



Горшков Андрей Сергеевич,

Студент 4 курса, СПбГУАП, г. Санкт-Петербург

ИССЛЕДОВАНИЕ ХАРАКТЕРИСТИЧЕСКИХ СВОЙСТВ РАЗЛИЧНЫХ ТИПОВ ЗВУКОВ, ИСПОЛЬЗУЕМЫХ В РЕЧИ

Аннотация: В статье описано исследование характеристических свойств различных типов звуков, используемых в речи. Изучение акустических свойств речи возможно с помощью различных методов, включая преобразование акустических колебаний в электрический сигнал. Для анализа изменений акустического сигнала во времени используются динамические спектрограммы. Изучение характеристик звуков речи может быть полезным для создания алгоритмов сегментации речевого сигнала, таких как VAD-алгоритмы, которые определяют речевую активность и фонемную сегментацию. Результатом исследования являются характеристики звуков речи, исследованных в программе Praat, которые могут быть применены для алгоритмов распознавания речи.

Ключевые слова: Речь, VAD, распознавание речи, сегментация речевого сигнала.

Изучение акустических свойств речи возможно осуществить различными способами. Изначально акустические колебания необходимо преобразовать в электрический сигнал. Эту функцию выполняет электроакустический прибор – микрофон. Наблюдать за изменением параметров электрического сигнала можно несколькими способами: при помощи спектрографа, осциллографа, компьютерного ПО и других. В таблице 1 приведены наиболее часто встречаемые способы визуального представления звуков.

При помощи осциллограммы можно измерить амплитудные и временные параметры сигнала, подаваемого на его вход, но не частотные составляющие.



Измерив в определенной точке акустического сигнала частоту, можно получить мгновенную спектрограмму сигнала, при этом получится зависимость амплитуды от частоты.

Таблица 1

Основные виды визуального представления звуков

Вид представления	Ось	
	Абсцисса	Ордината
Осциллограмма	время	амплитуда
Мгновенная спектрограмма	частота	амплитуда
Динамическая спектрограмма	время	частота
	амплитуда передается степенью затемнения	

Для анализа изменений акустического сигнала во времени пользуются динамическими спектрограммами. Спектральный анализ можно называть наиболее распространенным способом анализа звуков. С его помощью можно определить относительные амплитуды частотных составляющих звука. Было установлено, что механизм речеобразования является квазистационарным источником звуков, то есть близкий по свойствам к стационарному, отличающийся от него медленным изменением во времени характером возбуждения и частотной характеристикой. Поэтому любая спектральная характеристика, пригодная для речевого сигнала, должна отражать не только спектральные особенности воспринимаемых значений, но и временные особенности.

Речевой сигнал человека состоит из различных типов звуков, включая шумные, вокализованные и взрывные звуки, а также паузы между ними. Смычка - это пауза, которая обычно находится перед взрывным звуком и может присутствовать между частями одного слова. Знание средней длительности смычки может быть полезным для VAD-алгоритма, чтобы снизить количество ложных определений фрагментов как пауз.



Статистические данные, полученные в результате исследования характеристических свойств звуков речи, могут быть использованы для создания различных алгоритмов сегментации. Для VAD-алгоритма крайне важны данные о продолжительности непрерывной речевой активности и средней продолжительности смычки. Фонемная сегментация является более технически сложным уровнем сегментации, и для ее создания необходимы статистические данные о параметрах фонемы в зависимости от ее положения в слове, относительно других звуков и относительно ударного гласного.

Когда мы произносим гласные звуки, наши голосовые связки работают активно, и воздушный поток свободно проходит через них. Благодаря этому у голоса возникают гармоники частоты ОТ. На спектрограмме гласного звука "а" в слове "работа" на рисунке 1, можно видеть изменение спектра со временем. Если голосовые связки вибрируют с разной частотой при производстве звука, то это может привести к частично периодической составляющей в частотной и временной структуре звука. Основная частота звука, которую вызывает вибрация голосовых связок, зависит от пола и возраста человека и может варьироваться от 50 Гц у взрослых мужчин до 350 Гц у женщин и детей.

