ый]	A013	<i>Батя</i> [ба'т'ъ]
2	ш, ж, р, т, ц, с, д, з, н, л	A020	<i>Еда</i> [и'д'а]	A021	<i>Запад</i> [за'п'ат]	A022	<i>Дата</i> [да'т'ъ]	A023	<i>Тася</i> [та'с'ъ]
3	к, х, г, у, о, а, э, ы	A030	<i>Нога</i> [на'га]	A031	<i>Легавый</i> [л'и'га'в'ый]	A032	<i>Сказка</i> [с'ка'з'к'ъ]	A033	<i>Галя</i> [га'л'ъ]
4	т', с', д', з', н', л', ч', ш', р', п', ф', б', в', м', к', х', г', й, и	A040	<i>Хотя</i> [х'ат'а]	A041	<i>Октябрьской</i> [ак'т'а'бр'ск'ый]	A042	<i>Тяга</i> [т'а'г'ъ]	A043	<i>Тянет</i> [т'ан'ът]

В таблицах могут быть пропущены аллофоны (тогда вместо слова указан символ «-»), это связано с тем, что данные аллофоны не употребляются в русской речи (например, безударный гласный «ё» или парный звонкий согласный в позиции перед паузой) или не были найдены в корпусе RUSLAN. В последнем случае отсутствующий аллофон при синтезе речи будет заменен другим, наиболее похожим на него и найденным в базе данных аллофонов. Так как аллофон, отсутствующий в речевом корпусе, скорее всего является очень редким для русской речи, это не будет оказывать существенное влияние на качество

синтезируемой речи. Реальное количество аллофонов при этом получается несколько меньше, чем теоретическое значение в 420 аллофонов.

3 Технологический раздел

3.1 Описание используемого стека технологий

В качестве языка программирования был выбран язык Python3 по следующим причинам:

- для python3 существует большое количество библиотек для обработки естественного языка (nltk, scikit-learn, rymorphy, natasha и т. д.);
- библиотеки снабжены подробной документацией и многочисленными примерами их использования;
- быстрая скорость разработки программ на python3.

Для хранения всех необходимых словарей (сокращений, ударений, согласных и фонетических исключений) необходимо выбрать способ. Так как необходимо хранить словари, отлично подойдут NoSQL базы данных типа «ключ-значение».

Одной из популярных БД типа «ключ-значение» является RocksDB, который предоставляет драйверы для большого количества языков программирования, в том числе и Python. В данной работе для хранения словарей будет использоваться RocksDB. Она обеспечит быстрый доступ к необходимым данным по ключу.

3.2 Структура хранения данных

Для поддержки разработанной системы синтеза речи необходимо пять хранилищ данных: четыре словаря и БД исходных элементов синтеза.

Исходные элементы хранятся в файловой системе в виде звуковых файлов. Структура этого хранилища представлена на рисунке 20. В корневой директории audio_db находятся пять поддиректорий, содержащих аллофоны гласных, согласных, а также три типа аллослогов.

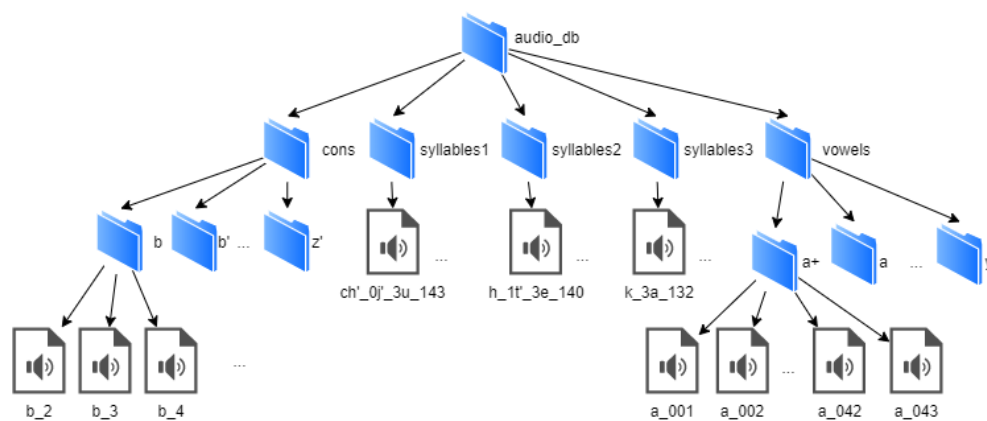


Рисунок 20 – Структура хранилища исходных элементов синтеза

Для хранения словарей используется NoSQL база данных типа «ключ-значение».

В БД ударений ключом является начальная форма слова, а значениями – все возможные формы этого слова с пометами ударений. Например, по ключу «аванпост» в БД хранится массив «аванпо'ст, аванпо'сты, аванпо'ста, аванпо'стов, аванпо'сту, аванпо'стам, аванпо'ст, аванпо'сты, аванпо'стом, аванпо'стами, аванпо'сте, аванпо'стах» (символом «'» обозначено ударение).

В БД аббревиатур ключом является сама аббревиатура. Значением является массив её свойств: произношение, склоняемость, род, расшифровка. Например, по ключу «НИУ» в БД хранится массив свойств: «[ни+у],нескл.,м.,Национальный исследовательский университет».

В БД фонетических исключений ключом является орфографически правильная форма слова, а значением – форма этого слова, но исправленная таким образом, что ее прочтение подчиняется стандартным правилам русского языка. Например, по ключу «альтернатива» в БД хранится значение «альтэрнатива».

В БД согласных ключом является согласная буква, а значением – ее произношение. Например, по ключу «в» в БД хранится значение «вэ».

3.3 Описание зависимостей

При разработке синтезатора речи были использованы следующие библиотеки для языка Python.

- pydub – библиотека для обработки звуковых файлов, в данной работе используется для склеивания исходных элементов синтеза в результирующий звуковой файл [16].
- nltk – набор библиотек для обработки естественного языка, в данной работе используется для разбиения предложений на слова [15].
- pymorphy2 – морфологический анализатор для русского языка, использует словари OpenCorpora, в данной работе используется для морфологической маркировки текста [6].
- Razdel – система токенизации предложений и слов в русском языке, являющаяся частью проекта Natasha. В данной работе используется для разбиения текста на предложения [20].
- rocksdb – библиотека для работы с базой данных RocksDB на языке Python. Используется для взаимодействия с базами данных аббревиатур, фонетических исключений, ударений и согласных [18].
- PyQt5 – графический фреймворк Qt для языка Python. Используется для реализации графического интерфейса пользователя [17].

3.4 Структура разработанного программного обеспечения

На рисунках 21 – 25 представлена диаграмма классов проекта в нотации UML. Диаграмма разбита на несколько блоков для удобства. На рисунке 21 представлена диаграмма классов блока «Лингвистический текстовый процессор». На рисунке 22 представлена диаграмма классов блока «Просодический процессор». На рисунке 23 представлена диаграмма классов блока «Фонетический процессор». На рисунке 24 представлена диаграмма

классов блока «Акустический процессор». На рисунке 25 представлена диаграмма классов всей системы преобразования текста в речь. Она включает в себя четыре блока: лингвистический текстовый процессор, просодический процессор, фонетический процессор и акустический процессор, классы которых были представлены на рисунках 21 – 24. Следует обратить внимание, что на диаграммах могут повторяться некоторые классы (Sentence, Allophone). Это связано с тем, что классы из нескольких блоков находятся с ними в отношениях зависимости.

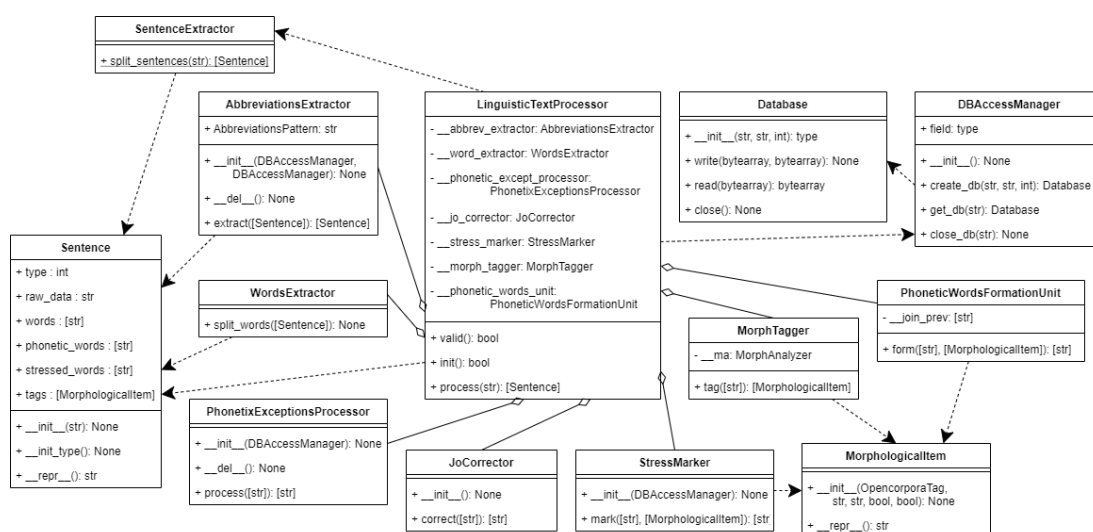


Рисунок 21 – Диаграмма классов лингвистического текстового процессора

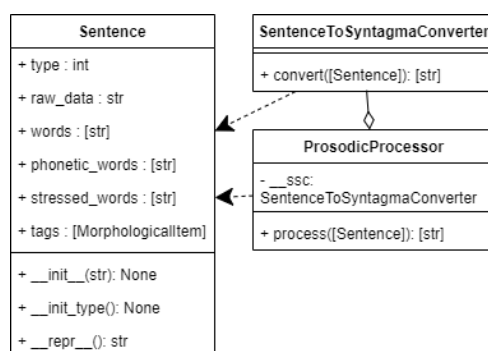


Рисунок 22 – Диаграмма классов просодического процессора

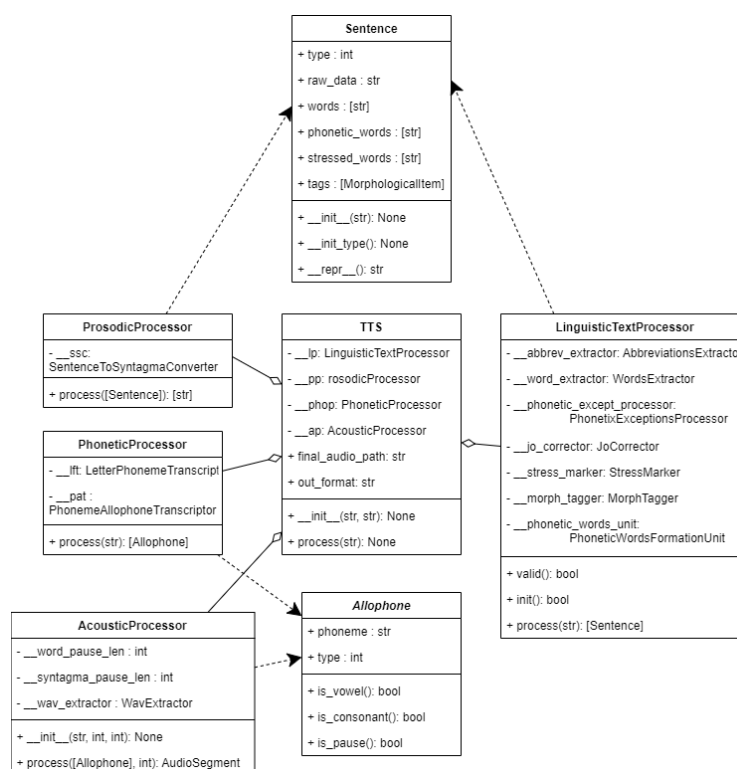


Рисунок 25 – Диаграмма классов системы преобразования текста в речь

В таблицах 9-13 кратко описано, для чего нужен каждый класс.

Таблица 9 – Классы блока «Лингвистический текстовый процессор»

Название класса	Описание
LinguisticTextProcessor	Предварительная обработка текста
SentenceExtractor	Деление текста на предложения
WordsExtractor	Деление предложений на слова
AbbreviationsExtractor	Поиск аббревиатур в предложениях и их дешифровка
PhonetixExceptionsProcessor	Поиск в предложениях слов – фонетических исключений – и замена их на эквиваленты
JoCorrector	Корректировка буквы «ё» в словах
StressMarker	Определение места ударения в словах
MorphTagger	Морфологическая разметка текста
PhoneticWordsFormationUnit	Формирование фонетических слов
MorphologicalItem	Хранение морфологической информации о слове
Database	Операции с БД (чтение, запись)
DBAccessManager	Соединение с БД (создание и разрыв соединения)

Таблица 10 – Классы блока «Просодический процессор»

Название класса	Описание
ProsodicProcessor	Выделение синтагм в тексте
SentenceToSyntagmaConverter	Выделение синтагм в предложении

Таблица 11 – Классы блока «Фонетический процессор»

Название класса	Описание
PhoneticProcessor	Преобразование орфографического текста в последовательность аллофонов
LetterPhonemeTranscriptor	Преобразование орфографического текста в последовательность фонем
PhonemeAllophoneTranscriptor	Преобразование последовательности фонем в последовательность аллофонов
Phoneme	Хранение информации о фонеме
Allophone	Хранение информации об аллофоне
ConsonantAllophone	Хранение информации об аллофоне-согласном
PauseAllophone	Хранение информации об аллофоне-паузе
VowelAllophone	Хранение информации об аллофоне-гласном

Таблица 12 – Классы блока «Акустический процессор»

Название класса	Описание
AcousticProcessor	Выбор необходимых элементов из базы данных исходных элементов и их склейка в общий выходной аудиофайл
WavExtractor	Поиск в БД необходимого звукового файла
SyllabicComplex1Unit	Деление последовательности аллофонов на слоговые комплексы первого типа
SyllabicComplex2Unit	Деление последовательности аллофонов на слоговые комплексы второго типа
SyllabicComplex3Unit	Деление последовательности аллофонов на слоговые комплексы третьего типа
SyllabicUnit	Родительский абстрактный класс для классов SyllabicComplex1Unit, SyllabicComplex2Unit, SyllabicComplex3Unit
OpenSyllableUnit	Деление последовательности аллофонов на открытые слоги

Таблица 13 – Классы системы преобразования текста в речь

Название класса	Описание
TTS	Преобразование текста в речь
Sentence	Хранение информации о предложении

3.5 Тестирование

При проведении тестирования в первую очередь необходимо показать работоспособность отдельных частей программы. Для этого используется модульное тестирование. Модульное тестирование (или unit-тестирование) – это метод тестирования программного обеспечения, при котором отдельные компоненты программы, называемые модулями, тестируются независимо со всеми необходимыми зависимостями [19].

