

ФЕДЕРАЛЬНОЕ ГОСУДАРСТВЕННОЕ АВТОНОМНОЕ ОБРАЗОВАТЕЛЬНОЕ УЧРЕЖДЕНИЕ ВЫСШЕГО ОБРАЗОВАНИЯ
«БЕЛГОРОДСКИЙ ГОСУДАРСТВЕННЫЙ НАЦИОНАЛЬНЫЙ
ИССЛЕДОВАТЕЛЬСКИЙ УНИВЕРСИТЕТ»
(Н И У « Б е л Г У »)

ИНСТИТУТ ИНЖЕНЕРНЫХ ТЕХНОЛОГИЙ И ЕСТЕСТВЕННЫХ НАУК

КАФЕДРА ИНФОРМАЦИОННО-ТЕЛЕКОММУНИКАЦИОННЫХ
СИСТЕМ И ТЕХНОЛОГИЙ

**ИССЛЕДОВАНИЕ АЛГОРИТМОВ, ИСПОЛЬЗУЕМЫХ ПРИ СИНТЕЗЕ
РЕЧИ ДЛЯ ОБЕСПЕЧЕНИЯ ЕСТЕСТВЕННОСТИ ВОСПРОИЗВЕДЕНИЯ**

Выпускная квалификационная работа
обучающегося по направлению подготовки 11.03.02 Инфокоммуникационные
технологии и системы связи
очной формы обучения, группы 07001410
Ха Куанг Нинь

Научный руководитель
канд. техн. наук, доцент
кафедры
Информационно-
телекоммуникационных
систем и технологий
НИУ «БелГУ» Прохоренко Е.И.

Рецензент
кан. техн. наук, начальник
Отдела серийного сопровождения
АО НПП «СпецРадио»
к.ф-м.н.ю Туяков С.В.

БЕЛГОРОД 2018

ФЕДЕРАЛЬНОЕ ГОСУДАРСТВЕННОЕ АВТОНОМНОЕ ОБРАЗОВАТЕЛЬНОЕ УЧРЕЖДЕНИЕ
ВЫСШЕГО ОБРАЗОВАНИЯ
**БЕЛГОРОДСКИЙ ГОСУДАРСТВЕННЫЙ НАЦИОНАЛЬНЫЙ
ИССЛЕДОВАТЕЛЬСКИЙ УНИВЕРСИТЕТ**
(НИУ «БелГУ»)

ИНСТИТУТ ИНЖЕНЕРНЫХ ТЕХНОЛОГИЙ И ЕСТЕСТВЕННЫХ НАУК
КАФЕДРА ИНФОРМАЦИОННО-ТЕЛЕКОММУНИКАЦИОННЫХ СИСТЕМ И ТЕХНОЛОГИЙ
Направление подготовки 11.03.02 Инфокоммуникационные технологии и системы связи
Профиль Сети связи и системы коммутации

Утверждаю
Зав. кафедрой
«____» _____ 201_ г.

**ЗАДАНИЕ
НА ВЫПУСКНУЮ КВАЛИФИКАЦИОННУЮ РАБОТУ**

Ха Куанг Нинь

1. Тема ВКР «Исследование алгоритмов, используемых при синтезе речи для обеспечения естественности воспроизведения»

Утверждена приказом по университету от «____» _____ 201_ г. № _____

2. Срок сдачи студентом законченной работы _____._____

3. Исходные данные к работе:

Объект: речевой сигнал

4. Содержание расчетно-пояснительной записки (перечень подлежащих разработке вопросов):

- 4.1. Разработка алгоритма оценивания динамики изменения формант
- 4.2. Подготовка речевого сигнала
- 4.3. Исследование динамического изменения формант речевого сигнала
- 4.4. Алгоритм динамического изменения формант речевого сигнала
- 4.5. Синтез звуков с использованием структуры фонем
- 4.6. Экономическая оценка результатов исследования

5. Консультанты по работе с указанием относящихся к ним разделов

Раздел	Консультант	Подпись, дата	
		Задание выдал	Задание принял
4.1. – 4.5	канд. техн наук, доцент каф. ИТСиТ Прохоренко Е.И.		
4.6	канд. техн наук, доцент каф. ИТСиТ Болдышев А.В.		

7. Дата выдачи задания _____

Руководитель

*Кандидат технических наук, доцент
кафедры Информационно-телекоммуникационных
систем и технологий»
НИУ «БелГУ» _____ Прохоренко Е.И.
(подпись)*

Задание принял к исполнению _____ *Xa Kuанг Нинь*
(подпись)

СОДЕРЖАНИЕ

ВВЕДЕНИЕ	3
1 МЕТОДЫ СИНТЕЗА РЕЧЕВЫХ СИГНАЛОВ	4
2 ФОНЕМЫ, СПОСОБ ИХ ОПИСАНИЯ И ДИФФЕРЕНЦИАЦИИ	9
2.1 Основы речеобразования	9
2.2 Описание фонем на основе методов оценивания основного тона	11
3 РАЗРАБОТКА АЛГОРИТМА ОЦЕНИВАНИЯ ДИНАМИКИ ИЗМЕНЕНИЯ ФОРМАНТ	18
3.1 Подготовка речевого сигнала	18
3.2 Исследование динамического изменения формант речевого сигнала	24
3.3 Алгоритм динамического изменения формант речевого сигнала	34
4 СИНТЕЗ ЗВУКОВ С ИСПОЛЬЗОВАНИЕМ СТРУКТУРЫ ФОНЕМ	53
5 ЭКОНОМИЧЕСКАЯ ОЦЕНКА РЕЗУЛЬТАТОВ ИССЛЕДОВАНИЯ	69
5.1 Планирование работ по исследованию	69
5.2 Расчет расходов на оплату труда на исследование	70
5.3 Расчет продолжительности исследования	72
5.4 Расчет стоимости расходных материалов	72
5.5 Расчет сметы расходов на исследование	73
ЗАКЛЮЧЕНИЕ	77
СПИСОК ИСПОЛЬЗОВАННЫХ ИСТОЧНИКОВ	78

Изм.	Лист	№ докум.	Подпись	Дата	11070006.11.03.02.369.ПЗВКР	Лит.	Лист	Листов
Разработал	Ха Куанг Нинь				Исследование алгоритмов, используемых при синтезе речи для обеспечения естественности воспроизведения		2	79
Проверил	Прохоренко Е.И							
Рецензент	Туяков С.В.							
Н. Контроль	Прохоренко Е.И							
Утвердил	Жиляков Е.Г.							

ВВЕДЕНИЕ

Процесс изучения и построения систем синтеза речи становится все более и более популярен в последнее время. Много подходов и алгоритмов предложено в этой области. Если в первых подобных системах акцент делался на разборчивость речи, то теперь особое внимание уделяется ее естественности, интонационной насыщенности, эмоциональной окраске. Голос довольно точно передает окружающим информацию о текущем состоянии человека, о его переживаниях, отношении к фактам, самочувствии, а нередко - и о темпераменте, о чертах характера. Уловить эмоции позволяет тон голоса. А для понимания сообщения важны как сила голоса, так и его высота.

С развитием технологий автоматического синтеза речи, синтезированная речь становится все более и более естественной, приближенной к речи человека. Однако системы синтеза речи в современных человеко-машинных интерфейсах, системах виртуальной реальности и мультимедийного общения, по-прежнему обладают рядом недостатков, которые утомляют слушателя, не давая ощущения того, что с ними общается живой человек. Для уменьшения количества дефектов, присущих синтезированной речи, различными научными коллективами разрабатываются методы, позволяющие повысить естественность речи. [1]

Цель данной выпускной квалификационной работы - исследование алгоритмов, используемых при синтезе речи для обеспечения естественности воспроизведения.

Задачи:

- 1) исследование методов описания и дифференциации фонем;
- 2) исследование динамического изменения формант на протяжении звучания фонем;
- 3) исследование методов синтеза звуков с использованием частотной структуры фонем.

Изм.	Лист	№ докум.	Подпись	Дата	Лист
					3

11070006.11.03.02.369.ПЗВКР

1 МЕТОДЫ СИНТЕЗА РЕЧЕВЫХ СИГНАЛОВ

Синтез звука - процесс генерации звука, представленного в виде дискретного сигнала (сигнала, который может принимать лишь конечное число значений).

Синтез речи может потребоваться во всех случаях, когда получателем информации является человек. О качестве синтезатора речи, прежде всего, судят по его сходству с человеческим голосом, а также способности быть понятным. Самую простую синтезированную речь можно создавать путём объединения частей записанной речи, которые затем будут храниться в базе данных. Такой способ синтеза в настоящее время используется повсеместно.

Синтез речи по тексту или коду сообщения может быть использован в информационно-справочных системах, для помощи слепым и немым, для управления человеком со стороны автомата, при объявлениях об отправлении поездов и тому подобное, для выдачи информации о технологических процессах: в военной и авиакосмической технике, в робототехнике, в акустическом диалоге человека с компьютером. Как звуковой эффект нередко используется в создании электронной музыки. [2]

В настоящее время наиболее известными примерами таких систем являются системы «электронных помощников», например система Siri (Speech Interpretation and Recognition Interface) - облачный персональный помощник и вопросно-ответная система, программный клиент которой входит в состав iOS, watchOS, macOS, и tvOS компании Apple. Данное приложение использует обработку естественной речи, чтобы отвечать на вопросы и давать рекомендации. Siri приспосабливается к каждому пользователю индивидуально, изучая его предпочтения в течение долгого времени. [3]

Пример российской разработки - «Алиса» - виртуальный голосовой помощник, созданный компанией «Яндекс». Имитирует живой диалог, распознавая естественную речь (устную и письменную) и выдавая ответы на

Изм.	Лист	№ докум.	Подпись	Дата
------	------	----------	---------	------

11070006.11.03.02.369.ПЗВКР

Лист

вопросы (произнося их вслух и выводя на экран).

Многие особенности личности «Алисы» заданы набором фраз, сочинённых редакторами «Яндекса». Однако создатели подчёркивают, что «Алиса» не ограничивается набором заранее заданных редакторских ответов: нейронная сеть помощницы обучена на большом массиве русскоязычных текстов, в том числе сетевых диалогов. «Алиса» умеет отвечать эмоционально: например, в зависимости от контекста она может проявлять жизнерадостность или грустить.

«Алиса» распознаёт и синтезирует речь с помощью платформы SpeechKit, разработанной в Яндексе. [4]

Существует много методов синтеза:

1) Субтрактивный синтез

Субтрактивный синтез (Subtractive Synthesis) - это метод синтеза, основанный на вычитании элементов друг из друга. В синтезаторах определяющим элементом субтрактивного синтеза является наличие фильтров. Фильтр «вырезает» часть спектра из звука, формируя нужную тембральную окраску.

Субтрактивный синтез часто называют «аналогово-субтрактивным», отдавая дань прародителям всех синтезаторов - аналоговым инструментам. [5]

2) Формантный синтез

Формантный синтез (Formant Shaping Synthesis) - синтез, в основе которого лежит воссоздание в спектре звука акустических резонансов, формирующихся из групп обертонов и называемых формантами (подробнее об обертонах здесь).

Изначально формантный синтез ассоциировался с синтезом речи. Речевой аппарат человека устроен таким образом, что нос и глотка являются, по сути, природными фильтрами. В процессе формирования звука эти фильтры чисто механически принимают определенные формы, исходя из которых в звуковом потоке образуются группы акустических резонансов, «сгустки» обертонов. В

Изм.	Лист	№ докум.	Подпись	Дата	Лист
					5

11070006.11.03.02.369.ПЗВКР

зависимости от частот, на которых появляются резонансы, и от структуры взаимодействия частот и амплитуд резонансов в спектре, наше ухо различает те или иные фонемы, то есть звуковое выражение букв. На картинке изображены формы резонаторов для воссоздания звуков разных гласных букв. Таким образом, имея возможность проанализировать спектры всевозможных фонем, мы получаем возможность их воссоздать средствами синтеза, или, проще говоря, синтезировать.

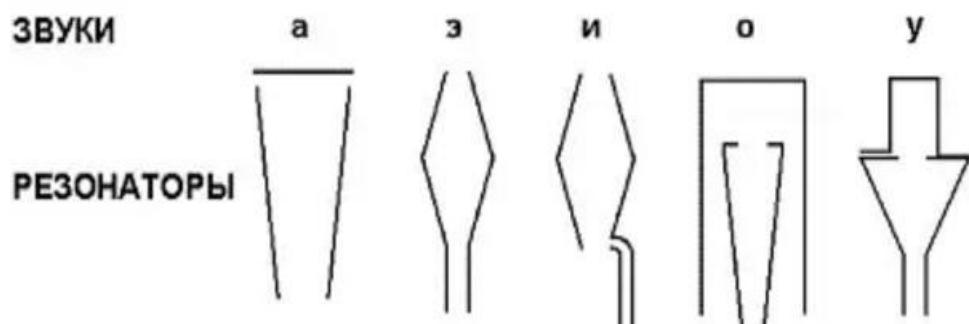


Рисунок 1.1 - Формы резонаторов для воссоздания звуков разных гласных букв

Польза этого метода была очевидна для передачи сигналов, несущих речевую информацию, на большие расстояния посредством кодирования их в пункте передачи и раскодирования по обратному алгоритму в пункте приема. На этом построен принцип работы вокодера, который при анализе речи разбивает спектр на полосы, выявляет в них резонансы, а затем преобразует обратно в звуковой сигнал. Конечно, полного сходства с источником сигнала получить невозможно, но добиться доступности для понимания передаваемой информации вполне реально.

Каждая фонема содержит в себе несколько формант. Чем больше формант воссоздает вокодер, тем яснее и доступнее передаваемая информация. Гласные звуки становятся разборчивыми при воссоздании уже первых трех групп резонансов. А при воспроизведении четырех уже можно говорить о довольно точной передаче речи. Пока человек произносит одно слово, резонансы его голоса постоянно перемещаются по всему частотному спектру в зависимости от фонемного состава этого слова.

Изм.	Лист	№ докум.	Подпись	Дата
------	------	----------	---------	------

11070006.11.03.02.369.ПЗВКР

Лист

6

Важной особенностью формант является неизменность их частот для каждой конкретной фонемы вне зависимости от высоты извлекаемой ноты. Например, форманты гласной «о» появляются на частотах 275, 850, 2400 герц, форманты «а» - на частотах 575, 900, 2450 герц, и так далее.

Зная это, совершенно очевидным становится факт, что для формирования фонем совершенно необязательно иметь источник речи в виде человека. Можно просто создать резонансы на нужных частотах для придания звуку соответствующей фонемной окраски. Функция выделения резонансов возлагается на специальные формантные фильтры, работающие по тому же принципу, что и человеческая носоглотка, однако уже не механическим способом, а электронным. [6]

3) Частотная модуляция

Частотная модуляция, FM-синтез (Frequency Modulation, FM-synthesis) - это тип синтеза, при котором тембр звука формируется воздействием одной простой волны на другую с целью изменения ее частоты. Под воздействием частотной модуляции возникают более спектрально богатые и сложные звуки, которые невозможно получить другими типами синтеза.