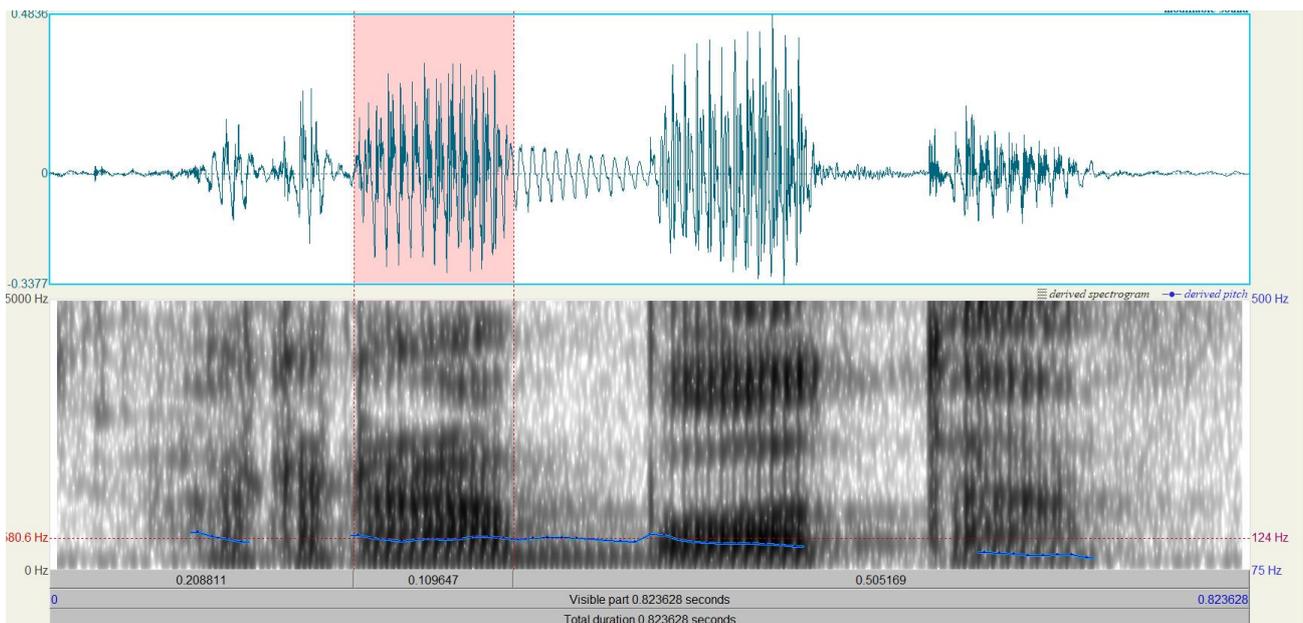


Рис. 1 – вокализованный гласный звук «а» слова «работа»



Не только гласные звуки, но и сонорные согласные, которые имеют меньше шума и больше тонового компонента, а также звонкие согласные могут создавать участки речи вокализованного типа. В случае использования вокализованных согласных звуков, формантная структура спектра сохранившаяся, но менее выраженными формантами, расположенными выше второй, может быть труднее различить. Кроме того, широкополосный шум на частотах от 2,5 кГц может наложиться на форманты звуков с высоким уровнем шума в своем составе.

На спектрограмме слова "Работа" показывается изменение спектра во времени для согласного дрожащего звука "р" на рисунке 2.