Для тестирования разработанного программного обеспечения были разработаны unit-тесты для классов LetterPhonemeTranscriptor, PhonemeAllophoneTranscriptor, SyllabicComplex1Unit, SyllabicComplex2Unit, SyllabicComplex3Unit, OpenSyllableUnit.

Примеры входных данных и ожидаемых результатов тестов для класса LetterPhonemeTranscriptor приведены в таблице 14. Примеры тест-кейсов для классов PhonemeAllophoneTranscriptor, OpenSyllableUnit, SyllabicComplex1Unit, SyllabicComplex2Unit и SyllabicComplex3Unit приведены в таблицах 15 – 19 соответственно.

Полный набор тест-кейсов для всех пяти классов приведен в приложении Б. В словах, которые подаются как входные данные для тестов, могут встретиться орфографические ошибки, это сделано намеренно для тестирования различных ситуаций.

Все тесты были успешно пройдены.

Таблица 14 – Примеры входных данных и ожидаемых результатов тестов для класса LetterPhonemeTranscriptor

№	Входные данные	Ожидаемый результат
1	"ПРИНЁ+С_ИГРУ+ШКУ"	["p", "r", "i", "n", "o", "+", "s", "_", "y", "g", "r", "u", "+", "sh", "k", "u"]
2	"ОТЦА+"	["a", "c", "c", "a", "+"]
3	"ОБЪЁ+М"	["a", "b", "j", "o", "+", "m"]
4	"БА+НТИК"	["b", "a", "+", "n", "t", "i", "k"]
5	"НО+ЛЬ"	["n", "o", "+", "l"]
6	"У+ТКА"	["u", "+", "t", "k", "a"]
7	"ВВЁ+З"	["v", "v", "o", "+", "s"]

Таблица 15 – Примеры входных данных и ожидаемых результатов тестов для класса PhonemeAllophoneTranscriptor

№	Входные данные	Ожидаемый результат
1	["#", "u"]	["pause_1", "u_100"]
2	["_", "a", "_"]	["pause_0", "a_100", "pause_0"]
3	["#", "y", "#"]	["pause_1", "y_100", "pause_1"]
4	["p", "a", "+", "p"]	["p_4", "a_011", "p_0"]
5	["sh", "a", "+", "t"]	["sh_4", "a_022", "t_0"]
6	["k'", "a", "+", "t"]	["k'_4", "a_043", "t_0"]
7	["p'", "a", "+"]	["p'_4", "a_040"]

Таблица 16 – Примеры входных данных и ожидаемых результатов тестов для класса OpenSyllableUnit

№	Входные данные	Ожидаемый результат
1	"КО+ЛОСА"	[[["k", "o"], ["l", "a"], ["s", "a"]]]
2	"А+ММО"	[[["a"], ["m", "m", "a"]]]
3	"МНО+ГО"	[[["m", "n", "o"], ["g", "a"]]]
4	"О+М"	[[["o", "m"]]]
5	"КО+Л"	[[["k", "o", "l"]]]
6	"О+"	[[["o"]]]
7	"МАЙК"	[[["m", "a", "j", "k"]]]

Таблица 17 – Примеры входных данных и ожидаемых результатов тестов для класса SyllabicComplexUnit

№	Входные данные	Ожидаемый результат
1	"АЛЛОФО+Н"	[[["a", "l", "l", "a"], ["f", "o", "n"]]]
2	"АЯ+КС"	[[["a", "j", "a", "k", "s"]]]
3	"ТАКА+Я"	[[["t", "a"], ["k", "a", "j", "a"]]]
4	"ЛИ+НИЯ"	[[["l", "i"], ["n", "i", "j", "a"]]]
5	"Я+МА"	[[["j", "a"], ["m", "a"]]]
6	"РА=ДИОА=ЭРОНАВИГА+ЦИЯ"	[[["r", "a"], ["d", "i", "a", "a", "e"], ["r", "a"], ["n", "a"], ["v", "i"], ["g", "a"], ["c", "y", "j", "a"]]]
7	"НАИ+ВНЫЙ"	[[["n", "a", "i"], ["v", "n", "y", "j"]]]

Таблица 18 – Примеры входных данных и ожидаемых результатов тестов для класса SyllabicComplex2Unit

№	Входные данные	Ожидаемый результат
1	"МА+ЙКА"	[[["m", "a", "j"], ["k", "a"]]]
2	"МА+ЙНА"	[[["m", "a"], ["j", "n", "a"]]]
3	"ЖА+РКО"	[[["zh", "a", "r"], ["k", "a"]]]
4	"МА+ЙК"	[[["m", "a", "j", "k"]]]
5	"РМА+ДА"	[[["r", "m", "a"], ["d", "a"]]]
6	"АРМА+ДА"	[[["a", "r", "m", "a"], ["d", "a"]]]

Таблица 19 – Примеры входных данных и ожидаемых результатов тестов для класса SyllabicComplex3Unit

№	Входные данные	Ожидаемый результат
1	"КО+ЛОСА"	[[["k", "o", "l", "a"], ["s", "a"]]]
2	"КОЛО+СА"	[[["k", "a"], ["l", "o"], ["s", "a"]]]
3	"КО+ЛОН"	[[["k", "o"], ["l", "a", "n"]]]
4	"КО+ЛА"	[[["k", "o", "l", "a"]]]
5	"КОЛА+"	[[["k", "a"], ["l", "a"]]]

3.6 Сборка и запуск разработанного программного обеспечения

Процесс сборки и установки описан для операционной системы Linux. В системе должна быть установлена RocksDB.

Для запуска инструмента необходимо:

1. установить интерпретатор python3;
2. установить систему управления пакетами pip для python3;
3. установить необходимые зависимости, которые находятся в файле requirements.txt командой *pip install -r requirements.txt*.

Запуск тестов производится командой *python3 ./tests/main.py* из корневой директории проекта. В разработанном программном обеспечении предусмотрено два режима – графический и консольный. Запуск консольного режима производится командой *python3 ./main.py filename* из корневой директории проекта. Запуск графического режима производится командой *python3 ./main_int.py* из корневой директории проекта.

3.7 Руководство пользователя

Графический интерфейс пользователя представлен на рисунке 25. Предусмотрен как ручной ввод текста, так и выбор текстового файла. В процессе синтеза на интерфейсе отображена шкала прогресса и информация о том, что в данный момент выполняется системой синтеза речи. После завершения процесса синтеза пользователь имеет возможность прослушать получившийся аудиофайл.

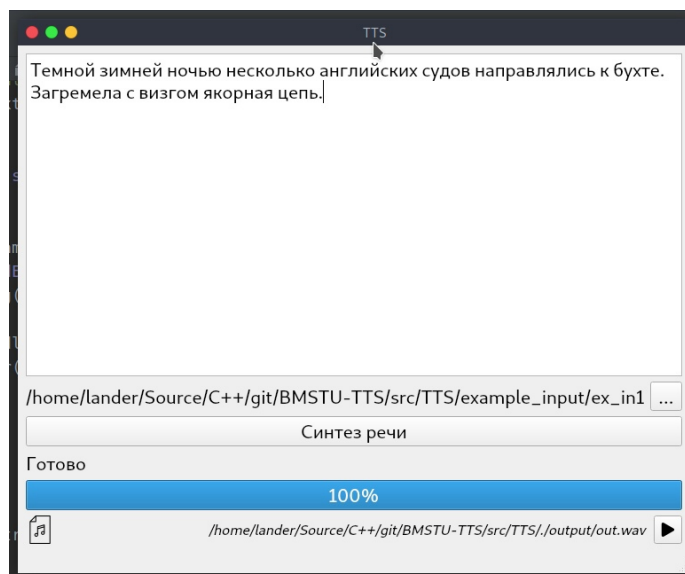


Рисунок 25 – Графический интерфейс пользователя

3.8 Руководство пользователя

Процесс использования ПО пользователями

Рассмотрим процесс использования разработанного ПО пользователями. Пользователь имеет возможность использовать инструмент в графическом и в консольном режимах. При использовании консольного режима для синтеза речи по тексту необходимо выполнить следующие шаги.

1. Подготовить входной текстовый файл с текстом, который необходимо озвучить. Проверить, что в файле корректно проставлена пунктуация и отсутствуют запрещенные символы.

2. Запустить скрипт *main.py* в корневой директории проекта с указанием пути к файлу, подготовленному на предыдущем шаге:
python3 ./main.py filename.
3. Результат работы программы – аудиофайл – будет расположен в директории */output* относительно корневой директории проекта и называться *out.wav*. Для того, чтобы прослушать получившийся файл, на компьютере должен быть установлен проигрыватель.

Использование графического режима предполагает следующие шаги (см. рисунок 26).

- Запустить скрипт *main_int.py* в корневой директории проекта:
python3 ./main.py.
- Ввод текста в поле (1) или выбор текстового файла кнопкой (2) (имя выбранного файла будет отображаться в поле (3)).
- После завершения предыдущего шага необходимо нажать на кнопку «Синтез речи» (4).
- После нажатия на кнопку появится шкала прогресса (6), на которой отображено, какая часть работы уже выполнена, и строка состояния (5), в которой отображается информация о том, что в данный момент делает система синтеза речи.
- После завершения синтеза появится поле (8) с названием результирующего файла, где была сохранена синтезированная речи. Результат можно прослушать, нажав на кнопку (7). Во время воспроизведения речи нажатием на эту же кнопку можно остановить воспроизведение.

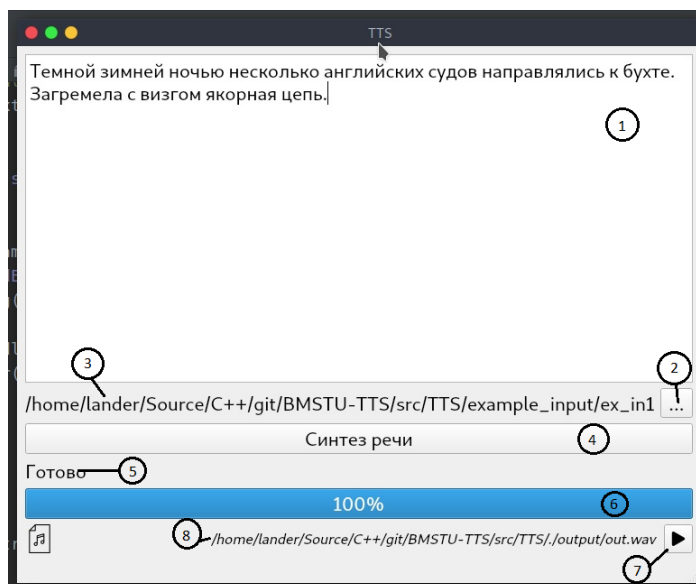


Рисунок 26 – Использование графического интерфейса пользователя

Процесс использования ПО другими приложениями

Рассмотрим процесс использования разработанного ПО другими приложениями. Необходимо выполнить следующие шаги.

1. Запустить скрипт *main.py*, находящийся в корневой директории проекта. Первым аргументом необходимо передать путь до файла, в котором находится текст, который нужно озвучить.
2. В результате успешной работы появится файл */output/out.wav* (относительно корневой директории проекта). В случае ошибки скрипт завершится с кодом возврата, отличным от нуля.

4 Исследовательский раздел

4.3 Оценка качества синтеза речи

Оценка качества синтезируемой речи является очень важной задачей. При оценке обычно используют четыре основных критерия:

- разборчивость;
- естественность;
- мультимодальность (эмоциональность);
- многоязычие (возможность синтеза речи на нескольких языках) [8].

В связи с тем, что не ставились задачи разработки многоязычного синтезатора речи и разработки синтезатора речи с просодическим оформлением, будем рассматривать только два первых критерия.

Разборчивость является основным критерием при оценке качества синтезируемой речи. Существует несколько способов оценки разборчивости:

- звуковая разборчивость — оценивается количество правильно воспринятых (бессмысленных) звуков;
- слоговая разборчивость — оценивается количество правильно воспринятых (бессмысленных) слогов;
- словесная разборчивость — оценивается количество правильно воспринятых никак не связанных между собой слов;
- фразовая разборчивость — оценивается количество правильно распознанных фраз;
- смысловая разборчивость — оценивается понимание содержания речи.

При оценке слушателям предлагается распознать соответственно звуки, слоги, слова, фразы или смысл того, что «сказала» машина.

Фразовая разборчивость должна быть на уровне 98-100%. При оценке словесной разборчивости требуется результат не менее 99% [8].

При оценке смысловой разборчивости речи выделяют пять градаций качества [1]:

- отлично – полное понимание речи, отсутствие переспросов;
- хорошо – понимание речи полное, но возможны переспросы необычных слов, фамилий и терминов;
- удовлетворительно – переспросы отдельных слов и фраз, но в целом восприятие речевой информации правильное;
- неудовлетворительно – отдельные слова непонятны даже при переспросе;
- плохо – смысл речевой информации понимается с трудом.

Естественность или натуральность речи является субъективной характеристикой, для её оценки нет объективных критериев. В то же время она зависит от многих факторов, таких как натуральность самих звуков, ритмическая сторона речи, интонация (может быть «роботная»), правильность постановки ударений и т. д. Оценка естественности сильно зависит от субъективного впечатления слушателя.

При оценке естественности речи можно выделить следующие градации по баллам [1]:

- 4,6 – 5,0 баллов – естественность звучания речи, полное отсутствие искажений;
- 4,0 – 4,5 – естественность звучания речи, малозаметные искажения;
- 3,5 – 3,9 – естественность звучания речи, слабое постоянное присутствие искажений;
- 3,0 – 3,4 – незначительное нарушение естественности, заметное присутствие искажений;
- 2,5 – 2,9 – заметное нарушение естественности, присутствие искажений или помех;
- 1,7 – 2,4 – существенное искажение естественности, постоянное присутствие искажений или помех;

- $< 1,7$ – сильные искажения, механический голос, потеря естественности.