Модуляцией называется воздействие одного сигнала на другой с целью изменения его формы. Сигнал, осуществляющий модуляцию, называется модулирующим, а сигнал, который подвергается воздействию модуляции - несущим. [7]

4) Аналоговое моделирование

Аналоговое моделирование (Analog Modeling Synthesis) - это имитация аналогового синтеза в цифровой среде на основе использования специализированных процессоров обработки сигнала (DSP - Digital Signal Processor).

Синтезаторы, построенные по принципу аналогового моделирования, называют также «виртуальными аналогами», или VA (Virtual Analog). Такое название происходит от латинского *virtus*, что означает «воображаемый»,

						Лист
Изм.	Лист	№ докум.	Подпись	Дата	11070006.11.03.02.369.ПЗВКР	7

«мнимый». [8]

5) Векторный синтез

Векторный синтез (Vector Synthesis) - тип синтеза, построенный на динамическом морфинге (перетекании) аудиосигналов для получения более богатых и сложных тембров.

По своей сути векторный синтез вряд ли можно назвать самостоятельным видом синтеза, потому что прямого отношения к генерации сигнала этот термин не имеет, оставляя за собой лишь функцию управления морфингом между тембрами.

В «классическом» представлении морфинг осуществляется между четырьмя различными тембрами посредством движения джойстиком, либо с помощью управления внутренними программируемыми алгоритмами. В крайних значениях джойстика по осям X и Y размещаются соответствующие программы (темперы), обозначаемые обычно как A, B, C, D. Двигая джойстик внутри обозначенной матрицы, можно получить различные соотношения сигналов этих тембров. Манипуляции джойстиком позволяют получить интересные звуковые эффекты и уникальные тембы. Программно морфинг осуществляется при помощи LFO или генераторов огибающих (EG). [9]

6) Линейно-арифметический синтез

Линейно-арифметический синтез (Linear Arithmetic Synthesis, LA) - тип синтеза, основанный на формировании звука путем совмещения семплированной атаки с генерируемой остальной частью сигнала.

Безусловно, как и некоторые другие типы, линейно-арифметический синтез является гибридным, поскольку совмещает в себе два метода синтеза для получения сложного тембра. [10]

Изм.	Лист	№ докум.	Подпись	Дата
------	------	----------	---------	------

11070006.11.03.02.369.ПЗВКР

Лист

8

2 ФОНЕМЫ, СПОСОБ ИХ ОПИСАНИЯ И ДИФФЕРЕНЦИАЦИИ

2.1 Основы речеобразования

Звуки речи образуются речевым аппаратом человека. Речевой аппарат - это совокупность и взаимодействие органов человека, необходимых для производства речи. Он состоит из двух отделов: центрального и периферического. Центральный отдел - это головной мозг с его корой, подкорковыми узлами, проводящими путями и ядрами соответствующих нервов. Перефериеский отдел - это вся совокупность исполнительных органов речи (глотка, ротовая полость с языком, лёгкие, носовая полость, губы, зубы), включающая в себя кости, хрящи, мышцы и связки, а также периферические чувственные и двигательные нервы, при помощи которых осуществляется управление работой указанных органов.

В зависимости от работы речевого аппарата звуки речи подразделяются на шумы и тоны: тоны в речи возникают в результате колебания голосовых складок; шумы образуются в результате непериодических колебаний выходящей из лёгких струи воздуха. Тонами являются обычно гласные; почти же все глухие согласные относятся к шумам. Звонкие согласные образуются путём слияния шумов и тонов.

Звуки речи человека генерируются, как правило, артикуляционным аппаратом. В общем его математическую модель можно представить в виде возбуждающих генераторов тонового и белого шума и группы фильтров, модуляторов и ключей (рот, нос, язык, губы), обеспечивающих фильтрацию и формирование определённого ощущения звука.

Речевой аппарат человека при генерации речи использует следующие физические принципы для получения различных типов звуков:

- гласный - в этом случае голосовая щель генерирует звуковые импульсы;

Изм.	Лист	№ докум.	Подпись	Дата
------	------	----------	---------	------

11070006.11.03.02.369.ПЗВКР

Лист

9

- шипящий согласный - в этом случае голосовая щель отключена и артикуляционный аппарат формирует шумовой сигнал;
- смешанные шипяще-тоновые звуки типа [з] [ж], где одновременно присутствует шумовая составляющая модулированная голосовой щелью, или типа [р], где модулируется тоновый сигнал;
- сонорные звуки типа [л] [м] [н];
- взрывной согласный - генерация звука основана на перекрытии потока воздуха артикуляционными органами и последующем акустическом ударе;
- пауза - отсутствие звука, длина паузы влияет на ощущение следующего за ней звука;
- изменение параметров артикуляции в процессе генерации (их динамика) также создаёт ощущение определённого звука (дифтонги);
- интонация - относительное изменение основного тона.

Шумы и тоны исследуются по их высоте, тембру, силе и многим другим характеристикам. Важнейшей характеристикой является частотный состав звука. Не маловажной характеристикой речевого сигнала является так же основной тон. Эта характеристика представляет собой низко-частотную модуляцию сигнала, параметры которой легко измеряются (установлено, что частота основного тона разных людей (мужчин, женщин, детей) находится в диапазоне 20 - 400 Гц). [11]

Фонема - основная единица звукового строя языка. Звуковой состав различных языков имеет свои особенности. В русском языке насчитывают 41 фонему, из них 6 гласных и 35 согласных (в английском - 20 гласных и 24 согласных, в французском - 15 гласных и 20 согласных). Можно сказать, что фонема - наименьшая языковая единица, имеющая смыслоразличительное значение. Из последовательности фонем строятся слова. Смысл высказывания выражается посредством цепочки слов.

Под формантами понимаются частотные резонансы (полюса передаточной функции) речевой акустической системы. Параметры формант

						Лист
Изм.	Лист	№ докум.	Подпись	Дата	11070006.11.03.02.369.ПЗВКР	10

(частота, ширина, уровень) определяются акустическими свойствами системы. Наиболее важный параметр - частота форманты, тесно связан с геометрической конфигурацией речевого тракта. Поскольку в процессе речи конфигурация речевого тракта меняется, то соответственно меняются формантные частоты.

Для удовлетворительного синтеза речи обычно нужны две - четыре формантные частоты. Они лежат в диапазоне от 200 (первая форманта мужского голоса) до 2000 Гц (третья форманта женского голоса). Точным расположением формантных частот в звуковом спектре и определяется звук, который мы интерпретируем как речь. Причем, все формантные частоты присутствуют в речи одновременно и непрерывно перемещаются вверх-вниз по частотному спектру в соответствии с особенностями произносимого слова. Поэтому, слушая говорящего человека, вы слышите звук не какой-либо одной частоты, а множество обертонов, которые образуются при фильтрации импульсов, формируемых на выходе голосового тракта. [12]

2.2 Описание фонем на основе методов оценивания основного тона

Звуки речи делятся на звонкие и глухие. Звонкие звуки образуются с участием голосовых связок, в этом случае находящихся в напряженном состоянии.

Под напором воздуха, идущего из легких, они периодически раздвигаются, в результате чего создается прерывистый поток воздуха. Импульсы потока воздуха, создаваемые голосовыми связками, с достаточной точностью могут считаться периодическими. Соответствующий период повторения импульсов называют периодом основного тона голоса T_0 - а обратную величину $1/T_0$ - частотой основного тона. Если связки тонкие и сильно напряжены, то период получается коротким и частота основного тона высокой; для толстых, слабо напряженных связок - низкой. Частота основного тона для всех голосов лежит в пределах 20 - 400 Гц.

Изм.	Лист	№ докум.	Подпись	Дата	Лист
					11070006.11.03.02.369.ПЗВКР 11

При произнесении речи она непрерывно изменяется в соответствии с ударением, подчеркиванием звуков и слов, а также с проявлением эмоций (вопрос, восклицание, удивление и т.д.). Изменение частоты основного тона называется интонацией. У каждого человека свой диапазон изменения основного тона (обычно он бывает немногим более октавы) и своя интонация. Последняя имеет большое значение для узнавания говорящего. [13]

Изменение частоты основного тона влечет за собой квазипропорциональное изменение других формант фонемы, как показано на рисунке 2.1.

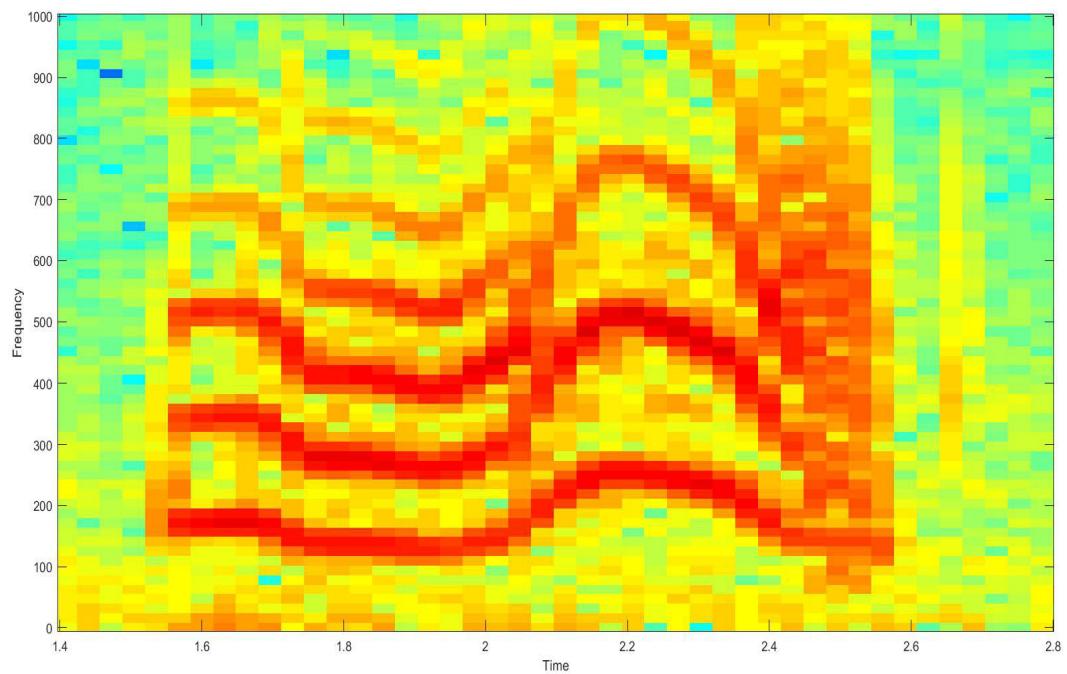


Рисунок 2.1 - Спектrogramma фонемы «нет»

На спектограмме отчетливо видны несколько формант, динамика изменения частоты которых схожа.

Таким образом, описание динамического изменения параметров формант на всем протяжении звучания фонемы, может являться основанием для синтеза звуков речи с сохранением естественности воспроизведения.

Для автоматического определения параметров формант, а именно частоты,

Изм.	Лист	№ докум.	Подпись	Дата
------	------	----------	---------	------

11070006.11.03.02.369.ПЗВКР

Лист

12

амплитуды и фазы, можно взять за основу один из существующих методов оценивания частоты основного тона.

Оценивание периода (или частоты) основного тона (ЧОТ) является одной из наиболее важных задач в обработке речи. Для решения проблем, связанных с измерением основного тона, были разработаны самые разнообразные методы.

В амплитудном методе на стационарном участке вокализованного звука при малом уровне шумов форма речевого колебания почти точно повторяется на каждом очередном периоде основного тона. Расстояние между глобальными максимумами (максимумами максиморумами) речевого сигнала можно приблизительно считать равным периоду основного тона. Основная трудность алгоритмов амплитудной селекции состоит в необходимости подавления локальных ложных максимумов. Этого можно добиться за счёт повышения порога срабатывания в схеме поиска максимумов. Однако при этом увеличивается вероятность пропуска истинного максимума. Очевидно, что как пропуск, так и потеря максимума может привести к существенным искажениям синтезированного звука. Повысить надёжность определения периода основного тона можно, например, добавив второй канал амплитудной селекции, выделяющий положение минимумов речевого сигнала. Главным достоинством устройств временной селекции является чрезвычайная простота реализации. Основные недостатки: низкая точность и неустойчивость определения основного тона (даже при относительно небольшом уровне шумов).

Корреляционные методы определения периода основного тона речевого сигнала базируются на оценке среднего значения периода пульсаций квазипериодической корреляционной функции (или, в частном случае, вычислении первого глобального максимума корреляционной функции). Частота основного тона f_1 рассчитывается по соотношению:

$$f_1 = \frac{f_d}{N_{p-1}} \quad (2.1)$$

Изм.	Лист	№ докум.	Подпись	Дата
------	------	----------	---------	------

11070006.11.03.02.369.ПЗВКР

Лист

13

где f_d - частота дискретизации; \overline{N}_{p-1} - среднее число отсчётов корреляционной функции, через которое пульсации повторяются; p - число глобальных максимумов корреляционной функции, взятых для усреднённой оценки \overline{N}_{p-1} .

В частном случае, основанном на поиске первого глобального максимума, применяется следующий подход. Пусть речевой сигнал представлен в виде последовательности отсчётов. Для вокализованных звуков можно считать, что временной вид речевого колебания почти точно повторяется на каждом очередном периоде основного тона. В качестве оценки периода основного тона $T = 1/f_1$, выраженной в числе отсчётов, выбирают значение, минимизирующее целевую функцию, которая определяется как сумма квадратов разностей между отсчётами сигнала и отсчётами сигнала, смещёнными на некоторое число отсчётов. Если предположить, что энергия речевого сигнала не меняется на участке квазистационарности, то оценка периода основного тона должна максимизировать корреляционную функцию. Данный подход обеспечивает существенно более высокую достоверность определения периода основного тона по сравнению с методами временной селекции. При этом следует отметить значительную вычислительную сложность данного алгоритма. Существуют его модификации, основанные на вычислении взаимной корреляционной функции с подобранный функцией, которая клиппирует речевой сигнал на три уровня $\{-1, 0, 1\}$ (трёхуровневый ограничитель). Таким образом, можно упростить алгоритм, сделав его пригодным для аппаратной реализации, тогда вычислитель взаимной корреляционной функции можно построить без умножителя. Рассмотренные корреляционные методы оценивания периода основного тона имеют общий недостаток: неустойчивую работу в случае, когда речевой сигнал модулирован по амплитуде. Энергия же реальной, т.е. эмоционально окрашенной речи изменяется даже на квазистационарных

Изм.	Лист	№ докум.	Подпись	Дата
------	------	----------	---------	------

11070006.11.03.02.369.ПЗВКР

Лист

14

участках, соответствующих одной фонеме. В этом случае применяется модифицированная целевая функция, в которой смещённые отсчёты сигнала умножаются на некоторый параметр, имеющий смысл коэффициента усиления. Метод позволяет получить достаточно точную оценку основного тона, которая плавно меняется во времени в соответствии с изменениями голоса. Поэтому данный алгоритм используется в стандарте G.723, регламентирующем способ сжатия речевого сигнала для видеоконференций.

