**Наименование проекта**

Программа для анализа и обработки речевых сигналов.

**Назначение проекта**

Программа предназначена для импорта, воспроизведения, обработки и анализа аудиозаписей речевых сигналов с возможностью сегментации на фонемы и выделения гласных, согласных и ударных звуков. Система обеспечивает визуализацию спектрограмм, работу с временными метками и экспорт результатов анализа.

**Пользовательские истории**

Как исследователь речи, я хочу сегментировать аудиозапись на фонемы, чтобы ускорить анализ акустических характеристик.

Как исследователь речи, я хочу автоматически выделять гласные и согласные фонемы, чтобы не тратить время на ручную разметку.

Как исследователь речи, я хочу определять ударные гласные, чтобы быстрее проводить фонетический анализ речи.

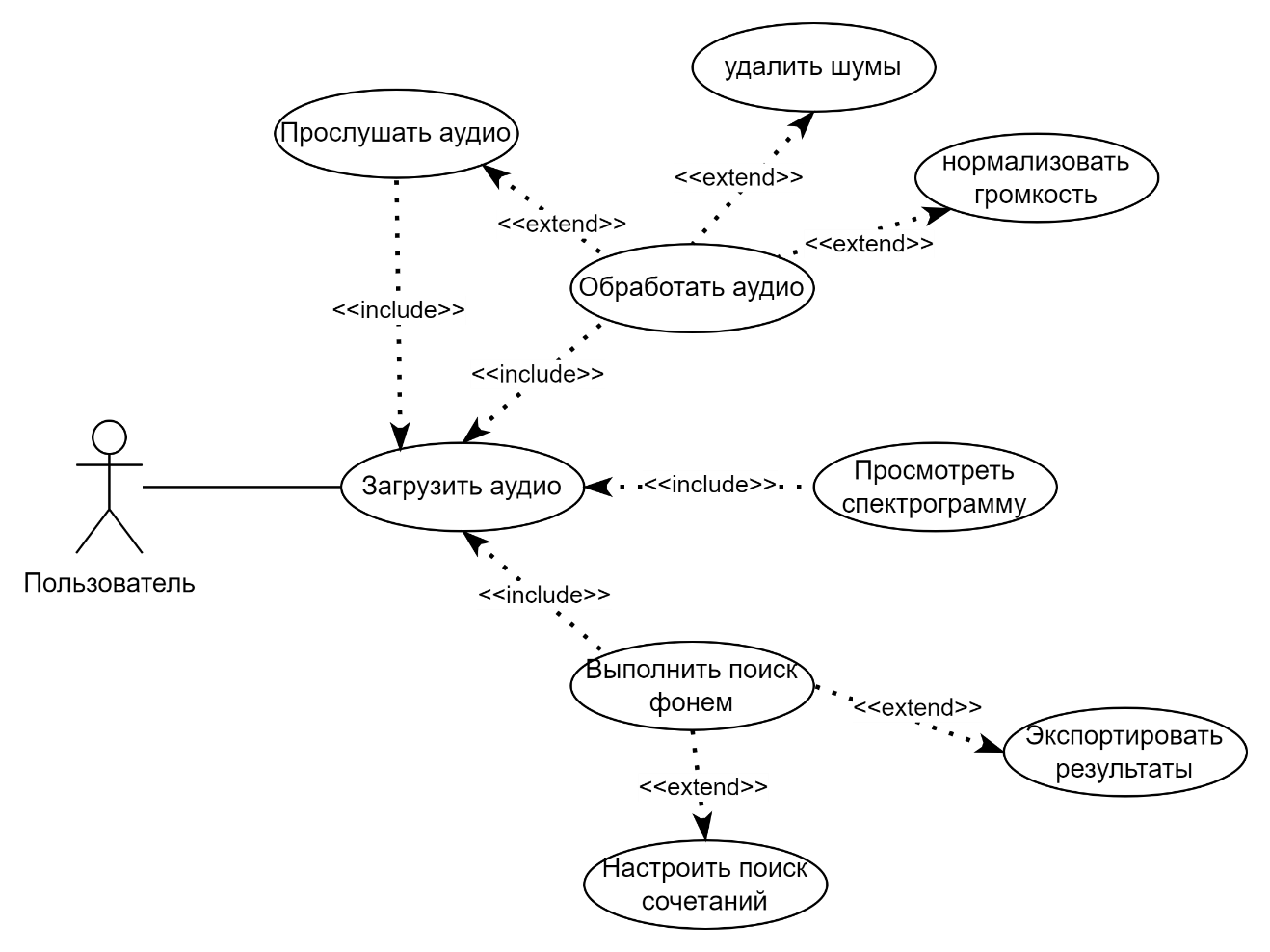
Как исследователь речи, я хочу искать звуковые сочетания и видеть их на спектрограмме, чтобы упростить работу с большим объёмом материала.