4.4 Проведение исследования

Будем проводить исследование фразовой и смысловой на семи контрольных примерах:

- «Темной зимней ночью несколько английских судов направлялись к бухте» (1);
- «Загремела с визгом якорная цепь» (2);
- «Молния блистала все чаще и ярче» (3);
- «На пароходе зажгли электричество» (4);
- «Они пошли в гостиную к роялю» (5);
- «Жара понемногу спадала» (6);
- «Жарко от летнего солнца и от теплой земли» (7).

Необходимо оценить фразовую и смысловую разборчивость и естественность озвученных фраз. К примеру, разобьем первую фразу на слоговые комплексы трех типов:

- [[[[[t'_4, o_041]]], [[m_2, n_3, a_123, j'_0]]], [[[z'_4, i_041]]], [[m_2, n'_3, e_143, j'_0]]], [[[n_3, a_123]]], [[ch'_0, j'_3, u_143]]], [[[n'_4, e_042]]], [[s_1, k_3, a_133, l'_1]], [[k_3, a_132]]], [[[a_132, n_2, g_2, l'_4, i_043, j'_1]]], [[s_1, k'_3, i_142, h_0]]], [[[s_4, u_022]]], [[d_4, o_021, v_0]]], [[[n_3, a_121]]], [[p_2, r_3, a_121]]], [[v_2, l'_4, a_043]]], [[l'_3, i_143, z'_0]]], [[[g_2, b_4, u_012]]], [[h_1, t'_3, e_140]]]] – слоговые комплексы первого типа;
- [[[[[t'_4, o_041]], [[m_2, n_3, a_123, j'_0]]], [[[z'_4, i_041]]], [[m_2, n'_3, e_143, j'_0]]], [[[n_3, a_123]]], [[ch'_0, j'_3, u_143]]], [[[n'_4, e_042]]], [[s_1, k_3, a_133, l'_1]], [[k_3, a_132]]], [[[a_132, n_2, g_2, l'_4, i_043, j'_1]]], [[s_1, k'_3, i_142, h_0]]], [[[s_4, u_022]]], [[d_4, o_021, v_0]]],

[[[n_3, a_121]], [[p_2, r_3, a_121]], [[v_2, l'_4, a_043]], [[l'_3, i_143, z'_0]]], [[[g_2, b_4, u_012]], [[h_1, t'_3, e_140]]]] – слоговые комплексы второго типа;

- [[t'_4, o_041], [m_2, n_3, a_123, j'_0]], [z'_4, i_041], [m_2, n'_3, e_143, j'_0]], [[n_3, a_123], [ch'_0, j'_3, u_143]], [[n'_4, e_042], [s_1, k_3, a_133, l'_1], [k_3, a_132]], [[a_132, n_2, g_2, l'_4, i_043, j'_1], [s_1, k'_3, i_142, h_0]], [[s_4, u_022], [d_4, o_021, v_0]], [[n_3, a_121], [p_2, r_3, a_121], [v_2, l'_4, a_043], [l'_3, i_143, z'_0]], [[g_2, b_4, u_012], [h_1, t'_3, e_140]]] – слоговые комплексы третьего типа.

Как можно увидеть из примера, разбиение фразы на слоговые комплексы различных типов практически совпадает (за исключением нескольких слогов), поэтому разница в качестве речи, синтезированной с использованием различных слоговых комплексов, будет невелика.

Для исследования качеств синтезируемой речи был приглашен независимый слушатель, которому предлагалось прослушать контрольные примеры.

При использовании только аллофонов слушатель фразовая разборчивость оценивается как 3 фразы из 7 (42%), что является неприемлемым результатом. При этом смысловая разборчивость оценивается как неудовлетворительная – отдельные слова непонятны даже при переспросе. Такую речь тяжело понимать и приходится из контекста догадываться, что было сказано. Речь характеризуется сильными искажениями, механическим голосом, потерей естественности (< 1,7 баллов).

При использовании слоговых комплексов фразовая разборчивость оценивается как 7 фраз из 7 (100%). При этом смысловая разборчивость отличная – полное понимание смысла фразы без переспросов. Речь стало намного легче воспринимать и отпала необходимость «достраивать» фразу из контекста. Речь характеризуется заметным присутствием искажений, однако естественность нарушена незначительно (3,0 – 3,4 балла).

Результаты исследования приведены в таблице 20.

Таблица 20 – Исследование фразовой и смысловой разборчивости

Исходные элементы синтеза	Фразовая разборчивость	Смысловая разборчивость	Словесная разборчивость	Естественность
Аллофоны	42%	Неудовлетворительная	28%	< 1,7 балла
Аллофоны+ аллослоги	100%	Отличная	100%	3,0 – 3,4 балла

Теперь оценим словесную разборчивость еще на нескольких примерах – словах, встречающихся в речи не очень часто:

- шаромыжник;
- свентицкой;
- смоковников;
- геннисон;
- энниок;
- эбергаль;
- коломб.

Независимому слушателю было предложено прослушать семь слов по отдельности.

При использовании в синтезе только аллофонов неподготовленному слушателю удалось опознать лишь 2 из 7 слов (28%), что говорит об очень низкой разборчивости речи, такой результат неприемлем. При использовании слоговых комплексов любых типов слушатель смог опознать все слова (7 из 7, 100%).

На рисунке 27 представлены процентные соотношения фразовой и словесной разборчивости при использовании различных исходных элементов синтеза.

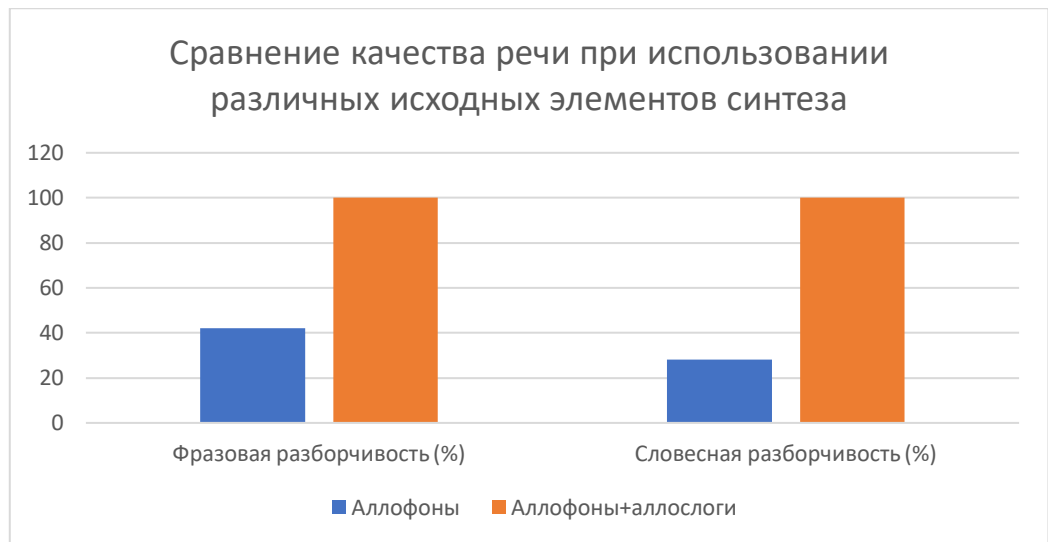


Рисунок 27 – Сравнение качества речи при использовании различных исходных элементов синтеза

Что касается объема базы данных исходных элементов, то при аллофонном синтезе с использованием мини-набора акустическая БД должна содержать 327 элементов синтеза (не 420, потому что не все аллофоны используются в языке). При использовании аллослогов третьего типа (легко сегментируемых) необходимо 433 элемента. На рисунке 28 представлена диаграмма для сравнения размеров БД при использовании различных исходных элементов синтеза. При этом при подготовке БД аллослогов были созданы только слоги, которые содержатся в контрольных примерах (7 фраз и 7 слов). Таким образом, можно сделать вывод, что увеличение качества синтезируемой речи посредством использования аллослогов приводит к значительному возрастанию объемов акустической базы данных.

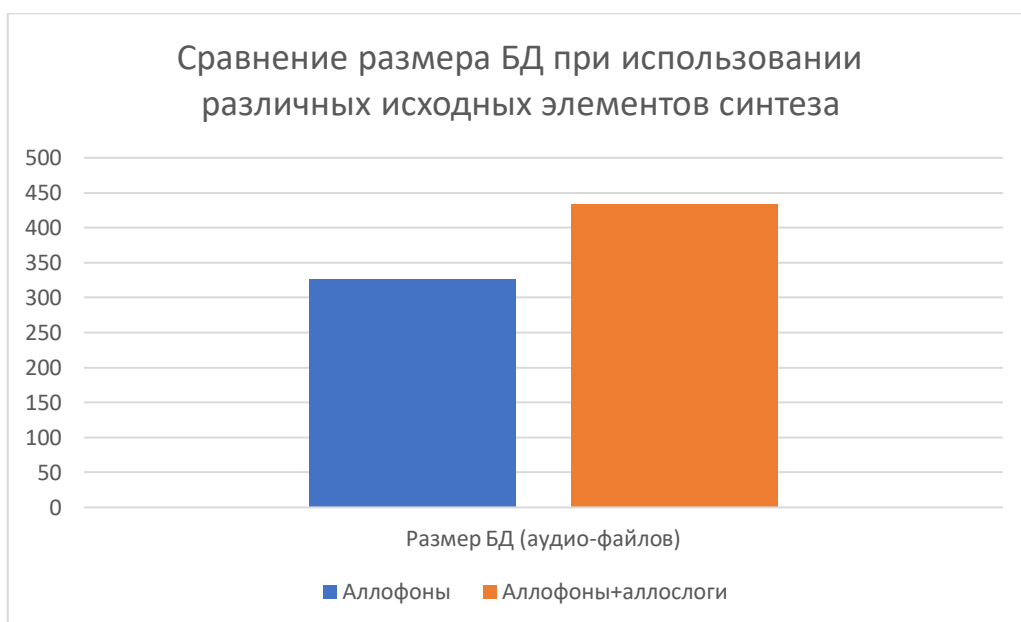


Рисунок 28 – Сравнение размера БД при использовании различных исходных элементов синтеза

4.5 Вывод

На основании проведенного исследования можно сделать вывод, что использование мини-набора аллофонов в синтезе дает недостаточно разборчивую речь из-за резких переходов между «склеенными» аллофонами. При использовании слоговых комплексов хотя бы третьего типа качество речи существенно повышается. При этом, однако, повышается и объем базы данных исходных элементов. Ценой разборчивой речи при синтезе по фонетическим правилам является большой объем БД. Необходимо найти баланс между объемом БД и качеством синтеза.

В рамках данной работы БД исходных элементов создавалась вручную, то есть проводилась «нарезка» каждого отдельного элемента БД из речевого корпуса. По этой причине создание БД большого объема было затруднительно, однако разработка инструмента автоматической разметки устных корпусов могла бы упростить этот процесс и, соответственно, дать возможность создавать качественные синтезаторы речи ценой меньших усилий.

ЗАКЛЮЧЕНИЕ

В результате выполненной работы были решены следующие задачи:

- было проведено аналитическое исследование существующих подходов к синтезу речи и выбран наиболее подходящий;
- на основе проведенного исследования был модифицирован выбранный метод для достижения поставленной цели;
- был разработан алгоритм по модифицированному методу;
- было создано программное обеспечение, реализующее данный алгоритм;
- были разработаны тесты для разработанного ПО;
- было проведено исследование разработанного ПО для оценки качества синтезируемой речи.

Была достигнута поставленная цель: был разработан алгоритм для синтеза речи по произвольному тексту на русском языке и создано ПО для озвучивания текста на основе этого алгоритма. На основе проведенного аналитического исследования был модифицирован метод аллофонного синтеза речи по фонетическим правилам, что привело к значительному улучшению качества озвучивания текстов.

Разработанное программное обеспечение было протестировано и прошло все тесты.

В качестве дальнейшего развития проекта можно предложить следующее:

- убрать ограничения на входной текст, а именно, добавить возможность озвучивать иностранные слова, специальные символы, многоразрядные и дробные числа, телефонные номера, сокращения, даты из римских цифр, обозначения времени и даты, интернет-адреса, математические выражения;
- добавить поддержку пользовательской разметки (ударения и просодическая разметка);

- разработать или интегрировать синтаксический и семантический анализатор для более качественного деления текста на синтагмы и для возможности озвучивать ограничения, приведенные в первом пункте;
- добавить интонационное оформление текста.

Разработанное программное обеспечение может быть использовано для создания систем разговорного языка и для упрощения взаимодействия человек-машина. Это позволит повысить доступность компьютеров и автоматизированных систем для широкого круга пользователей, а также повысить коммуникативные возможности слабослышащих, глухих, немых людей. Разработанное ПО может быть использовано в различных отраслях при создании голосовых роботов.

СПИСОК ИСПОЛЬЗОВАННЫХ ИСТОЧНИКОВ

1. ГОСТ Р 51061-97. Системы низкоскоростной передачи речи по цифровым каналам. Параметры качества речи и методы измерений. Принят и введен в действие Постановлением Госстандарта России от 24 июня 1997 г. №223: дата введения 01.01.1998. – М.: ИПК Издательство стандартов, 1997.
2. Интерполирующие цифровые фильтры [Электронный ресурс]. – Режим доступа: <https://digteh.ru/digital/Intrpltr.php>, свободный (дата обращения 18.05.2021).
3. Князев, С. В. Современный русский литературный язык: Фонетика, орфоэпия, графика и орфография: Учебное пособие для вузов. – 2-е изд., перераб. и доп. / С. В. Князев, С. К. Пожарицкая. – М.: Академический Проект; Гаудеамус, 2011. – 430 с.
4. Лингвистический энциклопедический словарь. [Электронный ресурс]. – Режим доступа: <http://tapemark.narod.ru/les/>, свободный (дата обращения 03.05.2021).
5. Лобанов, Б. М. Компьютерный синтез и клонирование речи. / Б. М. Лобанов, Л. И. Цирульник. – Минск: «Белорусская наука», 2008.
6. Морфологический анализатор pymorphy2. [Электронный ресурс]. – Режим доступа: <https://pymorphy2.readthedocs.io/en/stable/>, свободный (дата обращения 27.05.2021).
7. Правила разметки речевого корпуса на фонетические сегменты и стратегия выбора элементов компиляции при синтезе речи. [Электронный ресурс]. – Режим доступа: <http://www.dialog-21.ru/digests/dialog2007/materials/...>, свободный (дата обращения 03.05.2021).
8. Рыбин, С. В. Синтез речи. Учебное пособие по дисциплине «Синтез речи». / С. В. Рыбин. – СПб: Университет ИТМО, 2014. – 92 с.
9. Фонетико-акустическая база данных для многоязычного синтеза речи по тексту на славянских языках [Электронный ресурс]. – Режим доступа:

<http://www.dialog-21.ru/digests/dialog2006/materials/..>, свободный (дата обращения 31.01.2021).