Спектральный метод основан на том, что при вокализованном возбуждении речевого тракта в спектре сигнала присутствуют пики на частотах, кратных частоте основного тона. Если построить дискретное преобразование Фурье с достаточно малым шагом дискретизации по частоте, то можно попытаться в качестве оценки частоты основного тона использовать частоту, соответствующую максимальному значению энергии спектра. Поиск максимума следует производить в интервале 20 - 400 Гц. Однако часто возникает ситуация, когда в указанной полосе лежит и вторая гармоника основного тона, иногда даже с большей энергией. В этом случае она будет ошибочно принята за оценку основного тона. Чтобы избежать этого, обычно ищут максимум не спектра, а некоторой нелинейной функции от спектра. Эта функция, как правило, представляет собой сумму сжатых по частоте в несколько раз логарифмов спектра мощности. Суть идеи состоит в том, что для истинной частоты основного тона вторая гармоника второго слагаемого сложится с первой гармоникой первого слагаемого и усилит её. Аналогично для третьего слагаемого и т. д. В результате для вокализованного звука будет иметь место ярко выраженный пик функции (от спектра) на частоте основного тона, а для невокализованного звука суммирование будет иметь хаотический характер.

Заметим, что в общем случае оценка значений спектра является несостоительной и может иметь большие погрешности. Для уменьшения ошибки оценки спектральных составляющих, например, нормированной спектральной плотности мощности (также часто используют дискретное

Изм.	Лист	№ докум.	Подпись	Дата
------	------	----------	---------	------

11070006.11.03.02.369.ПЗВКР

Лист

15

преобразование Фурье, или быстрое преобразование Фурье), применяют методику спектральных окон. Выбор спектрального окна (весовой функции) при анализе определяется в результате компромисса между разрешающими способностями по частоте и во времени. Следует отметить, что для спектральных методов применение нелинейного преобразования спектра и окон может вносить большие смещения, что существенно ухудшает точность оценки.

Относительно новый развивающийся метод оценки частоты основного тона речевого сигнала **на основе вейвлет - преобразования** базируется на применении непрерывного или дискретного вейвлет - преобразования. Следует отметить положительную сторону данного метода: для генерированного эталонного четырёхгармонического сигнала относительная погрешность оценки частоты основного тона и формантных частот методом вейвлет-преобразования не превышала 0,38%. Недостатки данного метода: необходимость корректировки окна преобразования под каждую оцениваемую частоту; сложность алгоритма реализации метода; большие вычислительные затраты.

Метод оценивания основного тона **на основе кепстрального анализа** состоит в вычислении и анализе кепстраграфа - обратного преобразования Фурье логарифма спектра мощности сигнала. Однако данный метод имеет ряд существенных недостатков, таких как: необходимость применения дополнительной методики для вычисления порога для оценки периода основного тона в области возможных значений; работа в нереальном масштабе времени; необходимость применения временных окон и операций сглаживания; низкая точность оценки при сильной узкополосности гармоники основного тона.

Среди методов оценки **на основе линейного предсказания** обычно используют метод обратной фильтрации (обратный линейный фильтр). При приближении частоты повторения в обратном фильтре к частоте основного тона

Изм.	Лист	№ докум.	Подпись	Дата
------	------	----------	---------	------

11070006.11.03.02.369.ПЗВКР

Лист

16

происходит всё более и более сильное выравнивание спектра. Одна из трудностей - это постоянный расчёт спектра речевого сигнала при подстройке частоты повторения в обратном фильтре. Метод даёт удовлетворительные оценки, пока спектр выравнивается достаточно хорошо, однако здесь стоит новая задача определения степени равномерности спектра после обратной фильтрации. Также следует отметить, что при частотах основного тона выше 200 Гц данный метод оценки приводит к плохим результатам. [14]

В данной работе, при разработке алгоритма оценивания динамики изменения формант, были использованы амплитудный и спектральный методы.

Изм.	Лист	№ докум.	Подпись	Дата
------	------	----------	---------	------

11070006.11.03.02.369.ПЗВКР

Лист

17

3 РАЗРАБОТКА АЛГОРИТМА ОЦЕНИВАНИЯ ДИНАМИКИ ИЗМЕНЕНИЯ ФОРМАНТ

3.1 Подготовка речевого сигнала

Для исследования динамического изменения параметров формант (частоты, амплитуды, фазы) в фонемах речевых сигналов был разработан алгоритм для предварительной подготовки речевых сигналов. Алгоритм состоит из нескольких этапов, в результате выполнения которых на определенных временных отрезках выбираются информативные участки звука, а спектр сигнала ограничивается некоторым диапазоном частот, который подлежит описанию.

Этапы алгоритма обработки сигнала:

- разделение сигнала на отрезки;
- фильтрация;
- выделение информативных участков;
- спектральный анализ.

Этап 1. Разделение сигнала на отрезки

Исходный сигнал, длиной L отсчетов разделяется на равные отрезки длиной N .

В качестве примера исходного сигнала на рисунке 3.1 приведен сигнал “нет” длиной $L = 56000$ отсчетов.

Сигнал “нет” был разделен на $k = [L/N] = 7$ отрезков длительностью $N = 8000$ отсчетов (рисунок 3.2). Здесь квадратные скобки означают округление до целого.

На каждом из отрезков будет проводиться оценка параметров формант, длительность отрезка может быть выбрана из соображений необходимой точности динамической оценки, а так же исходя из естественной длительности

Изм.	Лист	№ докум.	Подпись	Дата	Лист
					18

11070006.11.03.02.369.ПЗВКР

звуков речи.

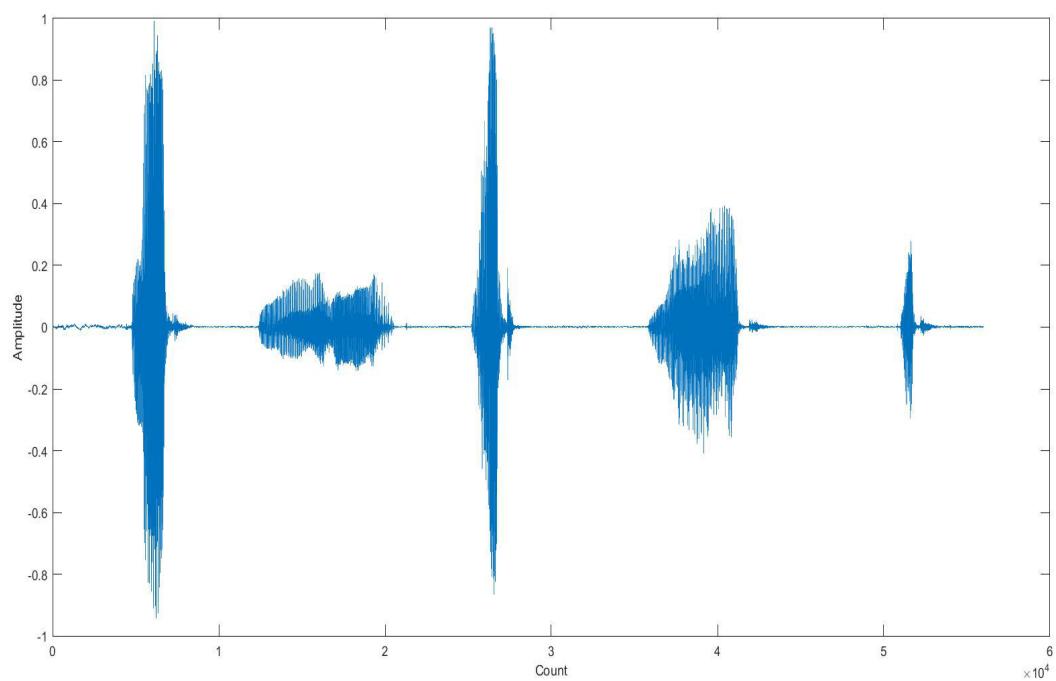


Рисунок 3.1 - Исходный сигнал “нет” длиной $L = 56000$ отсчётов

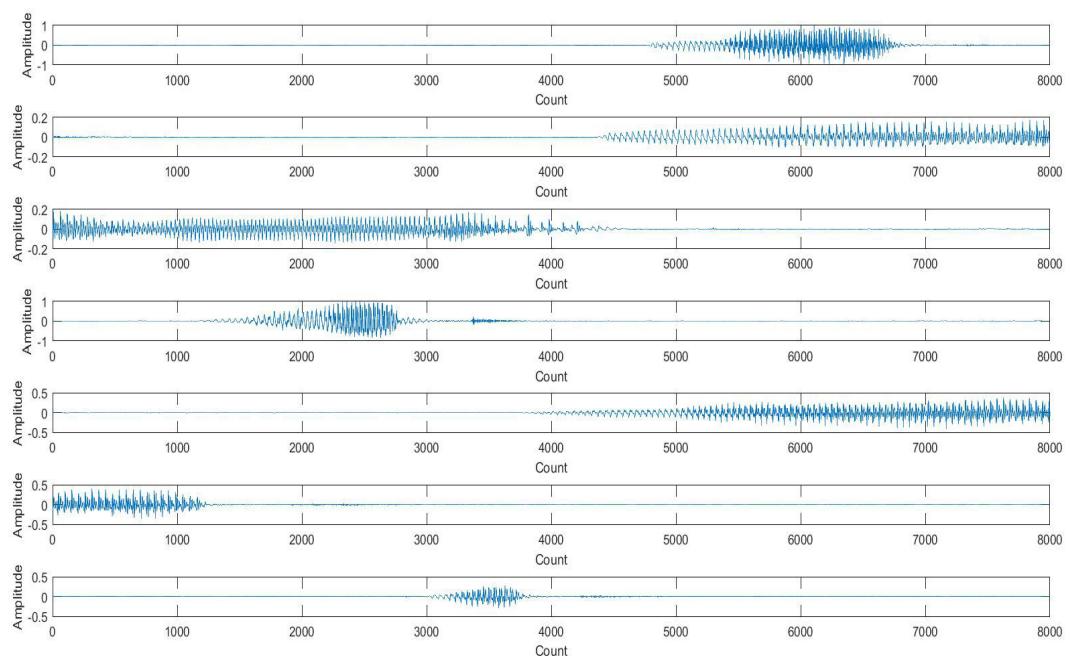


Рисунок 3.2 - Разделение сигнала на $k = 7$ отрезки

Изм.	Лист	№ докум.	Подпись	Дата
------	------	----------	---------	------

11070006.11.03.02.369.ПЗВКР

Лист

19

Этап 2. Фильтрация

Фильтрация осуществляется полосовым фильтром с полосой пропускания равной выбранному для исследования диапазону. Например для оценивания частоты основного тона диапазон составляет 20 - 400 Гц.

При этом рекомендуемые характеристики полосового КИХ-фильтра, определены эмпирически:

- длина импульсной характеристики фильтра $le = 512$;
- переходная полоса фильтра $c = 0.0025$;
- нижняя нормированная частота среза фильтра $v1 = 0.005$;
- верхняя нормированная частота среза фильтра $v2 = 0.100$.

На рисунке 3.3 показаны АЧХ и ФЧХ фильтра, полученного в соответствии с заданными характеристиками.

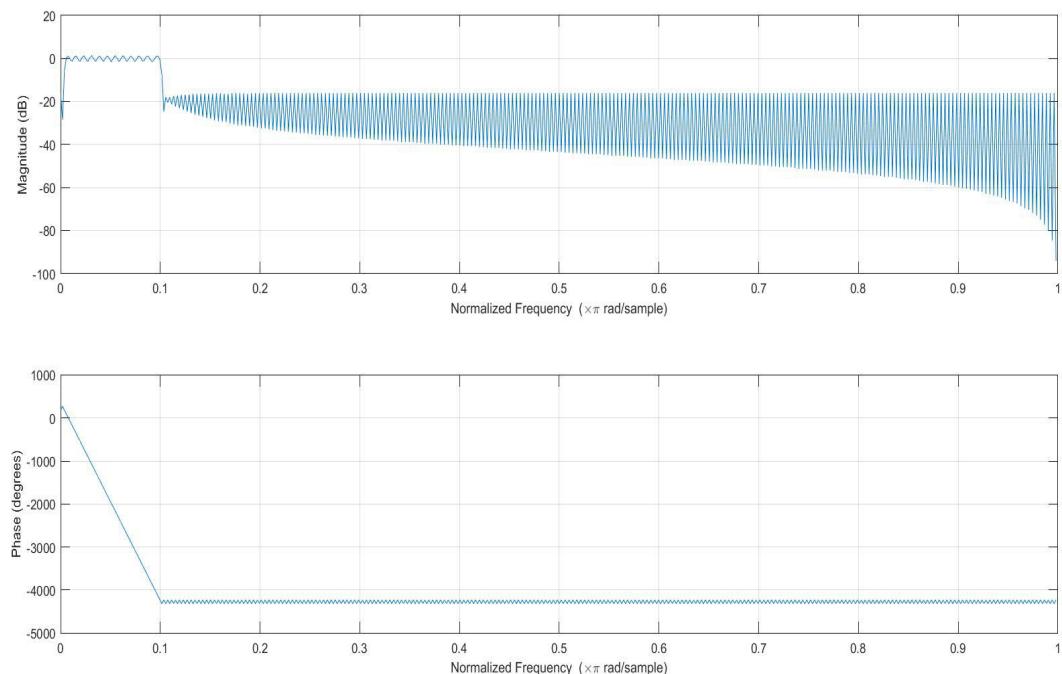


Рисунок 3.3 - АЧХ и ФЧХ фильтра

На рисунке 3.4 показаны отрезки сигнала, полученные в результате фильтрации.