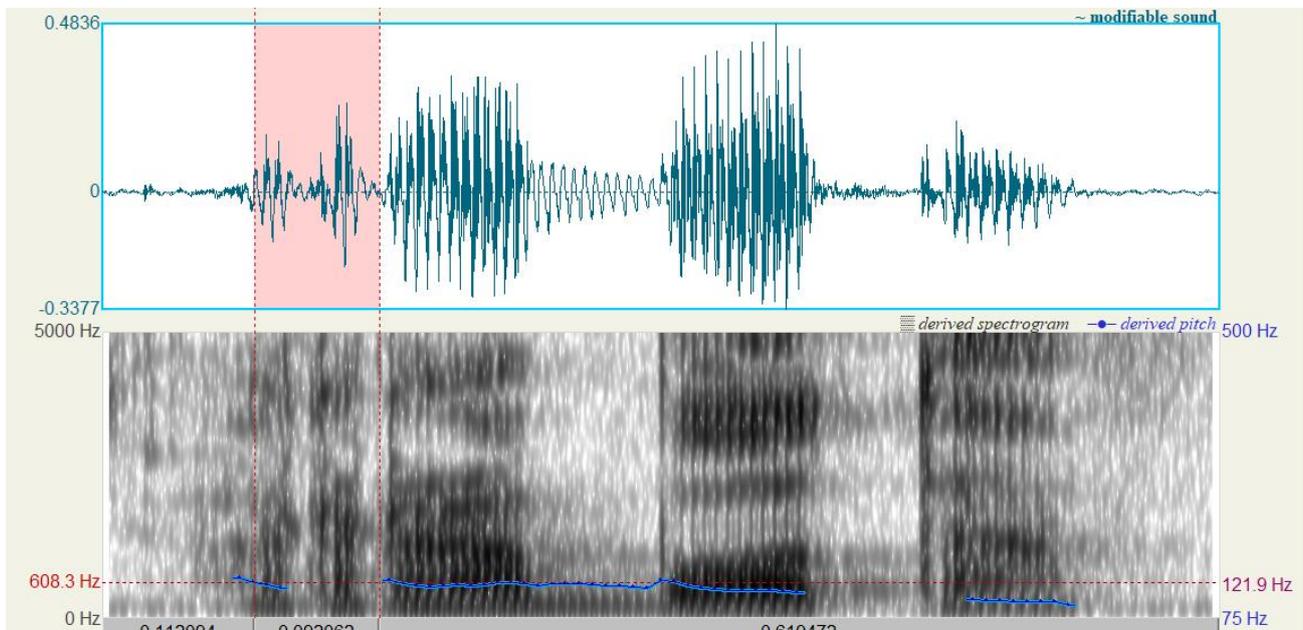


Рис. 2 – Вокализованный дрожащий звук «р» слова «работа»

Согласные звуки [ж], [з] и их палатализованные вариации являются аллофонами, которые обладают наибольшей шумовой составляющей среди вокализованных согласных. Их произношение вызывает значительные преграды в речевом аппарате, что приводит к преобладанию шумовой составляющей над тоном. Присутствие этих звуков в речи может отрицательно повлиять на точность работы алгоритма определения речевой активности, ухудшив ее качество.



Взрывные звуки, которые не содержат голоса, возникают благодаря кратковременному взрыву струи воздуха, который преодолевает полную преграду или смычок. Эти звуки имеют самую короткую длительность и вызывают резкое изменение особенностей сигнала, поэтому для их анализа используются менее короткие окна Фурье или другие методы. Роль резонансных полостей речевого аппарата, таких как глубина образования звука, играет важную роль при создании взрывных звуков. Звук [п], образовавшийся в заднеязычной полости, в спектре будут видны усиления интенсивности в формантных областях меньше, чем у звука [к], образовавшегося губами. Реализация звука [п] на рисунке 3 показывает его значительную низкочастотную составляющую и малую длительность по сравнению с другими звуками.