10. Adobe Audition [Электронный ресурс]. – Режим доступа: <https://www.adobe.com/ru/products/audition.html>, свободный (дата обращения 01.06.2021).

11. Amazon Polly [Электронный ресурс]. – Режим доступа: <https://aws.amazon.com/ru/polly/>, свободный (дата обращения 29.01.2021).

12. Cloud Text-to-Speech [Электронный ресурс]. – Режим доступа: <https://cloud.google.com/text-to-speech/>, свободный (дата обращения 29.01.2021).

13. Huang, Xuedong. Spoken language processing: a guide to theory, algorithm, and system development. / Xuedong Huang, Alex Acero, Hsiao-Wuen Hon. – Upper Saddle River, New Jersey: Prentice Hall PRT, 2001.

14. Microsoft Azure. Преобразование текста в речь [Электронный ресурс]. – Режим доступа: <https://azure.microsoft.com/ru-ru/services/cognitive-...>, свободный (дата обращения 29.01.2021).

15. Natural Language Toolkit. [Электронный ресурс]. – Режим доступа: <https://www.nltk.org/>, свободный (дата обращения 27.05.2021).

16. Pydub. [Электронный ресурс]. – Режим доступа: <https://github.com/jiaaro/pydub>, свободный (дата обращения 27.05.2021).

17. PyQt5. [Электронный ресурс]. – Режим доступа: <https://python-riverbankcomputing.com/software/pyqt/>, свободный (дата обращения 27.05.2021).

18. Python-rocksdb. [Электронный ресурс]. – Режим доступа: <https://python-rocksdb.readthedocs.io/en/latest/>, свободный (дата обращения 27.05.2021).

19. Pajankar, A. Python Unit Test Automation: Practical Techniques for Python Developers and Testers. / Ashwin Pajankar. – Nashik, Maharashtra, India, 2017.

20. Razdel. [Электронный ресурс]. – Режим доступа: <https://github.com/natasha/razdel>, свободный (дата обращения 27.05.2021).

21. RUSLAN: Russian Spoken Language Corpus For Speech Synthesis [Электронный ресурс]. – Режим доступа: <https://ruslan-corpus.github.io/>, свободный (дата обращения 10.02.2021).

22. Watson Text to Speech [Электронный ресурс]. – Режим доступа: <https://www.ibm.com/ru-ru/cloud/watson-text-to-speech>, свободный (дата обращения 29.01.2021).

ПРИЛОЖЕНИЕ А

Таблица А.1 – Мини-набор аллофонов для ударной гласной А

Третий индекс, правый контекст		0		1		2		3	
		пауза		п, ф, б, в, м		т, ц, с, д, з, н, л, ш, ж, р, к, х, г, у, о, а, э, ы		к', х', г', й, ч', ш', р', т', с', д', з', н', л', п', ф', б', в', м', и	
0	пауза	A000	<u>А</u> [а]	A001	<u>А</u> втор [а́фтър]	A002	<u>А</u> да [а́дъ]	A003	<u>А</u> ся [а́с'ь]
1	п, ф, б, в, м	A010	<u>Судьба</u> [суд'ба́]	A011	<u>Ба</u> ба [ба́бъ]	A012	<u>Хват</u> кой [хв́аткый]	A013	<u>Ба</u> тя [ба́т'ь]
2	ш, ж, р, т, ц, с, д, з, н, л	A020	<u>Еда</u> [и́да]	A021	<u>За</u> пад [за́път]	A022	<u>Да</u> та [да́тъ]	A023	<u>Та</u> ся [та́с'ь]
3	к, х, г, у, о, а, э, ы	A030	<u>Нога</u> [на́га]	A031	<u>Лега</u> вый [л'и́га́вый]	A032	<u>Ска</u> зка [ска́скъ]	A033	<u>Га</u> ля [га́л'ь]
4	т', с', д', з', н', л', ч', ш', р', п', ф', б', в', м', к', х', г', й, и	A040	<u>Хотя</u> [хат'а́]	A041	<u>Октя</u> брьской [акт'а́бр'скый]	A042	<u>Тя</u> га [т'а́гъ]	A043	<u>Тя</u> нет [т'ан'ьт]

Таблица А.2 – Мини-набор аллофонов для безударной гласной А

Третий индекс, правый контекст		0		1		2		3	
		пауза		п, ф, б, в, м		т, ц, с, д, з, н, л, ш, ж, р, к, х, г, у, о, а, э, ы		к', х', г', й, ч', ш', р', т', с', д', з', н', л', п', ф', б', в', м', и	
0	пауза	A100	-	A101	<u>А</u> втомобиль [а́фтмаб'и́л']	A102	<u>А</u> тмосфера [атмасф'ръ]	A103	<u>А</u> лименты [ал'им'нты]
1	п, ф, б, в, м	A110	<u>Ястреба</u> [ја́стр'ьбъ]	A111	<u>Трупа</u> м [тру́пъм]	A112	<u>Пасса</u> жир [пъсажы́р]	A113	<u>Спаси</u> бо [спас'ибъ]
2	ш, ж, р, т, ц, с, д, з, н, л	A120	<u>Работа</u> [рабо́тъ]	A121	<u>Забо</u> лел [зъбал'ёл]	A122	<u>Сделано</u> [зд'ёлънъ]	A123	<u>Остальное</u> [астал'но́ь]
3	к, х, г, у, о, а, э, ы	A130	<u>Филолога</u> [ф'ило́льгъ]	A131	<u>Капу</u> ста [капу́стъ]	A132	<u>Сканда</u> лист [скънда́л'и́ст]	A133	<u>Камер</u> тон [към'ирто́н]
4	т', с', д', з', н', л', ч', ш', р', п', ф', б', в', м', к', х', г', й, и	A140	<u>Время</u> [вр'ём'ь]	A141	<u>Слабостя</u> м [сла́бъс'т'ьм]	A142	<u>Лягну</u> ть [л'игну́тъ]	A143	<u>Взгля</u> ните [взгл'ин'и́тъ]

ПРОДОЛЖЕНИЕ ПРИЛОЖЕНИЯ А

Таблица А.3 – Мини-набор аллофонов для ударной гласной U

Третий индекс, правый контекст		0		1		2		3	
		пауза		п, ф, б, в, м		т, ц, с, д, з, н, л, ш, ж, р, к, х, г, у, о, а, э, ы		к', х', г', й, ч', ш', р', т', с', д', з', н', л', п', ф', б', в', м', и	
0	пауза	U000	-	U001	<i>Умным</i> [умнѣм]	U002	<i>Утром</i> [утрѣм]	U003	<i>Уйма</i> [уймѣ]
1	п, ф, б, в, м	U010	<i>Борьбу</i> [бар'бѹ]	U011	<i>Двум</i> [двѹм]	U012	<i>Перепутать</i> [п'ѣр'ипутѣт']	U013	<i>Буйно</i> [бѹйнѣ]
2	ш, ж, р, т, ц, с, д, з, н, л	U020	<i>Году</i> [гадѹ]	U021	<i>Крупная</i> [крѹпнѣѣ]	U022	<i>Шуткой</i> [шуткѣй]	U023	<i>Руки</i> [рук'и]
3	к, х, г, у, о, а, э, ы	U030	<i>Могу</i> [магѹ]	U031	<i>Скумбрия</i> [скѹмбр'ѣѣ]	U032	<i>Искусственно</i> [искѹств'ѣнѣ]	U033	<i>Купишь</i> [куп'ѣш]
4	т', с', д', з', н', л', ч', ш', р', п', ф', б', в', м', к', х', г', й, и	U040	<i>Звоню</i> [зван'ѹ]	U041	<i>Костюму</i> [кас'т'ѹму]	U042	<i>Повсюду</i> [пафс'ѹду]	U043	<i>Люди</i> [л'ѹд'и]

Таблица А.4 – Мини-набор аллофонов для безударной гласной U

Третий индекс, правый контекст		0		1		2		3	
		пауза		п, ф, б, в, м		т, ц, с, д, з, н, л, ш, ж, р, к, х, г, у, о, а, э, ы		к', х', г', й, ч', ш', р', т', с', д', з', н', л', п', ф', б', в', м', и	
0	пауза	U100	-	U101	<i>Упрёком</i> [упр'ѳкъм]	U102	<i>Упражняюсь</i> [упражн'ѣѣс']	U103	<i>Упиваясь</i> [уп'ивѣѣс']
1	п, ф, б, в, м	U110	<i>Голову</i> [голѣву]	U111	<i>Бумаги</i> [бумѣг'и]	U112	<i>Пустом</i> [пустѳм]	U113	<i>Путеводный</i> [пут'ивѳднѣй]
2	ш, ж, р, т, ц, с, д, з, н, л	U120	<i>Нашу</i> [нѣшу]	U121	<i>Ступайте</i> [ступѣйт'ѣ]	U122	<i>Журнал</i> [журнѣл]	U123	<i>Хорошую</i> [харѳшѹѹу]
3	к, х, г, у, о, а, э, ы	U130	<i>Руку</i> [рукѹ]	U131	<i>Кувыркком</i> [кувыркѳм]	U132	<i>Куском</i> [кусѳм]	U133	<i>Накупила</i> [нѣкуп'илѣ]
4	т', с', д', з', н', л', ч', ш', р', п', ф', б', в', м', к', х', г', й, и	U140	<i>Помню</i> [пѳмн'ѹ]	U141	<i>Дюма</i> [д'ѹмѣ]	U142	<i>Сюда</i> [с'ѹдѣ]	U143	<i>Авантюрист</i> [авѣн'т'ѹр'ѣст]

ПРОДОЛЖЕНИЕ ПРИЛОЖЕНИЯ А

Таблица А.5 – Мини-набор аллофонов для ударной гласной О

Третий индекс, правый контекст		0		1		2		3	
		пауза		п, ф, б, в, м		т, ц, с, д, з, н, л, ш, ж, р, к, х, г, у, о, а, э, ы		к', х', г', й, ч', ш', р', т', с', д', з', н', л', п', ф', б', в', м', и	
0	пауза	0000	<u>О</u> [о]	0001	<u>Образом</u> [о́брьзъм]	0002	<u>Острая</u> [о́стръь]	0003	<u>Очень</u> [о́ч'ьн']
1	п, ф, б, в, м	0010	<u>Письмо</u> [п'ис'мо́]	0011	<u>Помощью</u> [пóмъи'йу]	0012	<u>Помо́ста</u> [памóста]	0013	<u>Помо́щник</u> [памóш'н'ьк]
2	ш, ж, р, т, ц, с, д, з, н, л	0020	<u>Хорошо</u> [хърашó]	0021	<u>Особо</u> [асóбъ]	0022	<u>Смешон</u> [см'ишóн]	0023	<u>Король</u> [карóл']
3	к, х, г, у, о, а, э, ы	0030	<u>Высоко</u> [въсакó]	0031	<u>Телескоп</u> [т'ьл'искóп]	0032	<u>Скот</u> [скóт]	0033	<u>Ско́лько</u> [скóл'къ]
4	т', с', д', з', н', л', ч', ш', р', п', ф', б', в', м', к', х', г', й, и	0040	<u>Всё</u> [фс'ó]	0041	<u>Тёмный</u> [т'óмнъй]	0042	<u>Утесы</u> [ут'óсы]	0043	<u>Счетчике</u> <u>Огонечек</u> [аган'óч'ьк]

Таблица А.6 – Мини-набор аллофонов для безударной гласной О

Третий индекс, правый контекст		0		1		2		3	
		пауза		п, ф, б, в, м		т, ц, с, д, з, н, л, ш, ж, р, к, х, г, у, о, а, э, ы		к', х', г', й, ч', ш', р', т', с', д', з', н', л', п', ф', б', в', м', и	
0	пауза	0100	<u>О</u> [о]	0101	<u>Опусти́в</u> [апус'т'и́ф]	0102	<u>Особо</u> [асóбъ]	0103	<u>Опи́санного</u> [ап'и́сьнъвъ]
1	п, ф, б, в, м	0110	<u>Особо</u> [асóбъ]	0111	<u>Поправив</u> [пaпpáв'ьф]	0112	<u>Фонарь</u> [фанáръ']	0113	<u>Побежал</u> [пъб'ижáл]
2	ш, ж, р, т, ц, с, д, з, н, л	0120	<u>Скоро</u> [скóръ]	0121	<u>Пропадала</u> [пръпада́лъ]	0122	<u>Хорошо</u> [хърашó]	0123	<u>Игроки</u> [играк'и́]
3	к, х, г, у, о, а, э, ы	0130	<u>Ско́лько</u> [скóл'къ]	0131	<u>Епископ</u> [йип'и́скъп]	0132	<u>Хорошо</u> [хърашó]	0133	<u>Копиру́я</u> [кап'и́ръь]
4	т', с', д', з', н', л', ч', ш', р', п', ф', б', в', м', к', х', г', й, и	0140	-	0141	-	0142	-	0143	-

ПРОДОЛЖЕНИЕ ПРИЛОЖЕНИЯ А

Таблица А.7 – Мини-набор аллофонов для ударной гласной I

Третий индекс, правый контекст		0		1		2		3	
		пауза		п, ф, б, в, м		т, ц, с, д, з, н, л, ш, ж, р, к, х, г, у, о, а, э, ы		к', х', г', й, ч', ш', р', т', с', д', з', н', л', п', ф', б', в', м', и	
Второй индекс, левый контекст									
0	пауза	I000	<u>И</u> [и]	I001	<u>Им</u> [им]	I002	<u>Исповеди</u> [испѣв'ѣд'и]	I003	<u>Имя</u> [им'ѣ]
1	п, ф, б, в, м	I010	-	I011	-	I012	-	I013	-
2	ш, ж, р, т, ц, с, д, з, н, л	I020	-	I021	-	I022	-	I023	-
3	к, х, г, у, о, а, э, ы	I030	<u>Свои</u> [свай]	I031	<u>Своим</u> [свайм]	I032	<u>Воинственным</u> [вайнств'ѣнѣм]	I033	<u>Своими</u> [свайми]
4	т', с', д', з', н', л', ч', ш', р', п', ф', б', в', м', к', х', г', й, и	I040	<u>Провести</u> [прѣв'ис'т'и]	I041	<u>Учтливости</u> [уч'т'ивѣс'т'и]	I042	<u>Аппетитом</u> [ап'ит'итѣм]	I043	<u>Типичный</u> [т'ип'ич'нѣй]