Изм.	Лист	№ докум.	Подпись	Дата
------	------	----------	---------	------

11070006.11.03.02.369.ПЗВКР

Лист
20

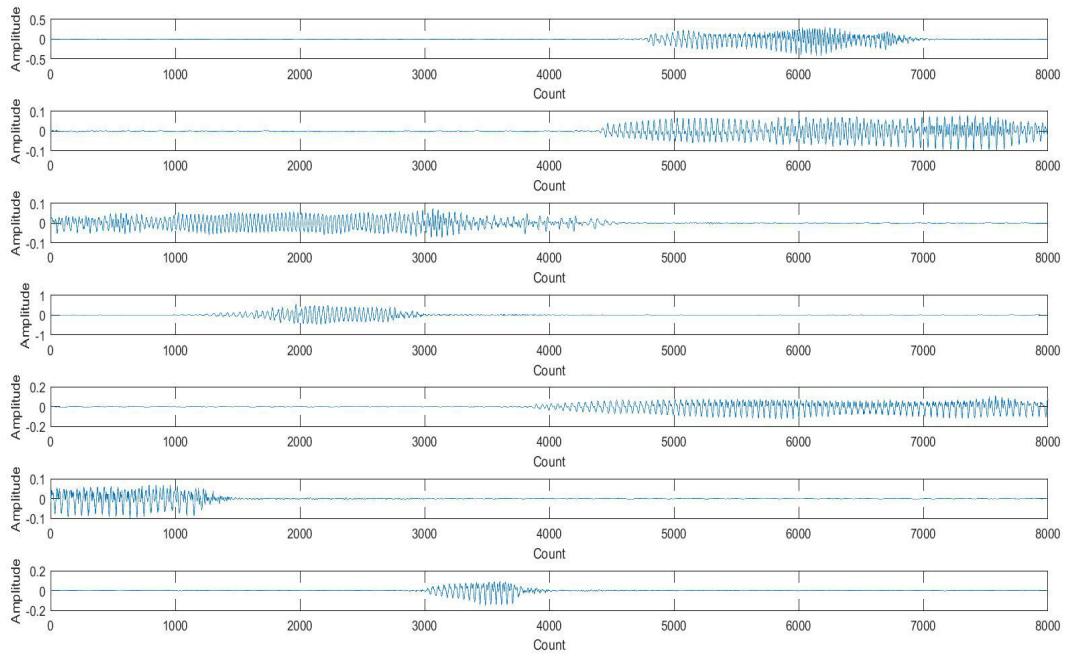


Рисунок 3.4 - Отрезки сигнала, полученные в результате фильтрации

Этап 3. Выделение информативных участков

Как было показано в разделе 1, основной тон проявляется на вокализованных участках речи. Но на произвольно выбранном отрезке речевого сигнала, даже значительной длительности, могут присутствовать как невокализованные звуки, так и паузы. Эти части звука вносят определенные погрешности, поэтому их целесообразно исключить из анализа.

Невокализованные звуки и паузы отличаются от вокализованных звуков в том числе и тем, что имеют значительно меньший размах, поэтому для выделения информативных вокализованных участков можно воспользоваться наиболее простой реализацией технологии VAD. То есть, с помощью энергетического порога, выделяются участки, его превышающие, а остальные отсчеты сигнала приравниваются к нулю.

В работе был эмпирическим путем выбран порог $H = 0.9 * X_{\max}$, где X_{\max} – максимальное значение амплитуды на отрезке сигнала.

На рисунке 3.5 показан результат работы алгоритма выделения

Изм.	Лист	№ докум.	Подпись	Дата	Лист 21
					11070006.11.03.02.369.ПЗВКР

вокализованных звуков на отрезках длительностью N .

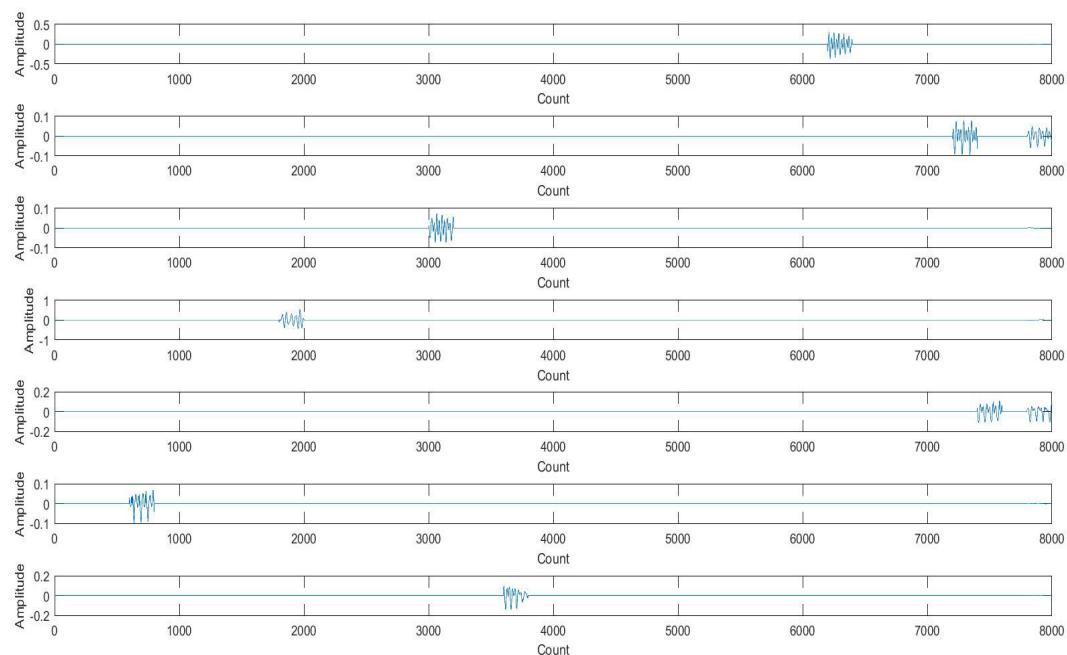


Рисунок 3.5 - Выделение вокализованных звуков

Далее, участки, соответствующие вокализованным звукам, записываются в отдельный массив для дальнейшего анализа. Было принято, что длительность $N_m = 256$ отсчётов достаточна. Пример полученных звуков показан на рисунке 3.6.

Изм.	Лист	№ докум.	Подпись	Дата
------	------	----------	---------	------

11070006.11.03.02.369.ПЗВКР

Лист
22

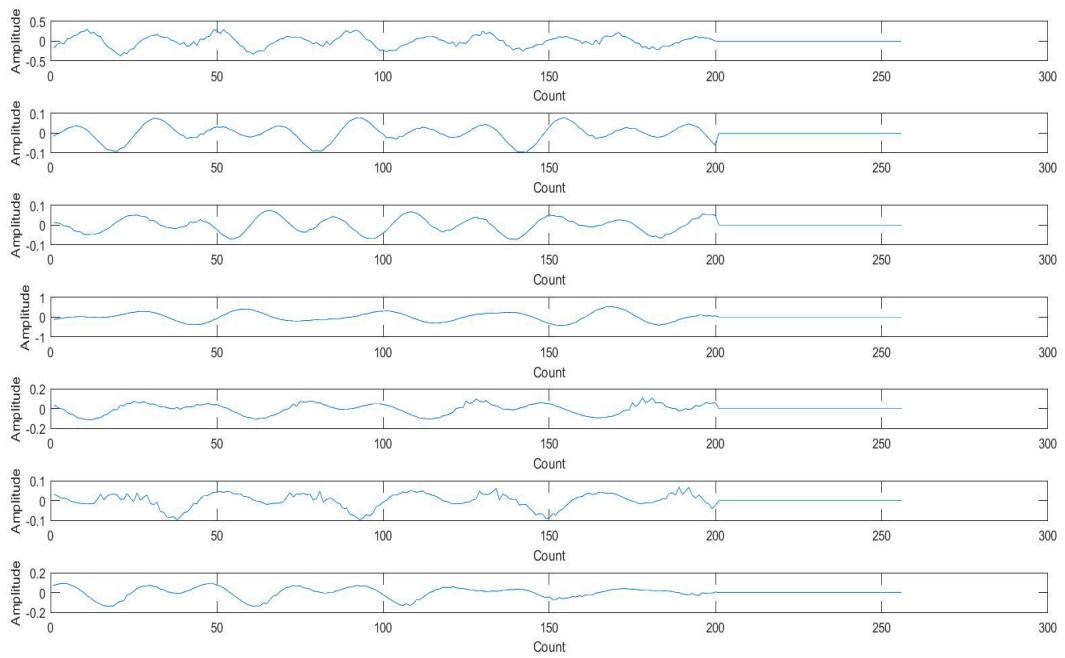


Рисунок 3.6 - Выделение отрезка размером 256 отсчётов

Для удобства сравнения и построения графиков, сигналы нормируются к единице (рисунок 3.7).

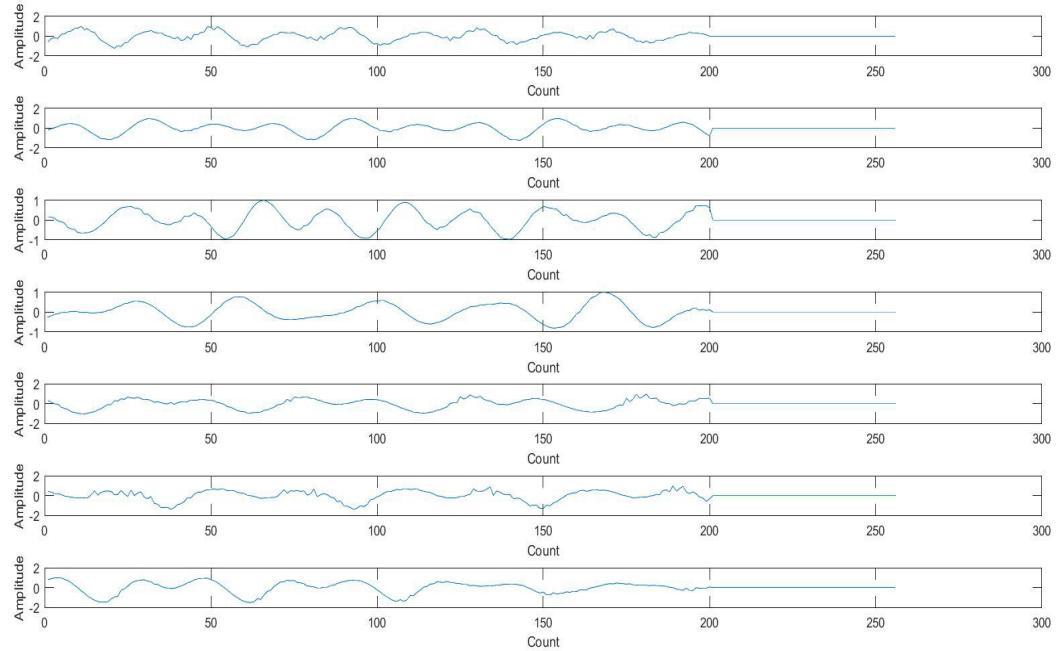


Рисунок 3.7 - Нормирование сигнала

Изм.	Лист	№ докум.	Подпись	Дата
------	------	----------	---------	------

11070006.11.03.02.369.ПЗВКР

Лист
23

Полученные данные могут быть использованы для реализации амплитудного метода оценивания периода основного тона.

Этап 4. Спектральный анализ

Для реализации спектрального метода оценивания основного тона, необходимо провести частотный анализ, который осуществляется методом ДПФ.

На рисунке 3.8 показан полученный результат.

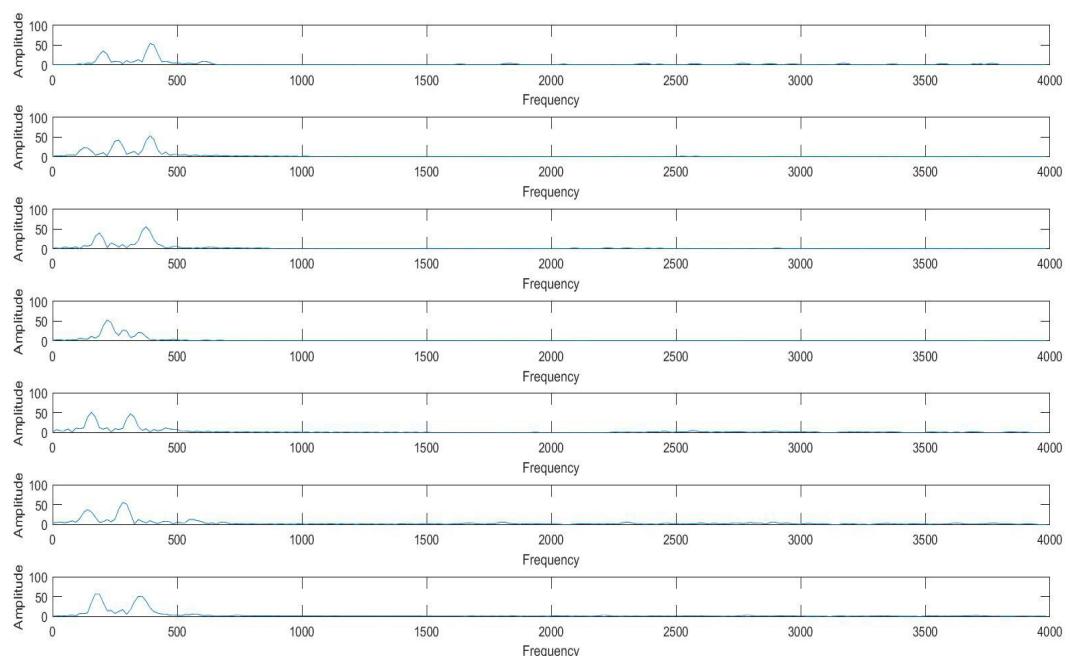


Рисунок 3.8 - Спектральный анализ

В результате выполнения всех этапов, для дальнейшего исследования получены ограниченный по частоте спектр информативных участков речевых сигналов.

3.2 Исследование динамического изменения формант речевого сигнала

В данном параграфе методика исследования показана на примере

Изм.	Лист	№ докум.	Подпись	Дата	Лист 24
					11070006.11.03.02.369.ПЗВКР

оценивания основного тона. Аналогичным образом можно исследовать поведение формант в любом выбранном диапазоне.

В работе используются два метода оценивания основного тона - амплитудный и спектральный.

Амплитудный метод основан на свойстве квазистационарности речевых сигналов, т.е. на стационарном участке вокализованного звука при малом уровне шумов форма речевого колебания почти точно повторяется на каждом очередном периоде основного тона. Таким образом, период основного тона определяется как расстояние между глобальными максимумами (максимумами максиморумами) речевого сигнала (рисунок 3.9).

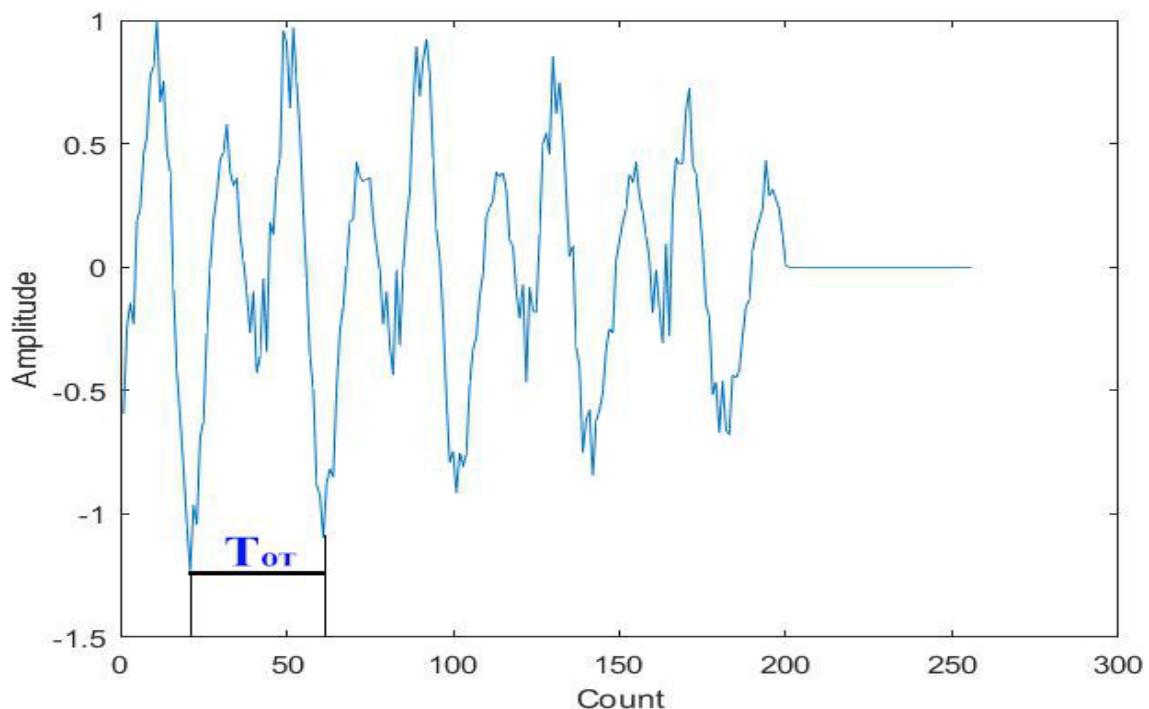


Рисунок 3. 9 - Определение периода основного тона амплитудным методом

Частота основного тона, определяется, как величина обратная периоду основного тона $T_{от}$.