**Пользовательские сценарии**

*Основной пользовательский сценарий:*

1. Пользователь загружает аудиофайл в систему.
2. Система обрабатывает аудио, нормализуя громкость и удаляя шумы.
3. Система отображает спектрограмму записи.
4. Пользователь выбирает критерии поиска сочетаний.
5. Система выполняет сегментацию аудиосигнала на фонемы.
6. Система определяет гласные, согласные и ударные гласные.
7. Система дает расшифровку аудио в виде транскрипции и текста.
8. На основе введённых критериев система находит требуемые звуковые сочетания.
9. Система отображает результаты на спектрограмме и формирует список сочетаний с тайм-кодами.
10. Пользователь при необходимости экспортирует результаты в удобный формат.

**UseCase**



**Основной функционал:**

|  |  |  |  |
| --- | --- | --- | --- |
| **Название функции** | **Входные параметры** | **Выходные параметры** | **Описание функции** |
| Импорт аудиофайла | Путь к файлу (WAV, MP3, FLAC) | Загруженный аудиофайл | Загружает аудиофайл в рабочее пространство программы. |
| Воспроизведение аудио | Аудиофайл | Звук через колонки/наушники | Воспроизводит аудиозапись внутри программы. |
| Нормализация громкости | Аудиофайл | Аудиофайл с нормализованной громкостью | Уравнивает уровень громкости сигнала. |
| Фильтрация шумов | Аудиофайл | Аудиофайл с удалённым шумом | Удаляет посторонние шумы из аудиозаписи. |
| Обрезка/выделение фрагментов | Аудиофайл, временной диапазон | Фрагмент аудиофайла | Позволяет выделить и сохранить отдельные части записи. |
| Сегментация на фонемы | Аудиофайл | Список фонем с тайм-кодами | Делит сигнал на фонемы (минимальные единицы речи). |
| Определение гласных/согласных | Список фонем | Список с классификацией фонем | Отмечает фонемы как гласные или согласные. |
| Определение ударных гласных | Список гласных фонем | Список ударных фонем с тайм-кодами | Выделяет ударные гласные в записи. |
| Поиск звуковых сочетаний | Список фонем | Список сочетаний с тайм-кодами | Находит определённые фонетические сочетания в сигнале. |
| Визуализация спектрограммы | Аудиофайл | Графическое отображение спектрограммы | Отображает спектральный анализ сигнала. |
| Экспорт результатов | Список данных (фонемы, тайм-коды, сочетания) | Файл (TXT, XML, PDF) | Сохраняет результаты анализа в удобном формате. |

**Аналоги:**

## Аналог №1 – «Praat»

Praat – это бесплатная программа для анализа и синтеза речи, разработанная в 1991 году Полом Боэрсма и Давидом Венинком. Она широко применяется в фонетике и лингвистике, предоставляет возможности спектрального анализа, ручной разметки и работы с формантами.