Таблица А.8 – Мини-набор аллофонов для безударной гласной I

Третий индекс, правый контекст		0		1		2		3	
		пауза		п, ф, б, в, м		т, ц, с, д, з, н, л, ш, ж, р, к, х, г, у, о, а, э, ы		к', х', г', й, ч', ш', р', т', с', д', з', н', л', п', ф', б', в', м', и	
Второй индекс, левый контекст									
0	пауза	I100	-	I101	<u>Импозантно</u> [импозантнѣ]	I102	<u>Итак</u> [итак]	I103	<u>Имею</u> [им'ѣйу]
1	п, ф, б, в, м	I110	-	I111	-	I112	-	I113	-
2	ш, ж, р, т, ц, с, д, з, н, л	I120	-	I121	-	I122	-	I123	-
3	к, х, г, у, о, а, э, ы	I130	<u>Перебои</u> [п'ѣр'ибѣи]	I131	<u>Устроим</u> [устроим]	I132	<u>Стоит</u> [стоит]	I133	<u>Покоились</u> [пакѣил'ѣс']
4	т', с', д', з', н', л', ч', ш', р', п', ф', б', в', м', к', х', г', й, и	I140	<u>Легкости</u> [л'ѣхкѣс'т'и]	I141	<u>Против</u> [прѣт'ѣф]	I142	<u>Ответит</u> [атѣ'ѣт'ѣт]	I143	<u>Типичный</u> [т'ип'ич'нѣй]

ПРОДОЛЖЕНИЕ ПРИЛОЖЕНИЯ А

Таблица А.9 – Мини-набор аллофонов для ударной гласной Е

Третий индекс, правый контекст		0		1		2		3	
		пауза		п, ф, б, в, м		т, ц, с, д, з, н, л, ш, ж, р, к, х, г, у, о, а, э, ы		к', х', г', й, ч', ш', р', т', с', д', з', н', л', п', ф', б', в', м', и	
0	пауза	E000	-	E001	<i>Эбергайля</i> <i>Эдфу</i> [étfu]	E002	<i>Этом</i> <i>Экая</i> <i>Эхом</i> [эхъм]	E003	<i>Этим</i> [ét'ъм] <i>Эй</i>
1	п, ф, б, в, м	E010	-	E011	-	E012	<i>Кармэн</i>	E013	-
2	ш, ж, р, т, ц, с, д, з, н, л	E020	<i>Шоссе</i> [шас:э]	E021	<i>Ржевского</i> [ржэ́фскэвэ]	E022	<i>С этой</i> <i>Генуэзцах</i> [г'ьнуэ́сцэх]	E023	<i>Модель</i> [мадэ́л']
3	к, х, г, у, о, а, э, ы	E030	-	E031	-	E032	<i>Силуэт</i> [с'ьлуэ́т] <i>Поэтому</i> <i>Менуэт</i> <i>Кэт</i>	E033	-
4	т', с', д', з', н', л', ч', ш', р', п', ф', б', в', м', к', х', г', й, и	E040	<i>Все</i> [фс'э]	E041	<i>Затем</i> [зат'эм]	E042	<i>Потеха</i> [пат'эхэ]	E043	<i>Теперь</i> [т'ип'э́р']

Таблица А.10 – Мини-набор аллофонов для безударной гласной Е

Третий индекс, правый контекст		0		1		2		3	
		пауза		п, ф, б, в, м		т, ц, с, д, з, н, л, ш, ж, р, к, х, г, у, о, а, э, ы		к', х', г', й, ч', ш', р', т', с', д', з', н', л', п', ф', б', в', м', и	
0	пауза	E100	<i>Э</i> [э]	E101	<i>Эпоптов</i> <i>Эфиопов</i> [иф'ио́пѣ] <i>Эполетами</i>	E102	<i>Этаже</i> [итажэ́] <i>Экран</i>	E103	<i>Элиав</i> <i>Эфиопского</i> <i>Эпизод</i> [ип'изо́т]
1	п, ф, б, в, м	E110	<i>Джузеппе</i> [джуз'эп':э]	E111	-	E112	-	E113	-
2	ш, ж, р, т, ц, с, д, з, н, л	E120	<i>Лучие</i> [лу́ч'иѣ]	E121	<i>Шепну</i> [шы́пну]	E122	<i>Экзекуции</i> [игз'ику́ѣѣ]	E123	<i>Шепелявляций</i> [ш'ьп'ил'áv'ьш'ьй]
3	к, х, г, у, о, а, э, ы	E130	<i>Алоэ</i> [ало́и]	E131	-	E132	<i>Аэрмона</i> [аирмо́нѣ]	E133	<i>Наэлектризованном</i> [нэ́ил'э́ктр'изо́вѣн:ъм]
4	т', с', д', з', н', л', ч', ш', р', п', ф', б', в', м', к', х', г', й, и	E140	<i>Будьте</i> [бу́т':э]	E141	<i>Тепла</i> [т'ипла́]	E142	<i>Лейтенантский</i> [л'ьйт'ина́нцк'ьй]	E143	<i>Теперь</i> [т'ип'э́р']

ПРОДОЛЖЕНИЕ ПРИЛОЖЕНИЯ А

Таблица А.11 – Мини-набор аллофонов для ударной гласной У

Третий индекс, правый контекст		0		1		2		3	
		пауза		п, ф, б, в, м		т, ц, с, д, з, н, л, ш, ж, р, к, х, г, у, о, а, э, ы		к', х', г', й, ч', ш', р', т', с', д', з', н', л', п', ф', б', в', м', и	
0	пауза	Y000	-	Y001	-	Y002	-	Y003	-
1	п, ф, б, в, м	Y010	<i>Головы</i> []	Y011	<i>Выпустили</i> []	Y012	<i>Вспыхнуть</i> []	Y013	<i>Выпив</i> []
2	ш, ж, р, т, ц, с, д, з, н, л	Y020	<i>Удивлены</i> []	Y021	<i>Посыпались</i> []	Y022	<i>Скрытого</i> []	Y023	<i>Тушите</i> []
3	к, х, г, у, о, а, э, ы	Y030	-	Y031	-	Y032	-	Y033	-
4	т', с', д', з', н', л', ч', ш', р', п', ф', б', в', м', к', х', г', й, и	Y040	-	Y041	-	Y042	-	Y043	-

Таблица А.12 – Мини-набор аллофонов для безударной гласной У

Третий индекс, правый контекст		0		1		2		3	
		пауза		п, ф, б, в, м		т, ц, с, д, з, н, л, ш, ж, р, к, х, г, у, о, а, э, ы		к', х', г', й, ч', ш', р', т', с', д', з', н', л', п', ф', б', в', м', и	
0	пауза	Y100	-	Y101	-	Y102	-	Y103	-
1	п, ф, б, в, м	Y110	<i>Чтобы</i> [што́бы]	Y111	<i>Бывало</i> [быва́ль]	Y112	<i>Быка</i> [быка́]	Y113	<i>Неглупый</i> [не'иглу́пый]
2	ш, ж, р, т, ц, с, д, з, н, л	Y120	<i>Ленты</i> [л'э́нты]	Y121	<i>Просыпается</i> [пръсыпа́ицъ]	Y122	<i>Зарыдал</i> [зърыда́л]	Y123	<i>Усыпит</i> [усып'ит]
3	к, х, г, у, о, а, э, ы	Y130	-	Y131	-	Y132	-	Y133	-
4	т', с', д', з', н', л', ч', ш', р', п', ф', б', в', м', к', х', г', й, и	Y140	-	Y141	-	Y142	-	Y143	-

ПРОДОЛЖЕНИЕ ПРИЛОЖЕНИЯ А

Таблица А.13 – Мини-набор аллофонов для согласной **Р**

Правый контекст (индекс аллофона)	0	1	2	3	4
	Пауза	Глухой согласный	Звонкий согласный	Безударный гласный	Ударный гласный
Язык					
Русский	<i>Вор</i> [вóр]	<i>Марка</i> [мáркъ]	<i>Кордон</i> [кардо́н]	<i>Караван</i> [кѣравáн]	<i>Парад</i> [парáт]

Таблица А.14 – Мини-набор аллофонов для согласной **Р**

Правый контекст (индекс аллофона)	0	1	2	3	4
	Пауза	Глухой согласный	Звонкий согласный	Безударный гласный	Ударный гласный
Язык					
Русский	<i>Скле́</i> [скл'éп]	<i>Просту́рка</i> [прасту́ркъ]	<i>Внеза́пно</i> [вн'изáпнъ]	<i>Попа́сть</i> [пaпá't']	<i>Попа́сть</i> [пaпá't']

Таблица А.15 – Мини-набор аллофонов для согласной **Р**

Правый контекст (индекс аллофона)	0	1	2	3	4
	Пауза	Глухой согласный	Звонкий согласный	Безударный гласный	Ударный гласный
Язык					
Русский	<i>Граф</i> [гра́ф]	<i>Кафтан</i> [кафта́н]	<i>Флота</i> [фло́тъ]	<i>Фасона</i> [фасо́нь]	<i>Факелами</i> [фáк'эльм'ь]

Таблица А.16 – Мини-набор аллофонов для согласной **В**

Правый контекст (индекс аллофона)	0	1	2	3	4
	Пауза	Глухой согласный	Звонкий согласный	Безударный гласный	Ударный гласный
Язык					
Русский	-	-	<i>Обду́мывал</i> [абду́мъвл]	<i>Поба́гровел</i> [пѣбъграв'él]	<i>Губа́ми</i> [зубáм'и]

Таблица А.17 – Мини-набор аллофонов для согласной **В**

Правый контекст (индекс аллофона)	0	1	2	3	4
	Пауза	Глухой согласный	Звонкий согласный	Безударный гласный	Ударный гласный
Язык					
Русский	-	<i>Ловко</i> [ло́фкъ]	<i>Взгля́ните</i> [взгл'ини́т'ь]	<i>Клюкѡа</i> [кл'у́квъ]	<i>Призѡа́ть</i> [пр'изѡáт']

ПРОДОЛЖЕНИЕ ПРИЛОЖЕНИЯ А

Таблица А.18 – Мини-набор аллофонов для согласной М

Правый контекст (индекс аллофона)	0	1	2	3	4
	Пауза	Глухой согласный	Звонкий согласный	Безударный гласный	Ударный гласный
Язык					
Русский	<i>Други<u>м</u></i> [dʁʊg'ím]	<i>Гро<u>м</u>кой</i> [grómkʲɪ]	<i>Мгнове<u>н</u>но</i> [mɣnav'én:ɤ]	<i>Ду<u>м</u>ал</i> [dú'mɐl]	<i>Ро<u>м</u>анс</i> [ramáns]

Таблица А.19 – Мини-набор аллофонов для согласной Р'

Правый контекст (индекс аллофона)	0	1	2	3	4
	Пауза	Глухой согласный	Звонкий согласный	Безударный гласный	Ударный гласный
Язык					
Русский	<i>Це<u>п</u>ь</i> [cɛp'ʲ]	-	<i>П<u>р</u>яные</i> [p'rjáńɤ]	<i>П<u>р</u>итай</i> [p'ítáɪ]	<i>С<u>п</u>ину</i> [sp'ínu]

Таблица А.20 – Мини-набор аллофонов для согласной F'

Правый контекст (индекс аллофона)	0	1	2	3	4
	Пауза	Глухой согласный	Звонкий согласный	Безударный гласный	Ударный гласный
Язык					
Русский	<i>Сула<u>м</u>и<u>ф</u>ь</i> [сулам'íf'ʲ]	-	<i>Ф<u>л</u>ейту</i> [fl'éɪtu]	<i>Э<u>ф</u>иопка</i> [иф'иопкʲ]	<i>А<u>ф</u>ишах</i> [аф'ишɤx]

Таблица А.21 – Мини-набор аллофонов для согласной В'

Правый контекст (индекс аллофона)	0	1	2	3	4
	Пауза	Глухой согласный	Звонкий согласный	Безударный гласный	Ударный гласный
Язык					
Русский	-	-	<i>Б<u>в</u>ют</i> [б'júʈ] Лю <u>б</u> ви	<i>Об<u>в</u>итатели</i> [аб'итát'ьл'ь]	<i>В<u>л</u>юб<u>в</u>лись</i> [вл'уб'íl'ьс'ʲ]

Таблица А.22 – Мини-набор аллофонов для согласной V'

Правый контекст (индекс аллофона)	0	1	2	3	4
	Пауза	Глухой согласный	Звонкий согласный	Безударный гласный	Ударный гласный
Язык					
Русский	-	-	<i>С<u>в</u>новья</i> [с'нав'já]	<i>В<u>в</u>ина</i> [в'iná]	<i>В<u>в</u>идишь</i> [в'íd'ьш]

ПРОДОЛЖЕНИЕ ПРИЛОЖЕНИЯ А

Таблица А.23 – Мини-набор аллофонов для согласной М'

Правый контекст (индекс аллофона)	0	1	2	3	4
	Пауза	Глухой согласный	Звонкий согласный	Безударный гласный	Ударный гласный
Язык					
Русский	<i>Семь</i> [с'эм']	-	<i>Семья</i> [с'им'já]	<i>Безумия</i> [б'изум'ьь]	<i>Смирно</i> [см'и́рнъ]

Таблица А.24 – Мини-набор аллофонов для согласной Т

Правый контекст (индекс аллофона)	0	1	2	3	4
	Пауза	Глухой согласный	Звонкий согласный	Безударный гласный	Ударный гласный
Язык					
Русский	<i>Гово́рят</i> [гъвар'át]	<i>Отказа́м</i> [атка́зм]	<i>Высо́чество</i> [высоч'ьствъ]	<i>Пожалу́йста</i> [пажалу́йсть]	<i>Ста́рый</i> [ста́рый]