Спектральный метод основан на том, что при вокализованном возбуждении речевого тракта в спектре сигнала присутствуют пики на частотах, кратных частоте основного тона. Если построить дискретное преобразование

Изм.	Лист	№ докум.	Подпись	Дата
------	------	----------	---------	------

11070006.11.03.02.369.ПЗВКР

Лист
25

Фурье с достаточно малым шагом дискретизации по частоте, то можно, в качестве оценки частоты основного тона использовать наименьшую частоту, соответствующую одному из максимальных значений энергии спектра. Поиск максимума осуществляется в интервале 20 - 400 Гц (рисунок 3.10).

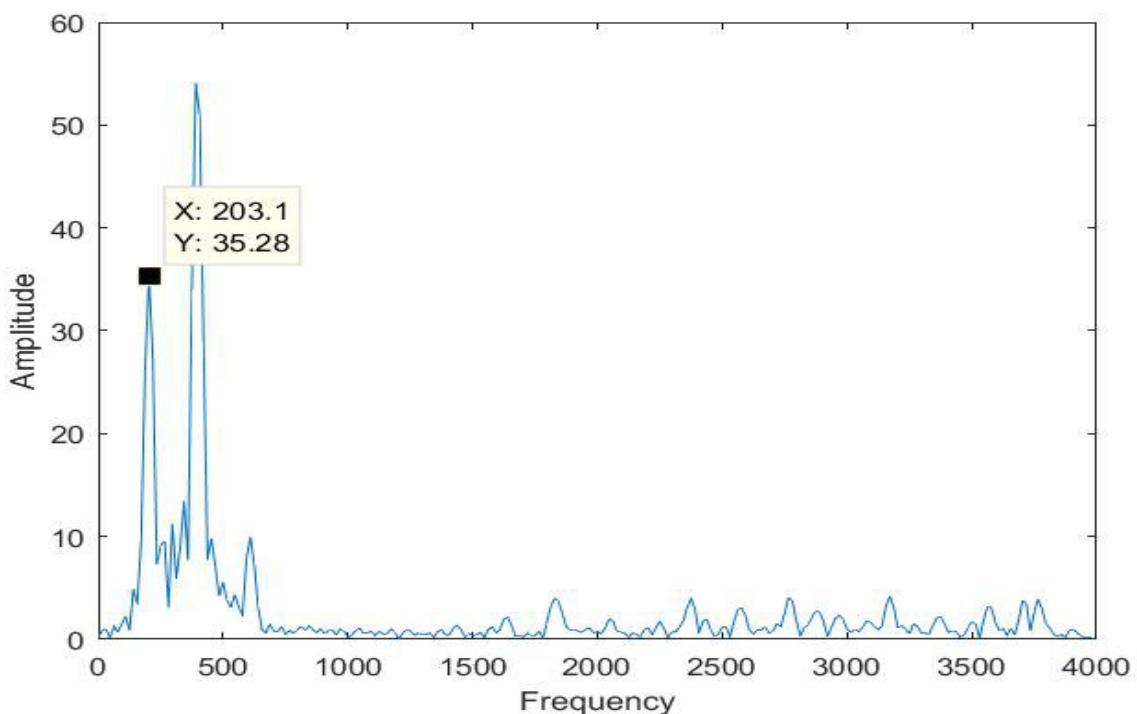


Рисунок 3.10 - Спектральный метод

На основе данных методов были проведены исследования речевого сигнала, показанного в качестве примера в п.3.1 и получены результаты отображенные в таблице 3.1 и на рисунках 3.11 - 3.17.

Таблица 3.1 - Результаты оценивания основного тона речевого сигнала 1 “нет”

Номер отрезка речевого сигнала	$F_{\text{от}}$, Гц	
	амплитудный метод	спектральный метод
1	2	3
1	200	203
2	127	125

Окончание таблицы 3.1

<i>I</i>	<i>2</i>	<i>3</i>
3	190.5	187.5
4	190	219
5	154	156
6	145.5	140.5
7	186	172

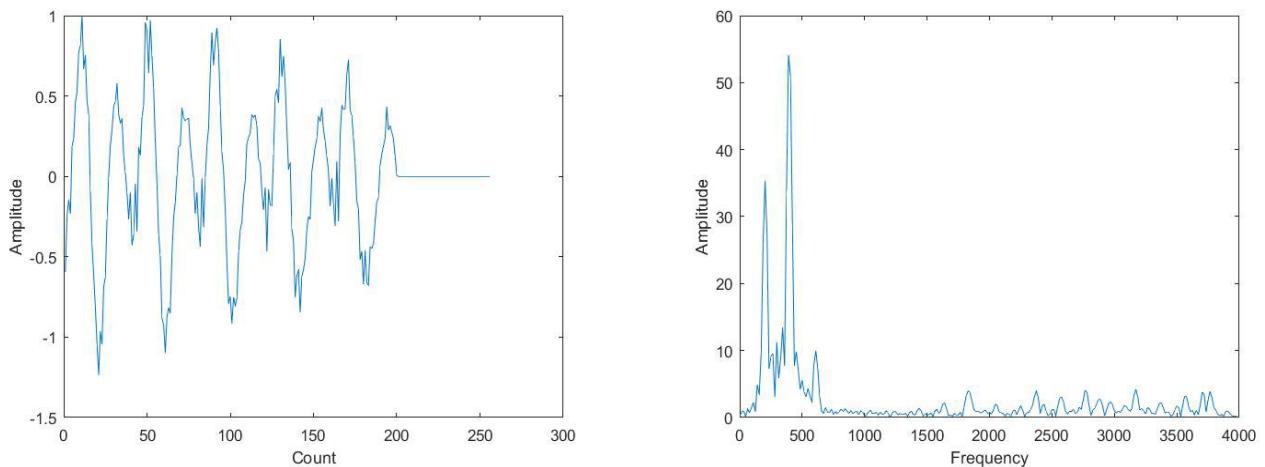


Рисунок 3.11 - Оценивание основного тона отрезка 1 речевого сигнала 1 “нет”

а) амплитудным методом, б) спектральным методом

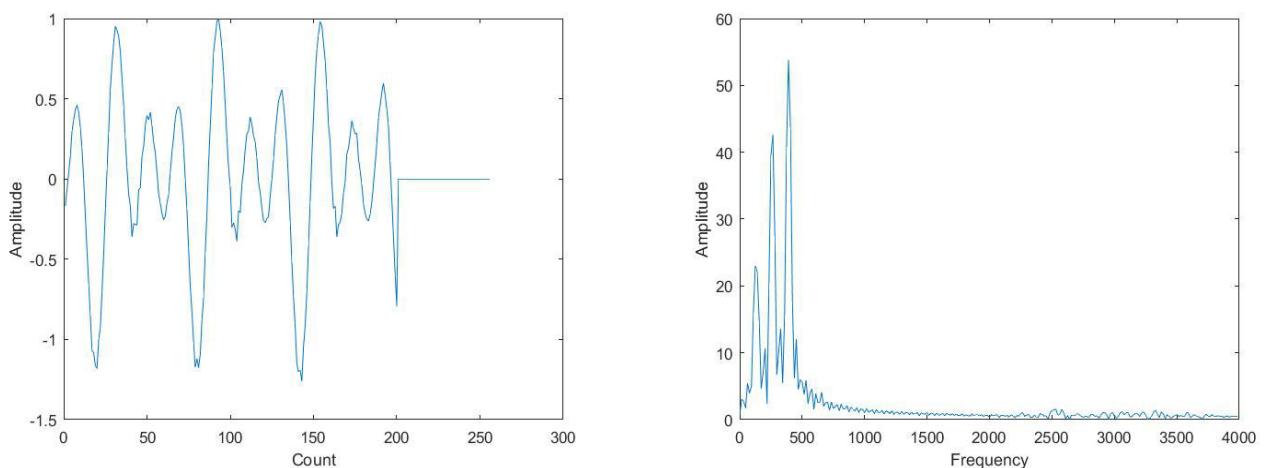


Рисунок 3.12 - Оценивание основного тона отрезка 2 речевого сигнала 1 “нет”

а) амплитудным методом, б) спектральным методом

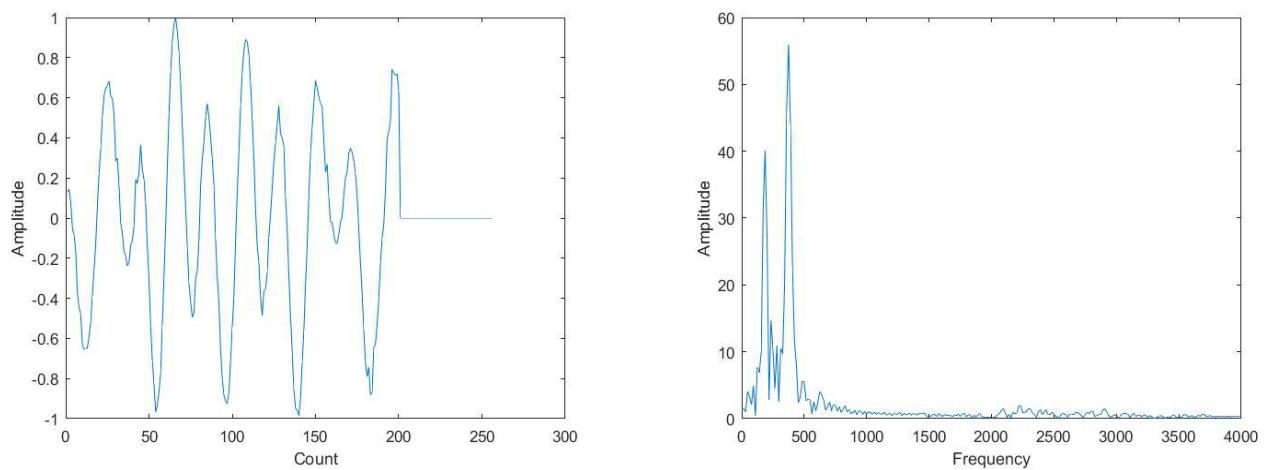


Рисунок 3.13 - Оценивание основного тона отрезка 3 речевого сигнала 1 “нет”
а) амплитудным методом, б) спектральным методом

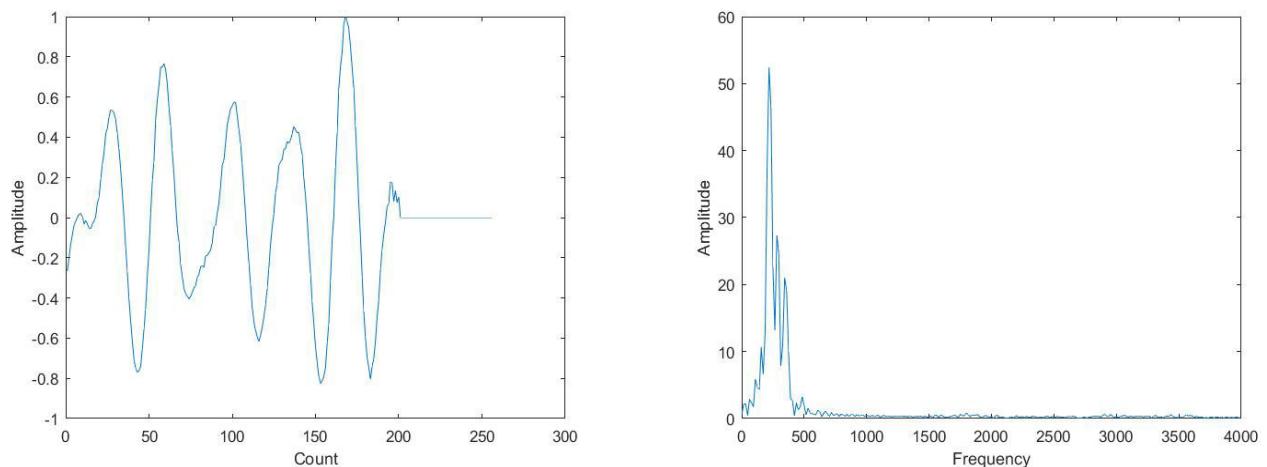


Рисунок 3.14 - Оценивание основного тона отрезка 4 речевого сигнала 1 “нет”
а) амплитудным методом, б) спектральным методом

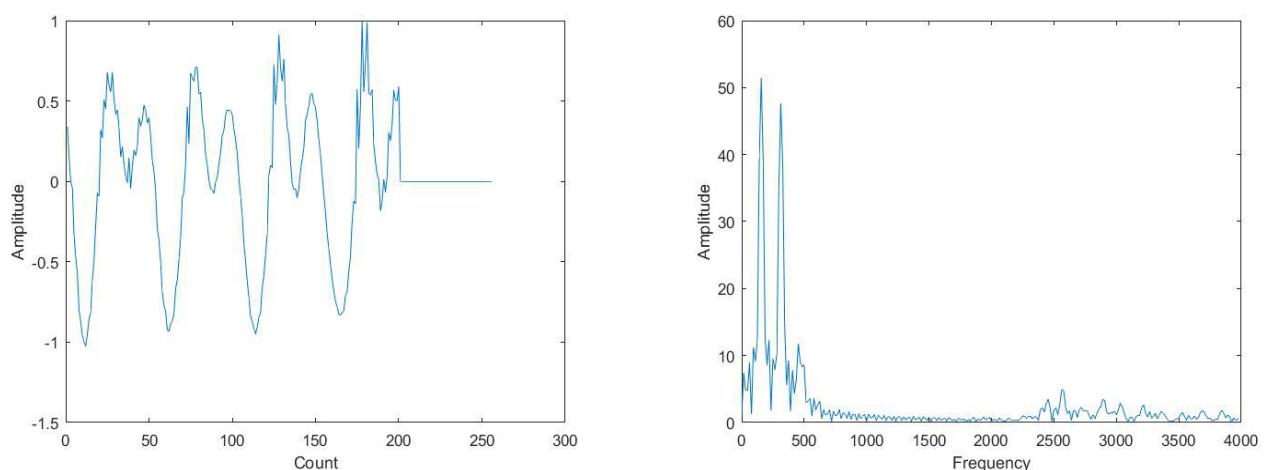


Рисунок 3.15 - Оценивание основного тона отрезка 5 речевого сигнала 1 “нет”
а) амплитудным методом, б) спектральным методом