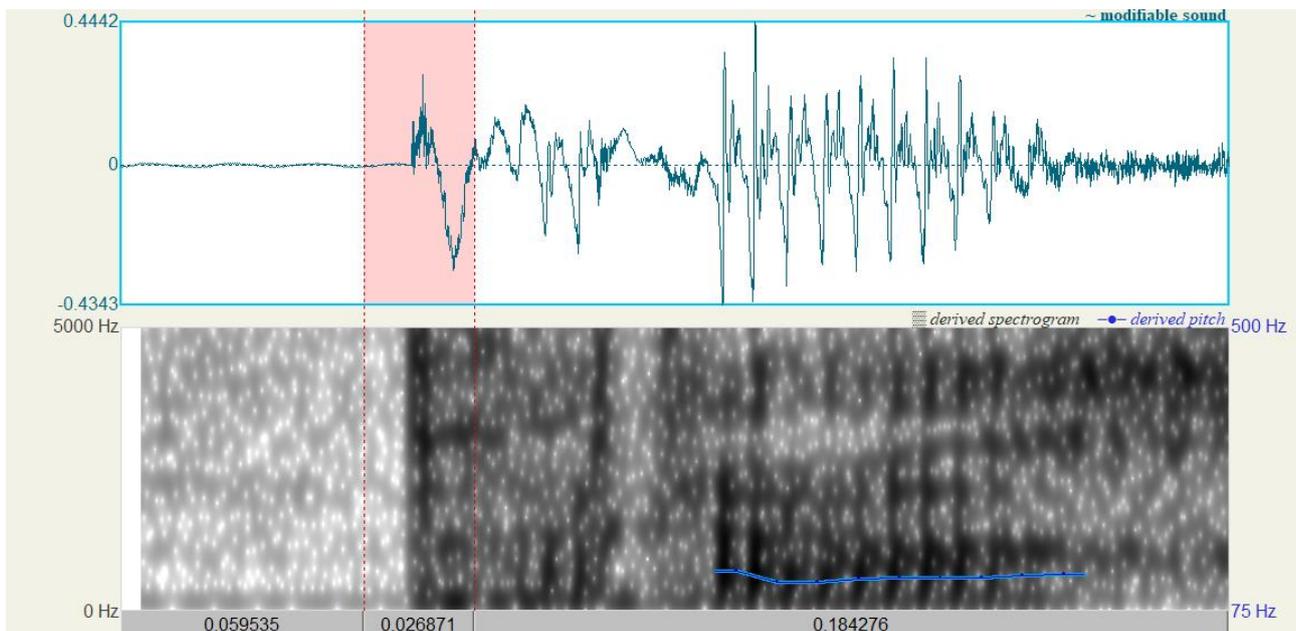


Рис. 3 - Невокализованный согласный взрывной звук «п»
в слове «Проход»

Гармоники ОТ отсутствуют в невокализованных шумных звуках, но если такой звук соединяется с соседним вокализованным звуком, то возможно обнаружить их присутствие. Шумные звуки не требуют участия голоса и



содержат только шумовые частоты выше 2 кГц. Из-за отсутствия голосовых связок формантная структура спектра будет слабо выражена. Однако резонансные полости речевого аппарата оказывают большое влияние на степень выраженности формант.

Заключение:

Исследование характеристических свойств звуков речи является важным направлением для понимания и анализа речевых данных. Различные типы звуков, такие как шумные, вокализованные и взрывные звуки, имеют уникальные особенности на акустических спектрограммах, которые могут быть использованы для разработки алгоритмов обработки речи. Изучение изменений акустического сигнала во времени с помощью динамических спектрограмм позволяет нам лучше понять структуру речи и выделить различные компоненты, такие как гласные, согласные и паузы. Это полезно для задач сегментации речевого сигнала и определения речевой активности. Характеристики звуков речи, такие как гармоники ОТ и формантная структура у гласных, высокочастотная шумовая составляющая у шумных звуков и резкие изменения особенностей сигнала у взрывных звуков, предоставляют ценную информацию для анализа и классификации звуков речи. Исследование акустических свойств звуков речи имеет широкий спектр применений, включая разработку систем автоматического распознавания речи, улучшение качества связи в телефонии, разработку алгоритмов удаления шума.

Список литературы:

1. Рабинер Л. Р., Шафер Р. В. Цифровая обработка речевых сигналов: Пер. с англ./Под ред. М. В. Назарова и Ю. Н. Прохорова. М.: Радио и связь, 1981. 496 с.
2. Джеймс Л. Фланаган. Анализ, синтез и восприятие речи: Пер. с англ./Под ред. А.А. Пирогова. М.: СВЯЗЬ, 1968. 395 с.
3. Акустика: Справочник / А.П. Ефимов, А.В. Никонов, М.А., Сапожков, В.И. Шоров; Под ред. М.А. Сапожкова. 2-е изд., перераб. и доп. М.: Радио и связь, 1989. 336 с.