Таблица А.25 – Мини-набор аллофонов для согласной С

Правый контекст (индекс аллофона)	0	1	2	3	4
	Пауза	Глухой согласный	Звонкий согласный	Безударный гласный	Ударный гласный
Язык					
Русский	<i>Отец</i> [ат'эц]	<i>Мужицко́й</i> [мужы́цкый]	<i>Цвета́м</i> [цв'итám]	<i>Птица́</i> [пт'и́ць]	<i>Отца́</i> [ати́ца]

Таблица А.26 – Мини-набор аллофонов для согласной S

Правый контекст (индекс аллофона)	0	1	2	3	4
	Пауза	Глухой согласный	Звонкий согласный	Безударный гласный	Ударный гласный
Язык					
Русский	<i>Ча́с</i> [ч'ás]	<i>Ма́сках</i> [ма́скъх]	<i>Высма́тривая</i> [высмáтр'ьвь]	<i>В са́ду</i> [фсаду́]	<i>Са́хар</i> [сáхър]

Таблица А.27 – Мини-набор аллофонов для согласной D

Правый контекст (индекс аллофона)	0	1	2	3	4
	Пауза	Глухой согласный	Звонкий согласный	Безударный гласный	Ударный гласный
Язык					
Русский	-	-	<i>Двое́</i> [дво́ь]	<i>Правда́</i> [пра́вдъ]	<i>Тогда́</i> [тагда́]

ПРОДОЛЖЕНИЕ ПРИЛОЖЕНИЯ А

Таблица А.28 – Мини-набор аллофонов для согласной Z

Правый контекст (индекс аллофона)	0	1	2	3	4
	Пауза	Глухой согласный	Звонкий согласный	Безударный гласный	Ударный гласный
Язык					
Русский	-	-	<i>Избавиться</i> [избáв'ьцъ]	<i>Заморская</i> [замóрскъ]	<i>Сказал</i> [сказáл]

Таблица А.29 – Мини-набор аллофонов для согласной N

Правый контекст (индекс аллофона)	0	1	2	3	4
	Пауза	Глухой согласный	Звонкий согласный	Безударный гласный	Ударный гласный
Язык					
Русский	<i>Нужен</i> [нужьн]	<i>Наизнанку</i> [нъизнáнку]	<i>Январе</i> [йьнвар'э]	<i>Надеюсь</i> [над'ейус']	<i>Наших</i> [нашгх]

Таблица А.30 – Мини-набор аллофонов для согласной L

Правый контекст (индекс аллофона)	0	1	2	3	4
	Пауза	Глухой согласный	Звонкий согласный	Безударный гласный	Ударный гласный
Язык					
Русский	<i>Причалил</i> [пр'ич'áл'ьл]	<i>Водки</i> [вóлк'и]	<i>Бодван</i> [балвáн]	<i>Разоблачить</i> [рзьзблач'ит']	<i>Славой</i> [слáвьй]

Таблица А.31 – Мини-набор аллофонов для согласной Т'

Правый контекст (индекс аллофона)	0	1	2	3	4
	Пауза	Глухой согласный	Звонкий согласный	Безударный гласный	Ударный гласный
Язык					
Русский	<i>Делать</i> [д'эльт']	-	-	<i>Чести</i> [ч'эс't'u]	<i>Схватила</i> [схват'иль]

Таблица А.32 – Мини-набор аллофонов для согласной S'

Правый контекст (индекс аллофона)	0	1	2	3	4
	Пауза	Глухой согласный	Звонкий согласный	Безударный гласный	Ударный гласный
Язык					
Русский	<i>Спрячусь</i> [спр'áч'ус']	<i>Чести</i> [ч'эс't'u]	-	<i>Седина</i> [с'ьд'ина]	<i>Беседке</i> [б'ис'этк'ь]

ПРОДОЛЖЕНИЕ ПРИЛОЖЕНИЯ А

Таблица А.33 – Мини-набор аллофонов для согласной D'

Правый контекст (индекс аллофона)	0	1	2	3	4
	Пауза	Глухой согласный	Звонкий согласный	Безударный гласный	Ударный гласный
Язык					
Русский	-	-	<i>Дьявол</i> [d'jávɐl]	<i>Будет</i> [búd'ɛt]	<i>Сердечные</i> [s'ird'ɛt'nʲɔ]

Таблица А.34 – Мини-набор аллофонов для согласной Z'

Правый контекст (индекс аллофона)	0	1	2	3	4
	Пауза	Глухой согласный	Звонкий согласный	Безударный гласный	Ударный гласный
Язык					
Русский	-	-	<i>Друзья</i> [druz'já]	<i>Образе</i> [óbrɛz'ɛ]	<i>Друзей</i> [druz'ɛj]

Таблица А.35 – Мини-набор аллофонов для согласной N'

Правый контекст (индекс аллофона)	0	1	2	3	4
	Пауза	Глухой согласный	Звонкий согласный	Безударный гласный	Ударный гласный
Язык					
Русский	<i>Очень</i> [óč'ɲn']	<i>Заинтересовался</i> [zɨin't'ɐr'ɛsávɐlɛ]	<i>Претензию</i> [pr'it'ɛn'z'ɨju]	<i>Сложения</i> [sləʒɛn'ɛ]	<i>Книга</i> [kn'ígɔ]

Таблица А.36 – Мини-набор аллофонов для согласной L'

Правый контекст (индекс аллофона)	0	1	2	3	4
	Пауза	Глухой согласный	Звонкий согласный	Безударный гласный	Ударный гласный
Язык					
Русский	<i>Служитель</i> [sluʒít'ɛl']	<i>Мальчик</i> [mal'č'ɲk]	<i>Револьвером</i> [r'ɛvɐl'v'érɐm]	<i>Мыслей</i> [mysl'ɛj]	<i>Лентами</i> [l'ɛntɐm'ɛ]

Таблица А.37 – Мини-набор аллофонов для согласной Sh

Правый контекст (индекс аллофона)	0	1	2	3	4
	Пауза	Глухой согласный	Звонкий согласный	Безударный гласный	Ударный гласный
Язык					
Русский	<i>Шалаш</i> [шалáш]	<i>Девушку</i> [d'ɛvuʃku]	<i>Швыряет</i> [швыр'áйт]	<i>Шалаш</i> [шалáш]	<i>Душа</i> [душá]

ПРОДОЛЖЕНИЕ ПРИЛОЖЕНИЯ А

Таблица А.38 – Мини-набор аллофонов для согласной Zh

Правый контекст (индекс аллофона)	0	1	2	3	4
	Пауза	Глухой согласный	Звонкий согласный	Безударный гласный	Ударный гласный
Язык					
Русский	-	-	<i>Службы</i> [службы]	<i>Мужа</i> [мужь]	<i>Продержав</i> [прѣд'ѣржѧф]

Таблица А.39 – Мини-набор аллофонов для согласной Ch'

Правый контекст (индекс аллофона)	0	1	2	3	4
	Пауза	Глухой согласный	Звонкий согласный	Безударный гласный	Ударный гласный
Язык					
Русский	<i>Михайлович</i> [м'ѣхѧйльѣ'ѣч']	<i>Очками</i> [ач'кам'ѣ]	<i>Ночью</i> [ноч'йу]	<i>Тотчас</i> [тоѣч'ѣс]	<i>Начальник</i> [нач'ѧл'н'ѣк]

Таблица А.40 – Мини-набор аллофонов для согласной Sh'

Правый контекст (индекс аллофона)	0	1	2	3	4
	Пауза	Глухой согласный	Звонкий согласный	Безударный гласный	Ударный гласный
Язык					
Русский	<i>Товарищ</i> [таѣар'ѣш']	-	<i>Помощью</i> [пѣмъш'йу]	<i>Внемлющая</i> [вн'ѣмл'уш'ѣѣ]	<i>Обещаю</i> [аб'ѣш'ѧйу]

Таблица А.41 – Мини-набор аллофонов для согласной R'

Правый контекст (индекс аллофона)	0	1	2	3	4
	Пауза	Глухой согласный	Звонкий согласный	Безударный гласный	Ударный гласный
Язык					
Русский	<i>Теперь</i> [т'ип'ѣр']	<i>Верьте</i> [ѣ'ѣр'т'ѣ]	<i>Серьезнее</i> [с'ир'јѣз'н'ѣѣ]	<i>Прекрасно</i> [пр'икрѧснѣ]	<i>Интересно</i> [ин'т'ир'ѣснѣ]

Таблица А.42 – Мини-набор аллофонов для согласной K

Правый контекст (индекс аллофона)	0	1	2	3	4
	Пауза	Глухой согласный	Звонкий согласный	Безударный гласный	Ударный гласный
Язык					
Русский	<i>Обидчик</i> [аб'йч'ѣк]	<i>Клякса</i> [кл'ѧксѣ]	<i>Клякса</i> [кл'ѧксѣ]	<i>Конечно</i> [кан'ѣч'нѣ]	<i>Таком</i> [такѣм]

ПРОДОЛЖЕНИЕ ПРИЛОЖЕНИЯ А

Таблица А.43 – Мини-набор аллофонов для согласной Н

Правый контекст (индекс аллофона)	0	1	2	3	4
	Пауза	Глухой согласный	Звонкий согласный	Безударный гласный	Ударный гласный
Язык					
Русский	<i>Глаза́х</i> [глаза́х]	<i>Оста́вшихс́я</i> [аста́фихс 'ь]	<i>Охраня́ют</i> [ахран 'а́йут]	<i>Ха́лате</i> [хала́т 'ь]	<i>Дыха́ние</i> [дыха́н 'ьь]

Таблица А.44 – Мини-набор аллофонов для согласной Г

Правый контекст (индекс аллофона)	0	1	2	3	4
	Пауза	Глухой согласный	Звонкий согласный	Безударный гласный	Ударный гласный
Язык					
Русский	-	-	<i>Тогда́</i> [тагда́]	<i>Трево́га</i> [тр 'иво́гъ]	<i>Друга́я</i> [друга́ь]

Таблица А.45 – Мини-набор аллофонов для согласной К'

Правый контекст (индекс аллофона)	0	1	2	3	4
	Пауза	Глухой согласный	Звонкий согласный	Безударный гласный	Ударный гласный
Язык					
Русский	-	-	-	<i>Удо́чке</i> [у́дьч 'к 'ь]	<i>Цве́тнике</i> [цв 'ьтн 'ик 'е́]

Таблица А.46 – Мини-набор аллофонов для согласной Н'

Правый контекст (индекс аллофона)	0	1	2	3	4
	Пауза	Глухой согласный	Звонкий согласный	Безударный гласный	Ударный гласный
Язык					
Русский	-	-	-	<i>Возду́хе</i> [во́здух 'ь]	<i>Хи́щника</i> [х 'и́шн 'н 'ькь]

Таблица А.47 – Мини-набор аллофонов для согласной Г'

Правый контекст (индекс аллофона)	0	1	2	3	4
	Пауза	Глухой согласный	Звонкий согласный	Безударный гласный	Ударный гласный
Язык					
Русский	-	-	-	<i>Доро́ге</i> [даро́г 'ь]	<i>Ге́нрихом</i> [г 'е́нр 'ьхьм]

ОКОНЧАНИЕ ПРИЛОЖЕНИЯ А

Таблица А.48 – Мини-набор аллофонов для согласной Ј’

Правый контекст (индекс аллофона)	0	1	2	3	4
	Пауза	Глухой согласный	Звонкий согласный	Безударный гласный	Ударный гласный
Язык					
Русский	<i>Торжественный</i> [таржэ́ств’ьн:ъй]	<i>Тайком</i> [тайко́м]	<i>Подойди</i> [пъдайд’и]	<i>Претензи<u>ю</u></i> [пр’ит’эн’з’ьйу]	<i>Прия<u>те</u>ль</i> [пр’ија́т’ьл’]

ПРИЛОЖЕНИЕ Б

Таблица Б.1 – Входные данные и ожидаемые результаты тестов для класса LetterPhonemeTranscriptor

№	Входные данные	Ожидаемый результат
1	'_'	['_']
2	'#'	['#']
3	'+'	['+']
4	'='	['=']
5	"	['']
6	"У+ТКА"	['u', '+', 't', 'k', 'a']
7	"ЮЛА+"	['j', 'u', 'l', 'a', '+']
8	"Э+ТОТ"	['e', '+', 't', 'a', 't']
9	"Е+ЛЬ"	['j', 'e', '+', 'l']
10	"ЛЁ+ТЧИК"	['l', 'o', '+', 'ch', 'i', 'k']
11	"О+СЕНЬ"	['o', '+', 's', 'e', 'n']
12	"АНАНА+С"	['a', 'n', 'a', 'n', 'a', '+', 's']
13	"Я+БЛОКА"	['j', 'a', '+', 'b', 'l', 'a', 'k', 'a']
14	"МЯ+ГКИЙ"	['m', 'a', '+', 'h', 'k', 'i', 'j']
15	"МОЛОКО+"	['m', 'a', 'l', 'a', 'k', 'o', '+']
16	"СЫ+ТЫЙ"	['s', 'y', '+', 't', 'y', 'j']
17	"ШИ+ЛО"	['sh', 'y', '+', 'l', 'a']
18	"ПРИНЁ+С_ИГРУ+ШКУ"	['p', 'r', 'i', 'n', 'o', '+', 's', ' ', 'y', 'g', 'r', 'u', '+', 'sh', 'k', 'u']
19	"ЩЕКА+"	['sh', 'e', 'k', 'a', '+']
20	"СЧА+СТЬЕ"	['sh', 'a', '+', 's', 't', 'j', 'e']
21	"ПЕРЕБЕ+ЖЧИК"	['p', 'e', 'r', 'e', 'b', 'e', '+', 'sh', 'i', 'k']
22	"ЧАСЫ+"	['ch', 'a', 's', 'y', '+']
23	"ЦА+ПЛЯ"	['c', 'a', '+', 'p', 'l', 'a']
24	"СПЕЦСВЯ+ЗЬ"	['s', 'p', 'e', 'c', 's', 'v', 'a', '+', 's']
25	"ПЕРЕВОЛНОВА+ТЬСЯ"	['p', 'e', 'r', 'e', 'v', 'o', 'l', 'n', 'o', 'v', 'a', '+', 'c', 'a']
26	"РУЧА+ЕТСЯ"	['r', 'u', 'ch', 'a', '+', 'j', 'e', 'c', 'a']
27	"БЛЮ+ДЦЕ"	['b', 'l', 'u', '+', 'c', 'c', 'e']
28	"ОТЦА+"	['a', 'c', 'c', 'a', '+']
29	"ТРЁ=ХХВО+СТКА"	['t', 'r', 'o', '=', 'h', 'h', 'v', 'o', '+', 's', 't', 'k', 'a']
30	"ЛЕГКО+"	['l', 'e', 'h', 'k', 'o', '+']
31	"ХИ+ТРЫЙ"	['h', 'i', '+', 't', 'r', 'y', 'j']
32	"ЛЁ+ГКИЙ"	['l', 'o', '+', 'h', 'k', 'i', 'j']
33	"ЯМА+ЙКА"	['j', 'a', 'm', 'a', '+', 'j', 'k', 'a']
34	"ВОРОБЬИ+"	['v', 'a', 'r', 'a', 'b', 'j', 'i', '+']
35	"БУЛЬО+Н"	['b', 'u', 'l', 'j', 'o', '+', 'n']
36	"КОНЯ+К"	['k', 'a', 'n', 'j', 'a', '+', 'k']
37	"ОБЪЁ+М"	['a', 'b', 'j', 'o', '+', 'm']
38	"МРА+МОР"	['m', 'r', 'a', '+', 'm', 'a', 'r']