Изм.	Лист	№ докум.	Подпись	Дата	Лист 28
					11070006.11.03.02.369.ПЗВКР

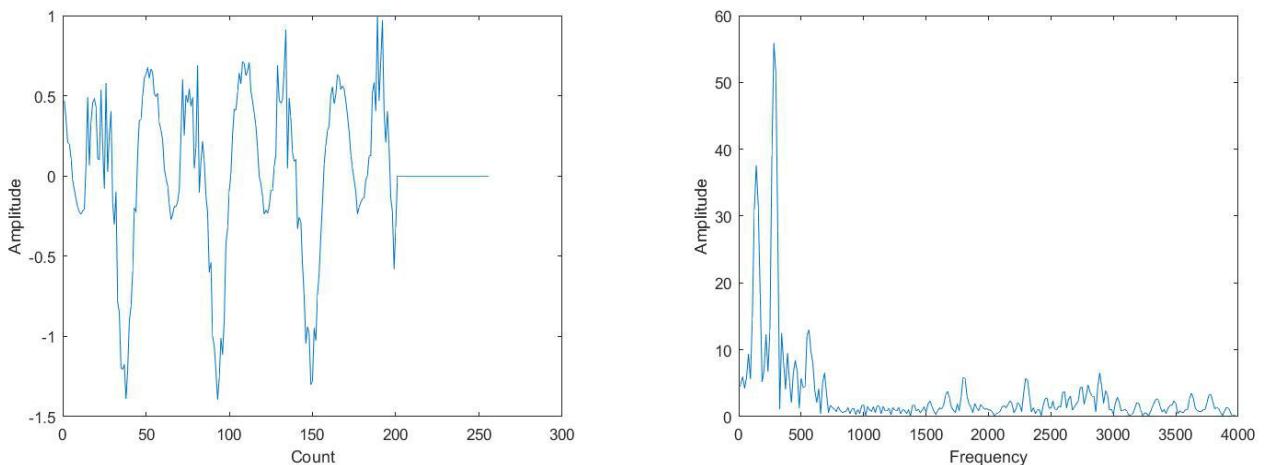


Рисунок 3.16 - Оценивание основного тона отрезка 6 речевого сигнала 1 “нет”
а) амплитудным методом, б) спектральным методом

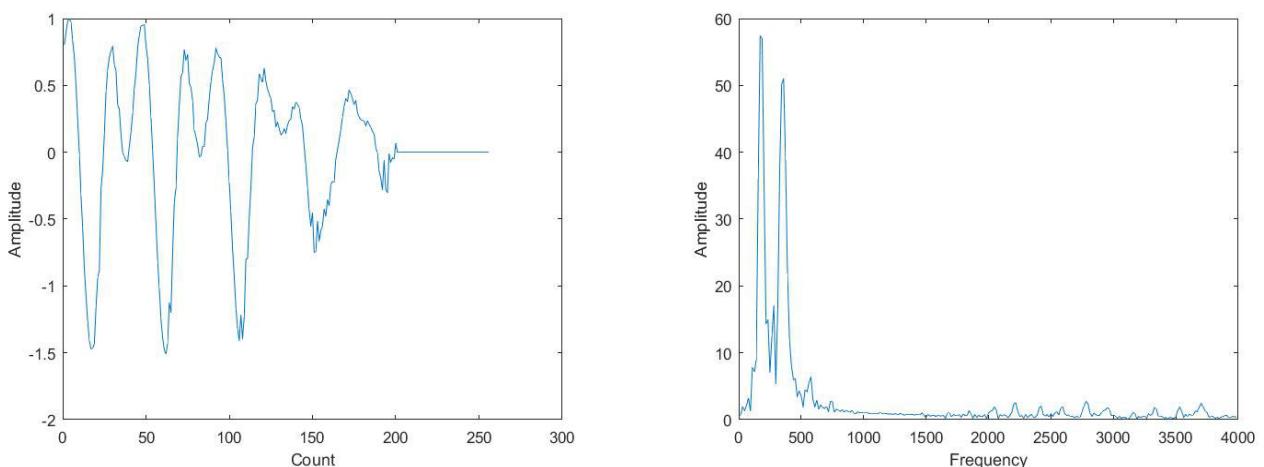


Рисунок 3.17 - Оценивание основного тона отрезка 7 речевого сигнала 1 “нет”
а) амплитудным методом, б) спектральным методом

Аналогичные исследования были проведены для второго сигнала “без высшего образования” длиной $L = 15798$ отсчётов. Сигнал “без высшего образования” был разделен на $k = 7$ отрезков длительностью $N = 2000$ отсчётов. Результаты представлены в таблице 3.2 и на рисунках 3.18-3.24.

Изм.	Лист	№ докум.	Подпись	Дата
------	------	----------	---------	------

11070006.11.03.02.369.ПЗВКР

Лист
29

Таблица 3.2 - Результаты оценивания основного тона речевого сигнала 2 “без высшего образования”

Номер отрезка речевого сигнала	$F_{\text{от}}$, Гц	
	амплитудный метод	спектральный метод
1	182	187.5
2	186	187.5
3	242.5	250
4	216	219
5	163	156
6	174	172
7	157	156.5

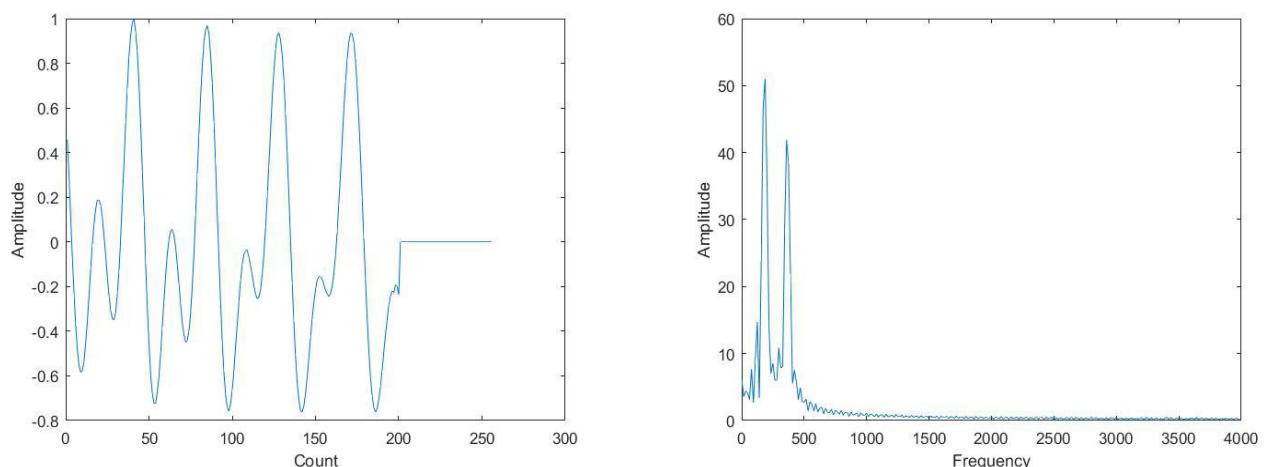


Рисунок 3.18 - Оценивание основного тона отрезка 1 речевого сигнала 2 “без высшего образования” а) амплитудным методом, б) спектральным методом

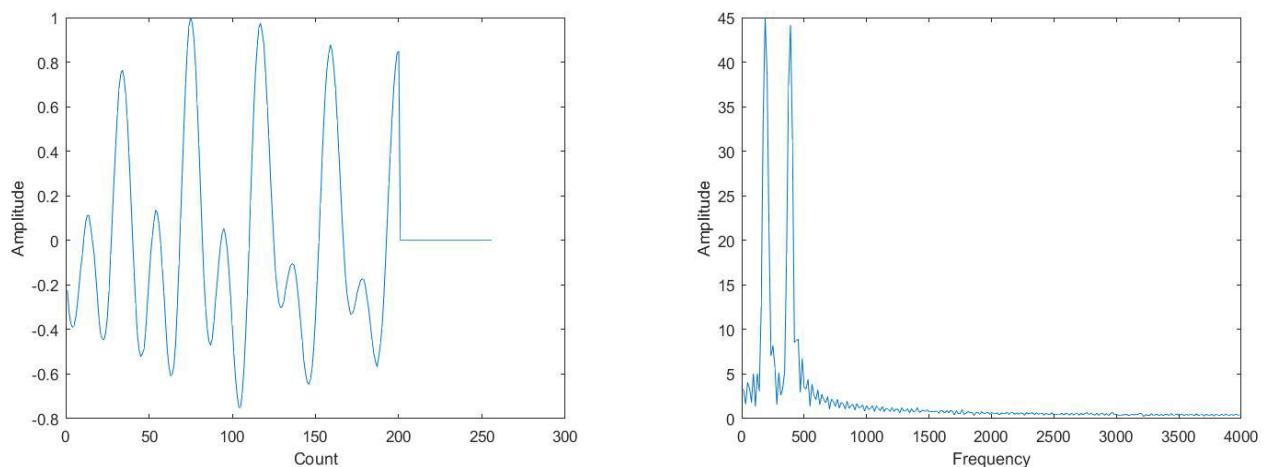


Рисунок 3.19 - Оценивание основного тона отрезка 2 речевого сигнала 2 “без высшего образования” а) амплитудным методом, б) спектральным методом

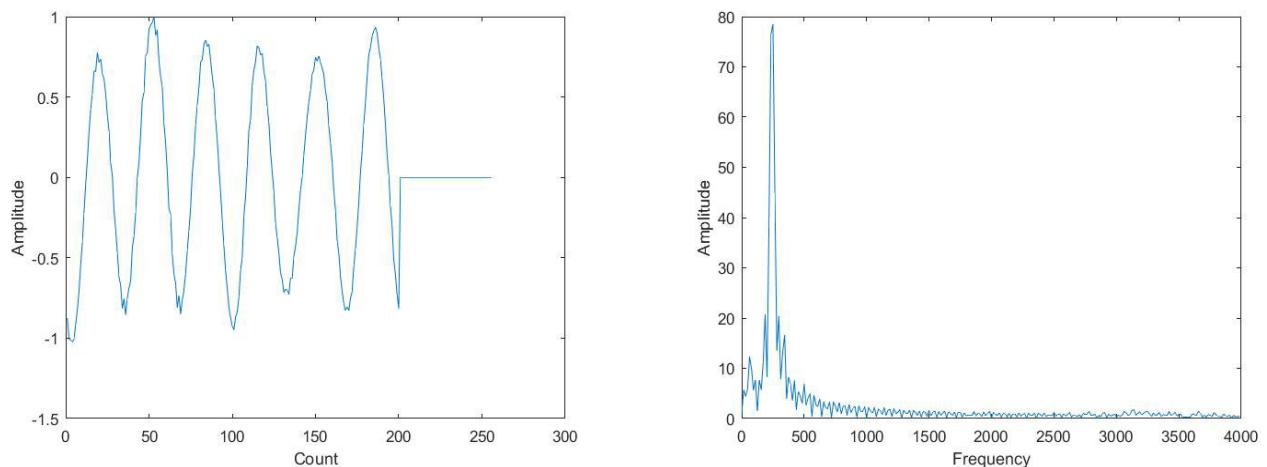


Рисунок 3.20 - Оценивание основного тона отрезка 3 речевого сигнала 2 “без высшего образования” а) амплитудным методом, б) спектральным методом

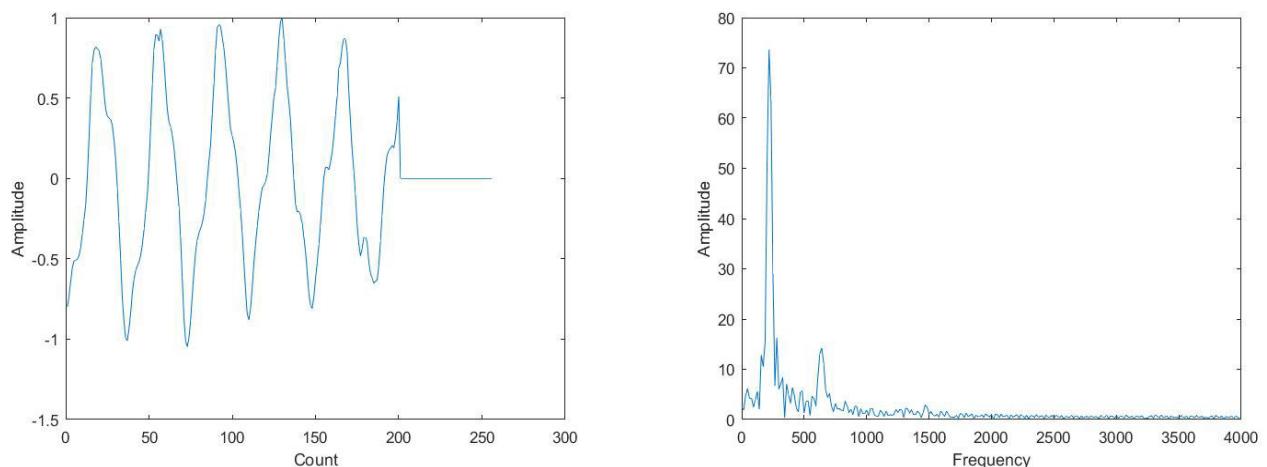


Рисунок 3.21 - Оценивание основного тона отрезка 4 речевого сигнала 2 “без высшего образования” а) амплитудным методом, б) спектральным методом