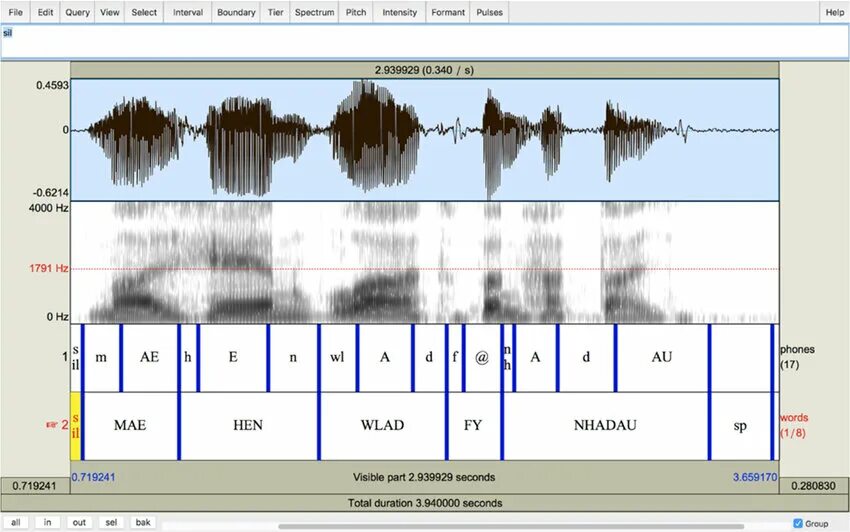


Рисунок 1 – Интерфейс программы «Praat»

Преимущества:

* программа бесплатна и активно используется в научных исследованиях;
* предоставляет большой набор инструментов для спектрального анализа;
* поддерживает русский язык в обработке аудиофайлов.

Недостатки:

* интерфейс устаревший и сложный для начинающих пользователей;
* отсутствует автоматическое определение фонем, гласных и согласных;
* сегментация речи выполняется вручную, что требует больших временных затрат.

## Аналог №2 – «Phonexia Speech Technologies»

Phonexia Speech Technologies – это коммерческий продукт, разрабатываемый одноимённой компанией с 2006 года. Система используется для распознавания речи, биометрической идентификации и анализа аудио. Имеет модуль Phoneme Recogniser, выполняющий разметку звуковых единиц.

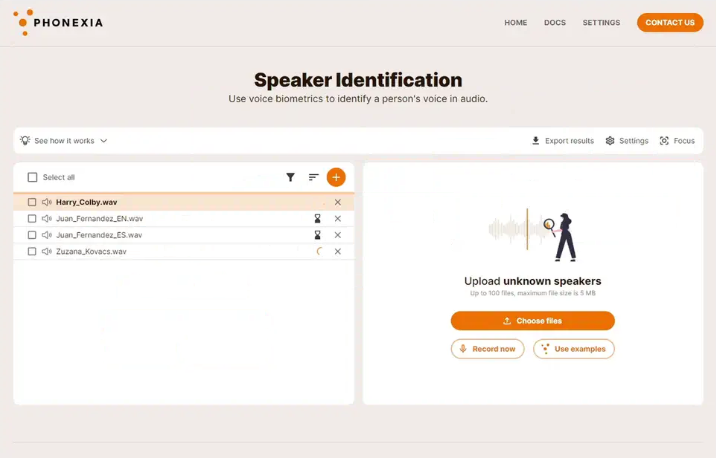


Рисунок 2 – Интерфейс программы «Phonexia Speech Technologies»

Преимущества:

* современный интерфейс и поддержка интеграции с другими системами;
* возможность автоматической работы с аудио и выделения фонем;
* применяется в криминалистике и корпоративной аналитике.

Недостатки:

* ограниченная поддержка русского языка;
* продукт является коммерческим и требует лицензии;
* программа не выполняет автоматического различения гласных, согласных и ударных фонем.

## Аналог №3 – «Gruut»

Gruut – это открытое программное обеспечение для фонематической транскрипции, разработанное сообществом энтузиастов. Оно предназначено в первую очередь для преобразования текста в последовательности фонем, что используется в системах синтеза речи.

Преимущества:

* полностью бесплатная и открытая лицензия;
* поддерживает множество языков, включая русский;
* интегрируется с проектами синтеза речи (например, TTS).

Недостатки:

* работает только по схеме «текст → фонемы», не анализирует аудио;
* отсутствует графический интерфейс, работа ведётся через код.