ПРОДОЛЖЕНИЕ ПРИЛОЖЕНИЯ Б

Продолжение таблицы Б.1

№	Входные данные	Ожидаемый результат
39	"ИЮ+НЬСКИМ_ВЕ+ТРОМ"	["i", "j", "u", "+", "n", "s", "k", "i", "m", "_", "v", "e", "+", "t", "r", "a", "m"]
40	"ТРИУМВИРА+Т"	["t", "r", "i", "u", "m", "v", "i", "r", "a", "+", "t"]
41	"НАЪТУ+МБЕ"	["n", "a", "t", "u", "+", "m", "b", "e"]
42	"МЯ+ЧИК"	["m", "a", "+", "ch", "i", "k"]
43	"НО+ША"	["n", "o", "+", "sh", "a"]
44	"ВА+СИН_ЧЕМОДА+Н"	["v", "a", "+", "s", "i", "n", "_", "ch", "e", "m", "a", "d", "a", "+", "n"]
45	"ЗАКО+НЧИТЬ"	["z", "a", "k", "o", "+", "n", "ch", "i", "t"]
46	"БА+НТИК"	["b", "a", "+", "n", "t", "i", "k"]
47	"КО+НЬ"	["k", "o", "+", "n"]
48	"СО+ЛНЕЧНЫЙ"	["s", "o", "+", "l", "n", "e", "ch", "n", "y", "j"]
49	"СО+ЛНЦЕ"	["s", "o", "+", "n", "c", "e"]
50	"ПО=ЛЪЛИСТА+"	["p", "o", "=", "l", "l", "i", "s", "t", "a", "+"]
51	"Э+ЛЛИПС"	["e", "+", "l", "l", "i", "p", "s"]
52	"НО+ЛЬ"	["n", "o", "+", "l"]
53	"КРОМЕ+ШНАЯ"	["k", "r", "o", "m", "e", "+", "sh", "n", "a", "j", "a"]
54	"КОРРИ+ДА"	["k", "o", "r", "r", "i", "+", "d", "a"]
55	"КО+РЬ"	["k", "o", "+", "r"]
56	"СТОЛБНЯ+К"	["s", "t", "o", "l", "b", "n", "y", "a", "+", "k"]
57	"ДУ+Б_ЗЕЛЁ+НЬИЙ"	["d", "u", "+", "b", "_", "z", "e", "l", "e", "n", "y", "j"]
58	"КРЕПДЫШИ+Н"	["k", "r", "e", "p", "d", "y", "sh", "i", "n"]
59	"СЕ+РП_ЗАБЛЕСТЕ+Л"	["s", "e", "+", "r", "p", "_", "z", "a", "b", "l", "e", "s", "t", "e", "+", "l"]
60	"СТО+ЛБ"	["s", "t", "o", "+", "l", "b"]
61	"ОБМЯ+К"	["o", "b", "m", "y", "a", "+", "k"]
62	"НАЪПНЕ+"	["n", "a", "p", "n", "e", "+"]
63	"ПО+Л"	["p", "o", "+", "l"]
64	"ОКО+П"	["o", "k", "o", "+", "p"]
65	"ПИ+Л"	["p", "i", "+", "l"]
66	"СКО+РБЬ"	["s", "k", "o", "+", "r", "b"]
67	"ГРА+Ф_ВИ+КТОР"	["g", "r", "a", "+", "f", "v", "i", "k", "t", "o", "r"]
68	"СА+МОГО"	["s", "a", "+", "m", "o", "g", "o"]
69	"ТВОЕГО+"	["t", "v", "o", "e", "g", "o", "+"]
70	"АФГАНИСТА+Н"	["a", "f", "g", "a", "n", "i", "s", "t", "a", "+", "n"]
71	"ВВЁ+З"	["v", "v", "o", "+", "z"]
72	"ВИ+ЛКА"	["v", "i", "+", "l", "k", "a"]
73	"КРО+ВЬ_ВРАГА+"	["k", "r", "o", "+", "v", "v", "r", "a", "g", "a", "+"]
74	"ЗА+ВТРА"	["z", "a", "+", "f", "t", "r", "a"]
75	"ФРУ+КТ"	["f", "r", "u", "+", "k", "t"]
76	"КРО+В"	["k", "r", "o", "+", "f"]
77	"ГОТО+ВЬСЯ"	["g", "o", "t", "o", "+", "f", "s", "y", "a"]

ПРОДОЛЖЕНИЕ ПРИЛОЖЕНИЯ Б

Продолжение таблицы Б.1

№	Входные данные	Ожидаемый результат
78	"ФИ+КУС"	["f", "i", "+", "k", "u", "s"]
79	"КРО+ВЬ"	["k", "r", "o", "+", "f"]
80	"ПОДУ+Л"	["p", "a", "d", "u", "+", "l"]
81	"БЕ+ЗДНА"	["b", "e", "+", "z", "n", "a"]
82	"ПО+ЗДНИЙ"	["p", "o", "+", "z", "n", "i", "j"]
83	"МУНДШТУ+К"	["m", "u", "n", "sh", "t", "u", "+", "k"]
84	"ЛАНДТА+Г"	["l", "a", "n", "t", "a", "+", "k"]
85	"ЯГДТА+Ш"	["j", "a", "k", "t", "a", "+", "sh"]
86	"СЕРДЧИ+ШКО"	["s", "e", "r", "ch", "i", "+", "sh", "k", "a"]
87	"УКЛА+ДЧИК"	["u", "k", "l", "a", "+", "ch", "i", "k"]
88	"БЛЮ+ДЦЕ"	["b", "l", "u", "+", "c", "c", "e"]
89	"ОБИ+ДЧИК"	["a", "b", "i", "+", "ch", "i", "k"]
90	"КО+Д ДО+МА"	["k", "o", "+", "d", "_", "d", "o", "+", "m", "a"]
91	"ДНЯ+МИ"	["d", "n", "a", "+", "m", "i"]
92	"ОТДИРА+ТЬ"	["a", "d", "d", "i", "r", "a", "+", "t"]
93	"ЧА+СТНЫЙ"	["ch", "a", "+", "s", "n", "y", "j"]
94	"ЗАСТЛА+ТЬ"	["z", "a", "s", "l", "a", "+", "t"]
95	"РЕНТГЕ+Н"	["r", "e", "n", "g", "e", "+", "n"]
96	"ХРИПОТЦА+ "	["h", "r", "i", "p", "a", "c", "c", "a", "+"]
97	"ЧУКО+ТСКИЙ"	["ch", "u", "k", "o", "+", "c", "k", "i", "j"]
98	"БРА+ТСТВО"	["b", "r", "a", "+", "c", "t", "v", "a"]
99	"КРУ+ЖИТСЯ"	["k", "r", "u", "+", "zh", "y", "c", "a"]
100	"КРУЖИ+ТЬСЯ"	["k", "r", "u", "zh", "y", "+", "c", "a"]
102	"МЛА+ДШЕ"	["m", "l", "a", "+", "t", "sh", "e"]
103	"КО+Д"	["k", "o", "+", "t"]
104	"СЧАСТЛИ+ВЫЙ"	["sh", "a", "s", "l", "i", "+", "v", "y", "j"]
105	"ПОДТИ+П"	["p", "a", "t", "t", "i", "+", "p"]
106	"ТИ+П"	["t", "i", "+", "p"]
107	"БУ+ДЬ"	["b", "u", "+", "t"]
108	"ВОКЗА+Л"	["v", "a", "g", "z", "a", "+", "l"]
109	"СТОКГО+ЛЬМ"	["s", "t", "a", "g", "g", "o", "+", "l", "m"]
110	"СВОЕГО+ "	["s", "v", "a", "j", "e", "v", "o", "+"]
111	"СА+МОГО"	["s", "a", "+", "m", "a", "v", "a"]
112	"ГО+Д"	["g", "o", "+", "t"]
113	"ХИ+МИК_ГО+ДА"	["h", "i", "+", "m", "i", "g", "_", "g", "o", "+", "d", "a"]
114	"ТРИ+ГГЕР"	["t", "r", "i", "+", "g", "g", "e", "r"]
115	"ТЕ+НА "	["g", "e", "+", "n", "a"]
116	"БЕ+ГСТВО"	["b", "e", "+", "k", "s", "t", "v", "a"]
117	"АКСИО+МА"	["a", "k", "s", "i", "o", "+", "m", "a"]
118	"СТО+Г"	["s", "t", "o", "+", "k"]
119	"СТО+Г <u>СЕ</u> +НА"	["s", "t", "o", "+", "k", "_", "s", "e", "+", "n", "a"]
120	"КИ+ПР"	["k", "i", "+", "p", "r"]
121	"СБА+ВИТЬ"	["z", "b", "a", "+", "v", "i", "t"]
122	"Е+ЗЖУ"	["j", "e", "+", "zh", "zh", "u"]

ПРОДОЛЖЕНИЕ ПРИЛОЖЕНИЯ Б

Продолжение таблицы Б.1

№	Входные данные	Ожидаемый результат
123	"ЗО+Л"	["z", "o", "+", "l"]
124	"ПРИНЁ+С_ЗУ+БЫ"	["p", "r", "i", "n", "o", "+", "z", "_", "z", "u", "+", "b", "y"]
125	"ЗЕЛЁ+НЫЙ"	["z", "e", "l", "o", "+", "n", "y", "j"]
126	"СДЕ+ЛАТЬ"	["z", "d", "e", "+", "l", "a", "t"]
127	"КОСЬБА+"	["k", "a", "z", "b", "a", "+"]
128	"БЕЗПОЩА+ДНО"	["b", "e", "s", "p", "a", "sh", "a", "+", "d", "n", "a"]
129	"РУЧА+ЕТСЯ"	["r", "u", "ch", "a", "+", "j", "e", "c", "a"]
130	"ПЕРЕВОЛНОВА+ТЬСЯ"	["p", "e", "r", "e", "v", "a", "l", "n", "a", "v", "a", "+", "c", "a"]
131	"ВЫ+ЛЕЗ"	["v", "y", "+", "l", "e", "s"]
132	"ВЫ+ЛЕЗ СО+М"	["v", "y", "+", "l", "e", "s", "_", "s", "o", "+", "m"]
133	"СЧА+СТЬЕ"	["sh", "a", "+", "s", "t", "j", "e"]
134	"РАСШИ+Б"	["r", "a", "sh", "sh", "y", "+", "p"]
135	"АГРЕССИ+ВНОСТЬ"	["a", "g", "r", "e", "s", "s", "i", "+", "v", "n", "a", "s", "t"]
136	"СИ+ЛА"	["s", "i", "+", "l", "a"]
137	"БРО+СЬ"	["b", "r", "o", "+", "s"]
138	"ЖО+РА"	["zh", "o", "+", "r", "a"]
139	"ЖИ+ТЬ"	["zh", "y", "+", "t"]
140	"ДРО+ЖЬ_ЗЕМЛИ+"	["d", "r", "o", "+", "zh", "_", "z", "e", "m", "l", "i", "+"]
141	"ВЪЕЗЖА+ТЬ"	["v", "j", "e", "zh", "zh", "a", "+", "t"]
142	"БРО+ШЬ ЖЕ+НИ"	["b", "r", "o", "+", "zh", "_", "zh", "e", "+", "n", "i"]
143	"МОЛОДЁ+ЖЬ"	["m", "a", "l", "a", "d", "o", "+", "sh"]
144	"РАСШИ+Б"	["r", "a", "sh", "sh", "y", "+", "p"]
145	"ШИ+ЛО"	["sh", "y", "+", "l", "a"]
146	"СКА+ЗКА"	["s", "k", "a", "+", "s", "k", "a"]
147	"ЛО+ДКА"	["l", "o", "+", "t", "k", "a"]
148	"РАСПИСА+ТЬ"	["r", "a", "s", "p", "i", "s", "a", "+", "t"]
149	"ВХО+Д"	["f", "h", "o", "+", "t"]
150	"ОТБИ+ТЬ"	["a", "d", "b", "i", "+", "t"]
151	"ПРО+СЬБА"	["p", "r", "o", "+", "z", "b", "a"]
152	"КОСЬБА+"	["k", "a", "z", "b", "a", "+"]
153	"СДА+ТЬ"	["z", "d", "a", "+", "t"]
154	"СВО+Й"	["s", "v", "o", "+", "j"]
155	"КВА+С"	["k", "v", "a", "+", "s"]
156	"ОТВЕ+Т"	["a", "t", "v", "e", "+", "t"]
157	"ХВАЛИ+ТЬ"	["h", "v", "a", "l", "i", "+", "t"]
158	"БРА+Т_ДРУ+ГА"	["b", "r", "a", "+", "d", "_", "d", "r", "u", "+", "g", "a"]
159	"ПРИНЁ+С_ДОМО+Й"	["p", "r", "i", "n", "o", "+", "z", "_", "d", "a", "m", "o", "+", "j"]
160	"ВСТРЯХНУ+ТЬ"	["f", "s", "t", "r", "a", "h", "n", "u", "+", "t"]