Изм.	Лист	№ докум.	Подпись	Дата	Лист
					31

11070006.11.03.02.369.ПЗВКР

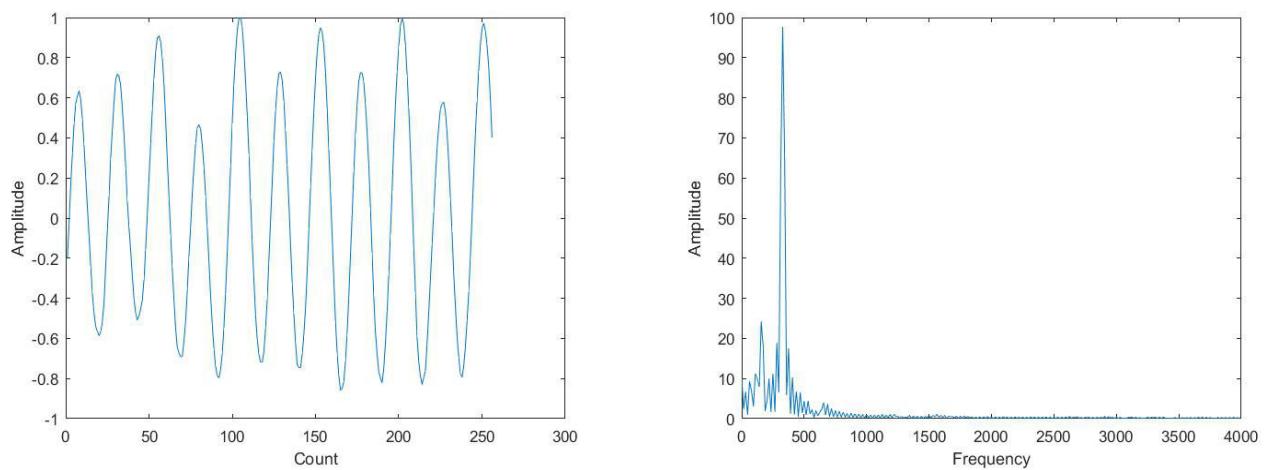


Рисунок 3.22 - Оценивание основного тона отрезка 5 речевого сигнала 2 “без высшего образования” а) амплитудным методом, б) спектральным методом

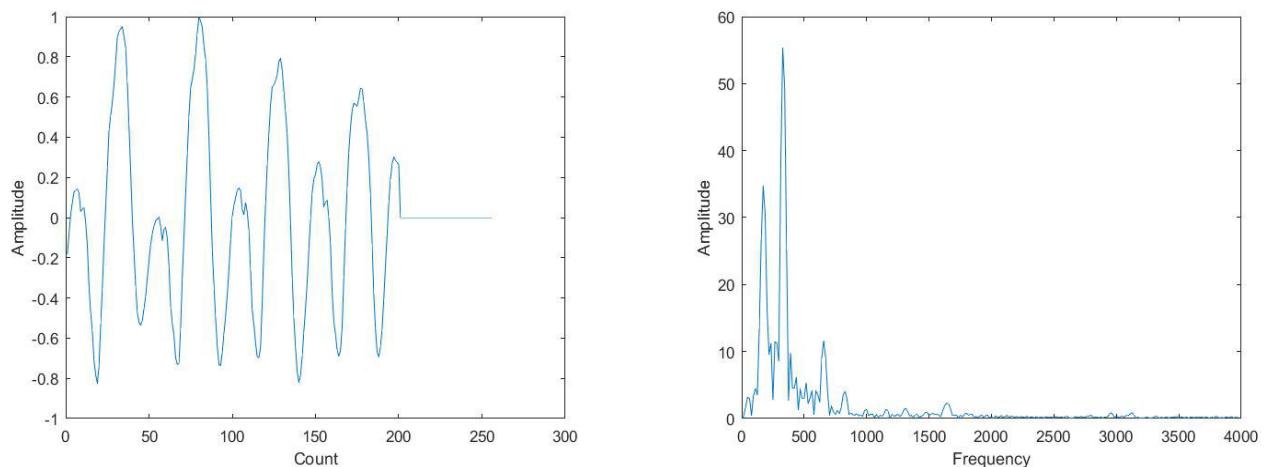


Рисунок 3.23 - Оценивание основного тона отрезка 6 речевого сигнала 2 “без высшего образования” а) амплитудным методом, б) спектральным методом

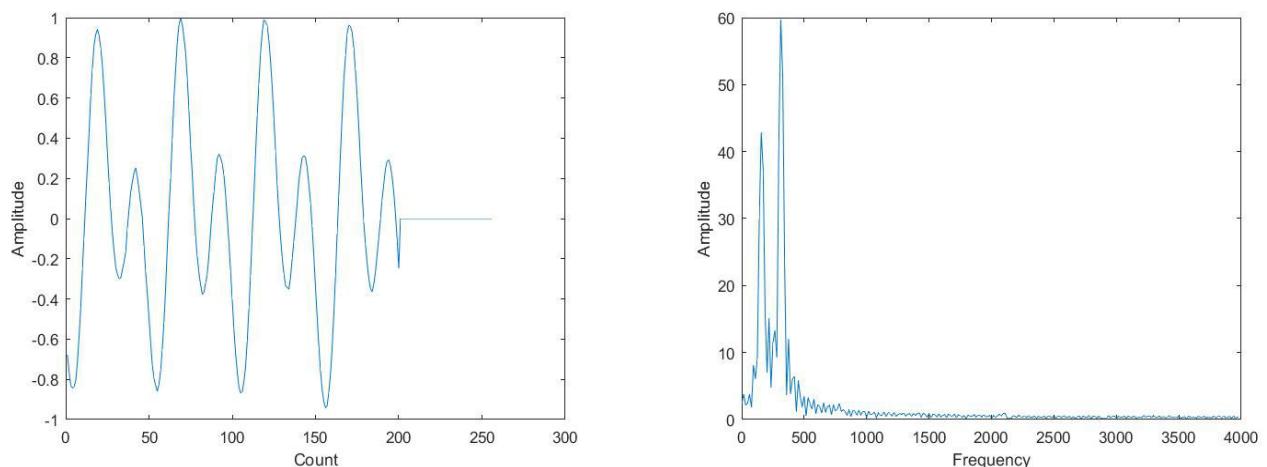


Рисунок 3.24 - Оценивание основного тона отрезка 7 речевого сигнала 2 “без высшего образования” а) амплитудным методом, б) спектральным методом

Изм.	Лист	№ докум.	Подпись	Дата	Лист
					11070006.11.03.02.369.ПЗВКР 32

Исследования показали достаточно близкое сходство этих двух методов, оба из которых не свободны от недостатков, основным из которых является сложность подавления неинформативных локальных максимумов.

Полученные результаты позволяют построить динамическую кривую изменения основного тона речевых сигналов (рисунки 3.25, 3.26).

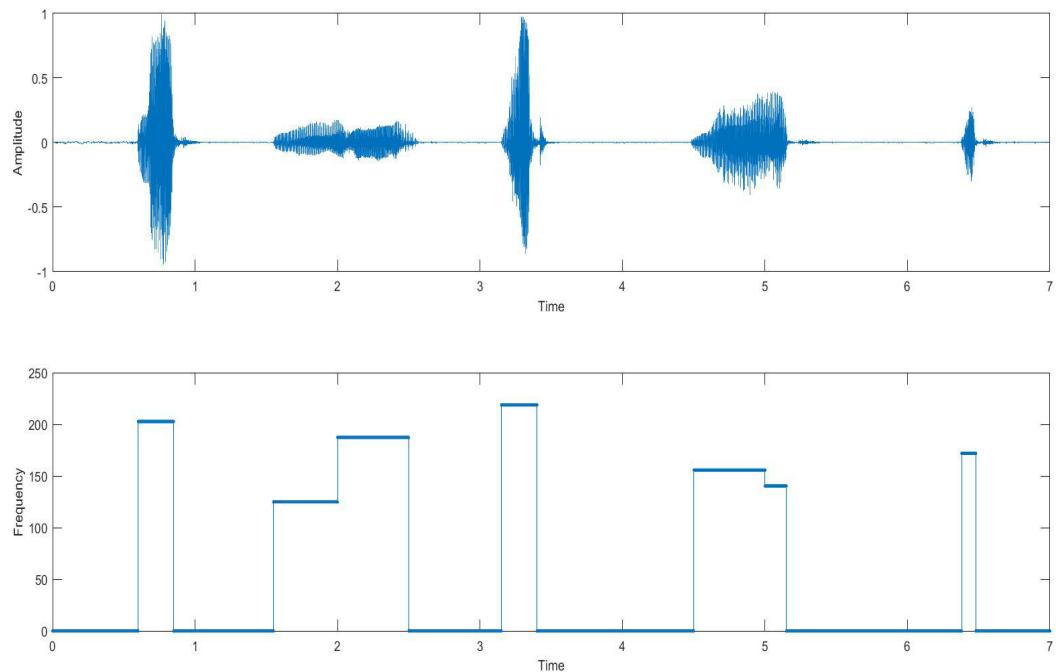


Рисунок 3.25 - Речевой сигнал 1 “нет” (а) и динамическое изменение частоты основного тона диктора (б)

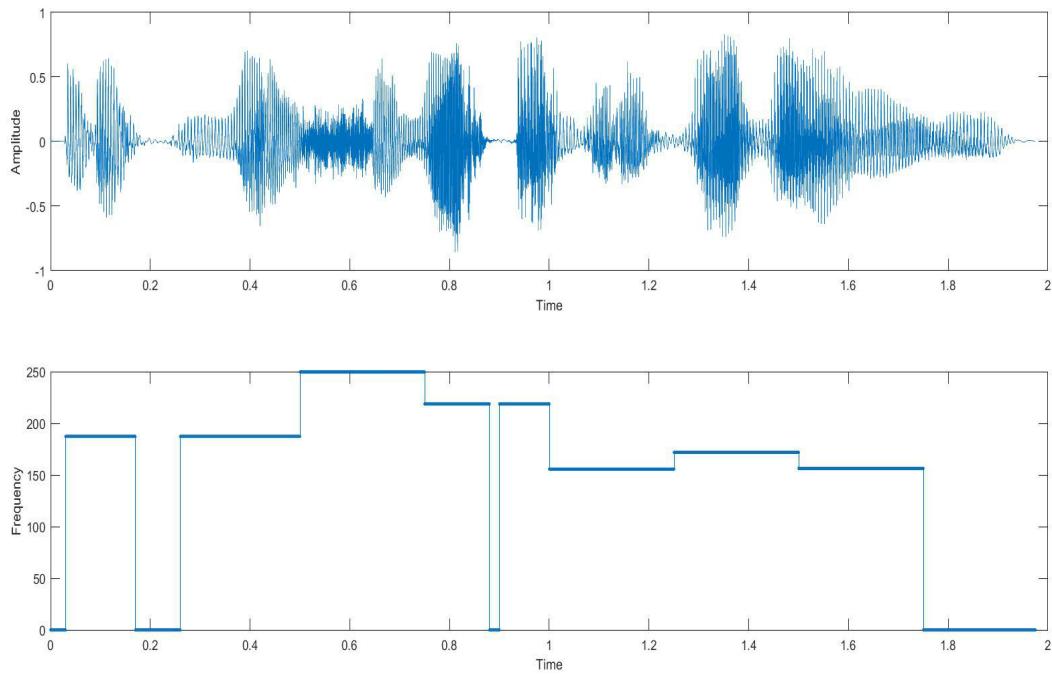


Рисунок 3.26 - Речевой сигнал 2 “без высшего образования” (а) и динамическое изменение частоты основного тона диктора (б)

Исследования показывают, что основного тона первого диктора (речевой сигнал 1, рисунок 3.25) изменяется в диапазоне 130 - 200 Гц, основного тона второго диктора (речевой сигнал 2, рисунок 3.26) - в диапазоне 160 - 250 Гц.

Такая вариативность обусловлена естественным произнесением и обеспечивает изменение интонации речи.

3.3 Алгоритм динамического изменения формант речевого сигнала

В данном разделе представлен метод автоматического оценивания динамического изменения формант речевого сигнала, на примере основного тона речевых сигналов.

Алгоритм состоит из нескольких этапов (рисунок 3.27).

Изм.	Лист	№ докум.	Подпись	Дата	Лист
					34

11070006.11.03.02.369.ПЗВКР



Рисунок 3.27 - Алгоритм динамического изменения формант речевого сигнала

Этап 1. Спектральный анализ

Как было показано в разделе 3.1, для реализации спектрального метода оценивания основного тона, необходимо провести частотный анализ, который осуществляется методом ДПФ.

На рисунке 3.28 показан полученный результат.

Изм.	Лист	№ докум.	Подпись	Дата

11070006.11.03.02.369.ПЗВКР

Лист
35

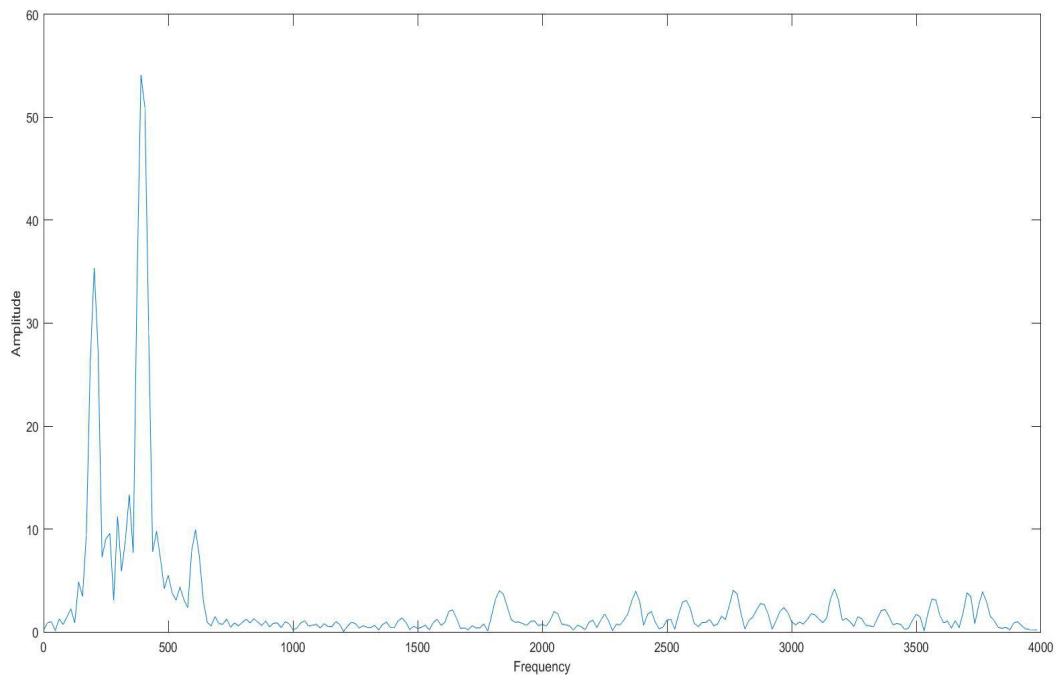


Рисунок 3.28 - Спектральный анализ отрезка 1 речевого сигнала 1 “нет”

Анализ показал, что в заданном частотном диапазоне присутствует несколько формант (в данном случае три) и определенное количество локальных максимумов, которые можно считать ложными для данной задачи.

Этап 2. Сортировка значений амплитуд по убыванию

Для селекции формант и ложных максимумов используется процедура сортировки значений амплитуд по убыванию. При этом сохраняются их параметры (частота, амплитуда, фаза).

На рисунке 3.29 показан полученный результат.