ПРОДОЛЖЕНИЕ ПРИЛОЖЕНИЯ Б

Продолжение таблицы Б.1

№	Входные данные	Ожидаемый результат
161	"ЧА+СТЬ_ГО+РОДА"	["ch", "a", "+", "z", "d", "_", "g", "o", "+", "r", "a", "d", "a"]
162	"ТЕ+КСТ_БИ+БЛИИ"	["t", "e", "+", "g", "z", "d", "_", "b", "i", "+", "b", "l", "i", "i"]
163	"ЛУ+Г"	["l", "u", "+", "k"]
164	"РО+З"	["r", "o", "+", "s"]
165	"ПРУ+Д"	["p", "r", "u", "+", "t"]
166	"ДУ+Б"	["d", "u", "+", "p"]
167	"ЛУ+Г ДО+МА"	["l", "u", "+", "g", "_", "d", "o", "+", "m", "a"]
168	"ЛУ+Г ВОДА+"	["l", "u", "+", "k", "_", "v", "a", "d", "a", "+"]
169	"ЧА+СТНЫЙ"	["ch", "a", "+", "s", "n", "y", "j"]
170	"СЧАСТЛИ+ВЫЙ"	["sh", "a", "s", "l", "i", "+", "v", "y", "j"]
171	"РЕНТГЕ+Н"	["r", "e", "n", "g", "e", "+", "n"]
172	"ПО+ЗДНО"	["p", "o", "+", "z", "n", "a"]
173	"ПОДЪУЗДЦЫ+"	["p", "a", "d", "u", "s", "c", "y", "+"]
174	"ГОЛЛА+НДЦЫ"	["g", "a", "l", "l", "a", "+", "n", "c", "y"]
175	"СЕ+РДЦЕ"	["s", "e", "+", "r", "c", "e"]
176	"ЛАНДША+ФТ"	["l", "a", "n", "sh", "a", "+", "f", "t"]
177	"ЯГДТА+Ш"	["j", "a", "k", "t", "a", "+", "sh"]
178	"СО+ЛНЦЕ"	["s", "o", "+", "n", "c", "e"]
179	"СЧА+СТЬЕ"	["sh", "a", "+", "s", "t", "j", "e"]
180	"ПЕРЕБЕ+ЖЧИК"	["p", "e", "r", "e", "b", "e", "+", "sh", "i", "k"]
181	"ПЕРЕВОЛНОВА+ТЬСЯ"	["p", "e", "r", "e", "v", "o", "l", "n", "a", "v", "a", "+", "c", "a"]
182	"РУЧА+ЕТСЯ"	["r", "u", "ch", "a", "+", "j", "e", "c", "a"]
183	"БЛЮ+ДЦЕ"	["b", "l", "u", "+", "c", "c", "e"]
184	"ОТЦА+"	["a", "c", "c", "a", "+"]
185	"ЛЁ+ГКИЙ"	["l", "o", "+", "h", "k", "i", "j"]
186	"ЛЕГКО+"	["l", "e", "h", "k", "o", "+"]
187	"СИ+НЕГО"	["s", "i", "+", "n", "e", "v", "a"]
188	"КРА+СНОГО"	["k", "r", "a", "+", "s", "n", "a", "v", "a"]
180	"МНО+ГО"	["m", "n", "o", "+", "g", "a"]
190	"ДО+РОГО"	["d", "o", "+", "r", "a", "g", "a"]
191	"РАСШИ+Б"	["r", "a", "sh", "sh", "y", "+", "p"]
192	"ВЪЕЗЖА+ТЬ"	["v", "j", "e", "zh", "zh", "a", "+", "t"]

ПРОДОЛЖЕНИЕ ПРИЛОЖЕНИЯ Б

Таблица Б.2 – Входные данные и ожидаемые результаты тестов для класса
PhonemeAllophoneTranscriptor

№	Входные данные	Ожидаемый результат
1	["a"]	["a_100"]
2	["o"]	["o_100"]
3	["u"]	["u_100"]
4	["i"]	["i_100"]
5	["e"]	["e_100"]
6	["y"]	["y_100"]
7	["_", "a"]	["pause_0", "a_100"]
8	["_", "o"]	["pause_0", "o_100"]
9	["_", "u"]	["pause_0", "u_100"]
10	["_", "i"]	["pause_0", "i_100"]
11	["_", "e"]	["pause_0", "e_100"]
12	["_", "y"]	["pause_0", "y_100"]
13	["#", "a"]	["pause_1", "a_100"]
14	["#", "o"]	["pause_1", "o_100"]
15	["#", "u"]	["pause_1", "u_100"]
16	["#", "i"]	["pause_1", "i_100"]
17	["#", "e"]	["pause_1", "e_100"]
18	["#", "y"]	["pause_1", "y_100"]
19	["a", "_"]	["a_100", "pause_0"]
20	["o", "_"]	["o_100", "pause_0"]
21	["u", "_"]	["u_100", "pause_0"]
22	["i", "_"]	["i_100", "pause_0"]
23	["e", "_"]	["e_100", "pause_0"]
24	["y", "_"]	["y_100", "pause_0"]
25	["a", "#"]	["a_100", "pause_1"]
26	["o", "#"]	["o_100", "pause_1"]
27	["u", "#"]	["u_100", "pause_1"]
28	["i", "#"]	["i_100", "pause_1"]
29	["e", "#"]	["e_100", "pause_1"]
30	["y", "#"]	["y_100", "pause_1"]
31	["_", "a", "_"]	["pause_0", "a_100", "pause_0"]
32	["_", "o", "_"]	["pause_0", "o_100", "pause_0"]
33	["_", "u", "_"]	["pause_0", "u_100", "pause_0"]
34	["_", "i", "_"]	["pause_0", "i_100", "pause_0"]
35	["_", "e", "_"]	["pause_0", "e_100", "pause_0"]
36	["_", "y", "_"]	["pause_0", "y_100", "pause_0"]
37	["#", "a", "#"]	["pause_1", "a_100", "pause_1"]
38	["#", "o", "#"]	["pause_1", "o_100", "pause_1"]
39	["#", "u", "#"]	["pause_1", "u_100", "pause_1"]
40	["#", "i", "#"]	["pause_1", "i_100", "pause_1"]
41	["#", "e", "#"]	["pause_1", "e_100", "pause_1"]
42	["#", "y", "#"]	["pause_1", "y_100", "pause_1"]
43	["a", "+"]	["a_000"]

ПРОДОЛЖЕНИЕ ПРИЛОЖЕНИЯ Б

Продолжение таблицы Б.2

№	Входные данные	Ожидаемый результат
44	["o", "+"]	["o_000"]
45	["u", "+"]	["u_000"]
46	["i", "+"]	["i_000"]
47	["e", "+"]	["e_000"]
48	["y", "+"]	["y_000"]
49	["p", "a"]	["p_3", "a_110"]
50	["p", "a", "+"]	["p_4", "a_010"]
51	["_", "p", "a"]	["pause_0", "p_3", "a_110"]
52	["_", "p", "a", "+"]	["pause_0", "p_4", "a_010"]
53	["#", "p", "a"]	["pause_1", "p_3", "a_110"]
54	["#", "p", "a", "+"]	["pause_1", "p_4", "a_010"]
55	["sh", "a"]	["sh_3", "a_120"]
56	["sh", "a", "+"]	["sh_4", "a_020"]
57	["k", "a"]	["k_3", "a_130"]
58	["k", "a", "+"]	["k_4", "a_030"]
59	["k'", "a"]	["k'_3", "a_140"]
60	["k'", "a", "+"]	["k'_4", "a_040"]
61	["a", "p"]	["a_101", "p_0"]
62	["a", "+", "p"]	["a_001", "p_0"]
63	["a", "p", "_"]	["a_101", "p_0", "pause_0"]
64	["a", "+", "p", "_"]	["a_001", "p_0", "pause_0"]
65	["a", "p", "3"]	["a_101", "p_0"]
66	["a", "+", "p", "#"]	["a_001", "p_0", "pause_1"]
67	["a", "t"]	["a_102", "t_0"]
68	["a", "+", "t"]	["a_002", "t_0"]
69	["a", "t'"]	["a_103", "t'_0"]
70	["a", "+", "t'"]	["a_003", "t'_0"]
71	["p", "a", "p"]	["p_3", "a_111", "p_0"]
72	["p", "a", "+", "p"]	["p_4", "a_011", "p_0"]
73	["p", "a", "t"]	["p_3", "a_112", "t_0"]
74	["p", "a", "+", "t"]	["p_4", "a_012", "t_0"]
75	["p", "a", "t'"]	["p_3", "a_113", "t'_0"]
76	["p", "a", "+", "t'"]	["p_4", "a_013", "t'_0"]
77	["sh", "a", "p"]	["sh_3", "a_121", "p_0"]
78	["sh", "a", "+", "p"]	["sh_4", "a_021", "p_0"]
79	["sh", "a", "t"]	["sh_3", "a_122", "t_0"]
80	["sh", "a", "+", "t"]	["sh_4", "a_022", "t_0"]
81	["sh", "a", "t'"]	["sh_3", "a_123", "t'_0"]
82	["sh", "a", "+", "t'"]	["sh_4", "a_023", "t'_0"]
83	["k", "a", "p"]	["k_3", "a_131", "p_0"]
84	["k", "a", "+", "p"]	["k_4", "a_031", "p_0"]
85	["k", "a", "t"]	["k_3", "a_132", "t_0"]
86	["k", "a", "+", "t"]	["k_4", "a_032", "t_0"]
87	["k", "a", "t'"]	["k_3", "a_133", "t'_0"]
88	["k", "a", "+", "t'"]	["k_4", "a_033", "t'_0"]

ПРОДОЛЖЕНИЕ ПРИЛОЖЕНИЯ Б

Продолжение таблицы Б.2

№	Входные данные	Ожидаемый результат
89	["k", "a", "p"]	["k'_3", "a_141", "p_0"]
90	["k", "a", "+", "p"]	["k'_4", "a_041", "p_0"]
91	["k", "a", "t"]	["k'_3", "a_142", "t_0"]
92	["k", "a", "+", "t"]	["k'_4", "a_042", "t_0"]
93	["k", "a", "t"]	["k'_3", "a_143", "t_0"]
94	["k", "a", "+", "t"]	["k'_4", "a_043", "t_0"]
95	["p"]	["p_0"]
96	["p", "_"]	["p_0", "pause_0"]
97	["p", "#"]	["p_0", "pause_1"]
98	["p"]	["p'_0"]
99	["p", "_"]	["p'_0", "pause_0"]
100	["p", "#"]	["p'_0", "pause_1"]
101	["p", "t"]	["p_1", "t_0"]
102	["p", "d"]	["p_2", "d_0"]
102	["p", "a"]	["p_3", "a_110"]
104	["p", "a", "+"]	["p_4", "a_010"]
105	["p", "t"]	["p'_1", "t_0"]
106	["p", "d"]	["p'_2", "d_0"]
107	["p", "a"]	["p'_3", "a_140"]
108	["p", "a", "+"]	["p'_4", "a_040"]

Таблица Б.3 – Входные данные и ожидаемые результаты тестов для класса OpenSyllableUnit

№	Входные данные	Ожидаемый результат
1	"КО+ЛОСА"	[["k", "o"], ["l", "a"], ["s", "a"]]
2	"КО+Л"	[["k", "o", "l"]]
3	""	[[]]
4	"О+"	[["o"]]
5	"К"	[["k"]]
6	"О+М"	[["o", "m"]]
7	"МО+"	[["m", "o"]]
8	"МНО+ГО"	[["m", "n", "o"], ["g", "a"]]
9	"МАЙК"	[["m", "a", "j", "k"]]
10	"А+ММО"	[["a"], ["m", "m", "a"]]

ОКОНЧАНИЕ ПРИЛОЖЕНИЯ Б

Таблица Б.4 – Входные данные и ожидаемые результаты тестов для класса SyllabicComplex3Unit

№	Входные данные	Ожидаемый результат
1	"КО+ЛОСА"	[[["k", "o", "l", "a"], ["s", "a"]]]
2	"КОЛО+СА"	[[["k", "a"], ["l", "o"], ["s", "a"]]]
3	"КО+ЛОН"	[[["k", "o"], ["l", "a", "n"]]]
4	"КО+ЛА"	[[["k", "o", "l", "a"]]]
5	"КОЛА+"	[[["k", "a"], ["l", "a"]]]

Таблица Б.5 – Входные данные и ожидаемые результаты тестов для класса SyllabicComplex2Unit

№	Входные данные	Ожидаемый результат
1	"МА+ЙКА"	[[["m", "a", "j"], ["k", "a"]]]
2	"МА+ЙНА"	[[["m", "a"], ["j", "n", "a"]]]
3	"ЖА+РКО"	[[["zh", "a", "r"], ["k", "a"]]]
4	"МА+ЙК"	[[["m", "a", "j", "k"]]]
5	"РМА+ДА"	[[["r", "m", "a"], ["d", "a"]]]
6	"АРМА+ДА"	[[["a", "r", "m", "a"], ["d", "a"]]]

Таблица Б.6 – Входные данные и ожидаемые результаты тестов для класса SyllabicComplex1Unit

№	Входные данные	Ожидаемый результат
1	"АЛЛОФО+Н"	[[["a", "l", "l", "o", "f", "o", "n"]]]
2	"АЯ+КС"	[[["a", "j", "a", "k", "s"]]]
3	"ТАКА+Я"	[[["t", "a"], ["k", "a", "j", "a"]]]
4	"ЛИ+НИЯ"	[[["l", "i"], ["n", "i", "j", "a"]]]
5	"Я+МА"	[[["j", "a"], ["m", "a"]]]
6	"РА=ДИОА=ЭРОНАВИГА+ЦИЯ"	[[["r", "a"], ["d", "i", "a", "a", "e"], ["r", "a"], ["n", "a"], ["v", "i"], ["g", "a"], ["c", "y", "j", "a"]]]
7	"НАИ+ВНЫЙ"	[[["n", "a", "i"], ["v", "n", "y", "j"]]]
8	"ДИОРА+МА"	[[["d", "i", "a"], ["r", "a"], ["m", "a"]]]
9	"БО+А"	[[["b", "o", "a"]]]
10	"РА=ДИОА=ЭРОНАВИГА+ЦИЯ"	[[["r", "a"], ["d", "i", "a", "a", "e"], ["r", "a"], ["n", "a"], ["v", "i"], ["g", "a"], ["c", "y", "j", "a"]]]