Изм.	Лист	№ докум.	Подпись	Дата
------	------	----------	---------	------

11070006.11.03.02.369.ПЗВКР

Лист
36

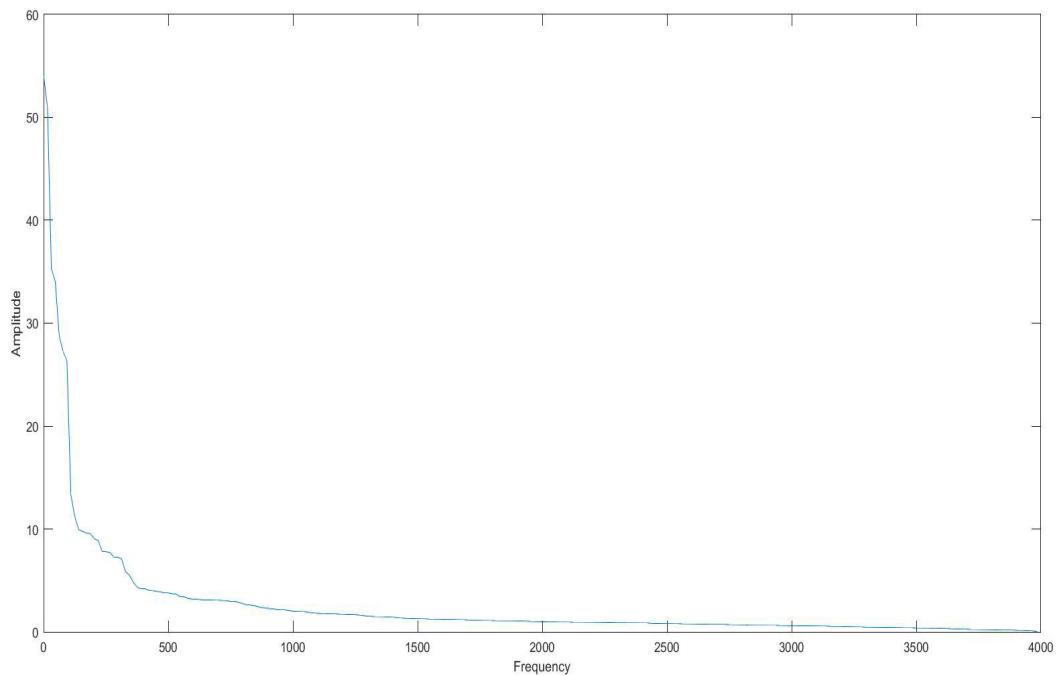


Рисунок 3.29 - Сортировка значений амплитуд по убыванию

Этап 3. Выбор значений максимальных амплитуд

Количество значений максимумов зависит от ширины исследуемого диапазона. В примере, показанном на рисунке 3.28, значений должно быть не менее трех. На рисунке 3.30 показаны значения семи максимумов.

54.0926	50.9634	35.2772	34.0031	28.7999	27.3059	26.3560
53.7813	43.6983	42.5580	40.7793	39.5445	27.5029	23.0265
55.9853	46.0030	43.9481	40.0983	31.3845	26.6593	22.8904
52.3875	46.3271	37.4537	27.3329	24.5735	22.5355	21.0372
51.4883	47.6375	39.4279	39.0410	38.6725	34.6712	14.7241
55.8529	51.5394	38.9382	37.6230	31.5574	30.9240	27.9666
57.4488	56.9365	50.9915	50.2020	38.8517	37.2096	36.0649

Рисунок 3.30 - Выбор значений максимальных амплитуд

Этап 4. Поиск значения амплитуды с минимальной частотой

Из значений максимальных амплитуд, выбирается значение амплитуды с минимальной частотой, в данном случае - это частота основного тона.

Результат исследования речевого сигнала “нет” отображен в таблице 3.3.

Изм.	Лист	№ докум.	Подпись	Дата	Лист
					11070006.11.03.02.369.ПЗВКР 37

Таблица 3.3 - Результаты оценивания основного тона речевого сигнала 1 “нет”

Номер отрезка речевого сигнала	$F_{\text{от}}$, Гц		
	амплитудный метод	спектральный метод	
		вручную	автоматически
1	200	203	187.5
2	127	125	125
3	190.5	187.5	172
4	190	219	203
5	154	156	141
6	145.5	140.5	125
7	186	172	156

Выданное ранее значение длительности отрезков $N = 8000$, не дало достаточной точности для формирования «рисунка» динамического изменения частоты основного тона, поэтому речевой сигнал 1 “нет” длиной $L = 56000$ был разделен на $k = 140$ отрезков длительностью $N = 400$ отсчётов.

Были проведены аналогичные исследования для нескольких формант. Результаты показаны в виде динамической кривой изменения основного тона и других формант речевых сигналов на спектрограмме речевого сигнала (рисунки 3.31, 3.32, 3.33, 3.34).

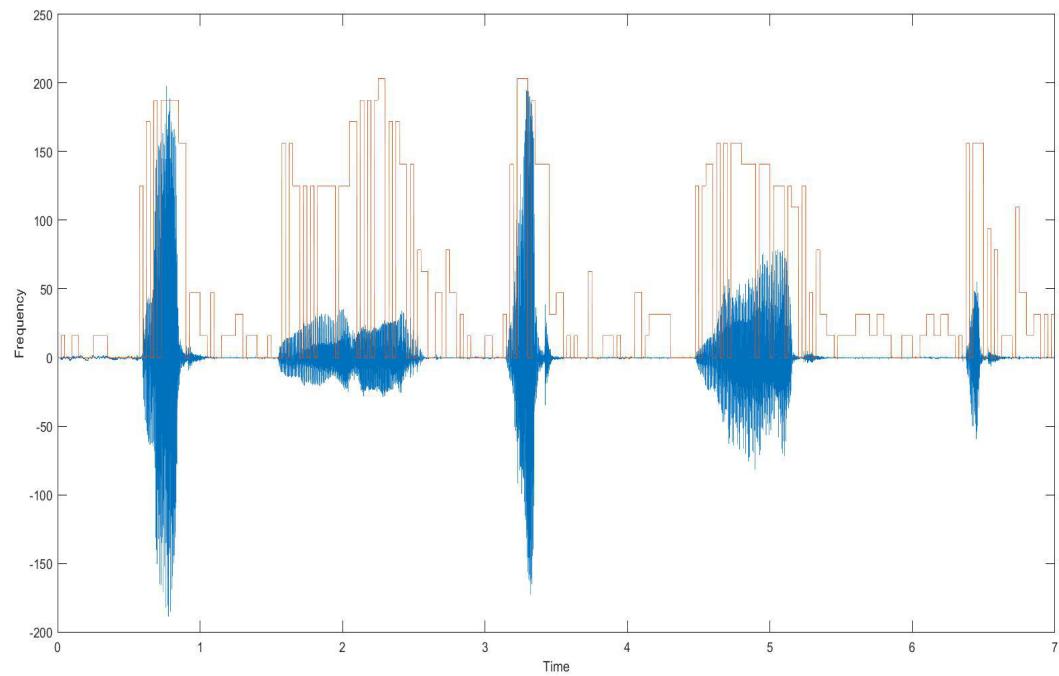


Рисунок 3.31 - Речевой сигнал 1 “нет” и динамическое изменение частоты основного тона диктора

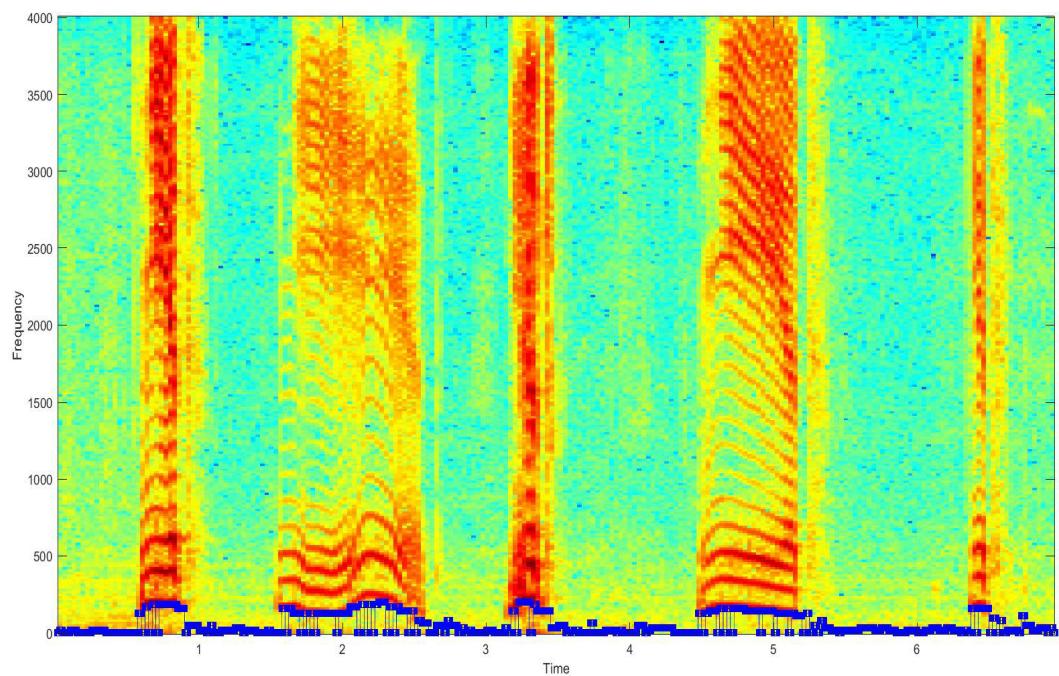


Рисунок 3.32 - Спектр речевого сигнала 1 “нет” и динамическое изменение частоты основного тона диктора (первый тон)

Изм.	Лист	№ докум.	Подпись	Дата
------	------	----------	---------	------

11070006.11.03.02.369.ПЗВКР

Лист
39

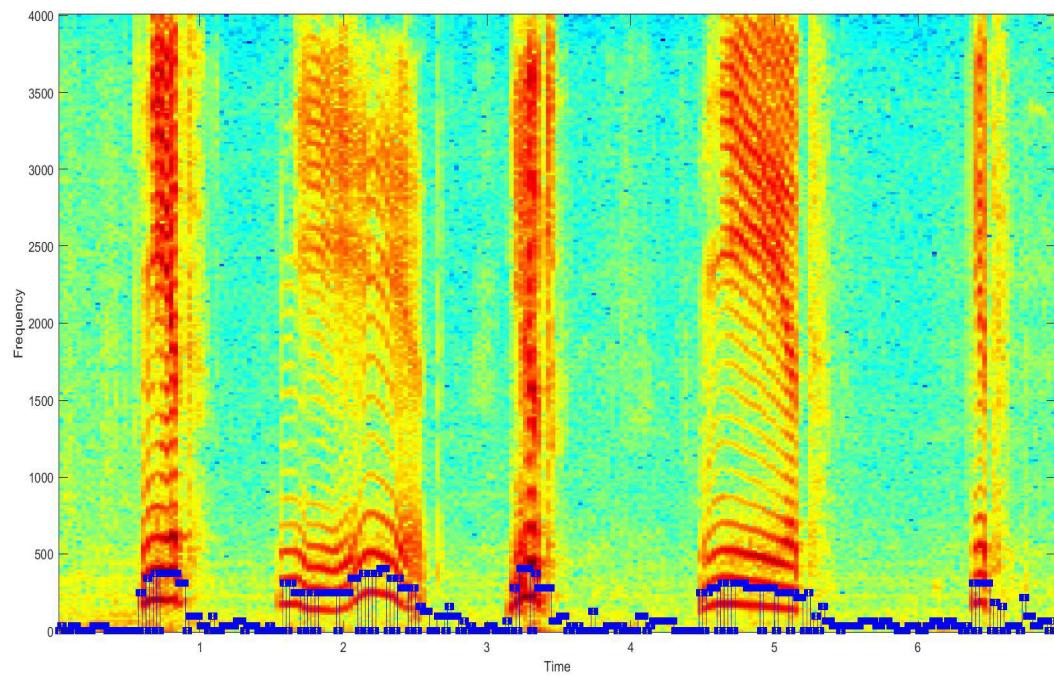


Рисунок 3.33 - Спектр речевого сигнала 1 “нет” и динамическое изменение частоты основного тона диктора (второй тон)

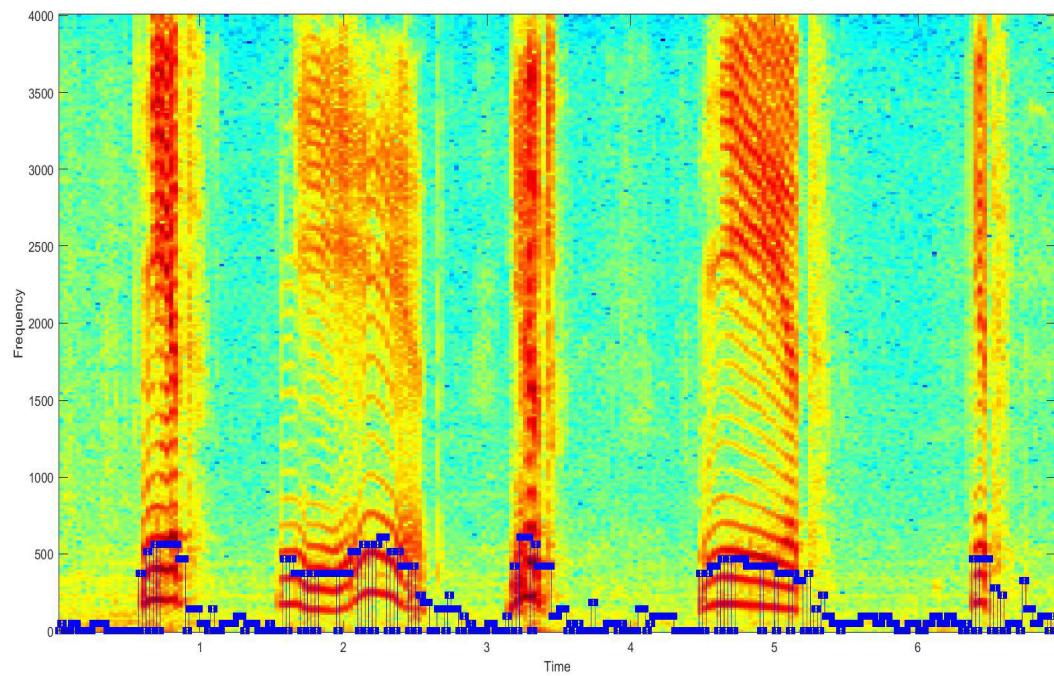


Рисунок 3.34 - Спектр речевого сигнала 1 “нет” и динамическое изменение частоты основного тона диктора (третий тон)

Изм.	Лист	№ докум.	Подпись	Дата
------	------	----------	---------	------

11070006.11.03.02.369.ПЗВКР

Лист
40

Разработанный алгоритм (как видно из рисунков) дает устойчивые результаты и позволяет создавать описание динамического изменения частотного состава звуков речи, как отдельных фонем, так и слов и даже фраз.

Современные доступные средства обработки звука, например, такой как Adobe Audition, не позволяют производить столь точный анализ. Например, функция Adobe Audition «выделить тон» дает результаты, показанные на рисунках 3.35, 3.36.

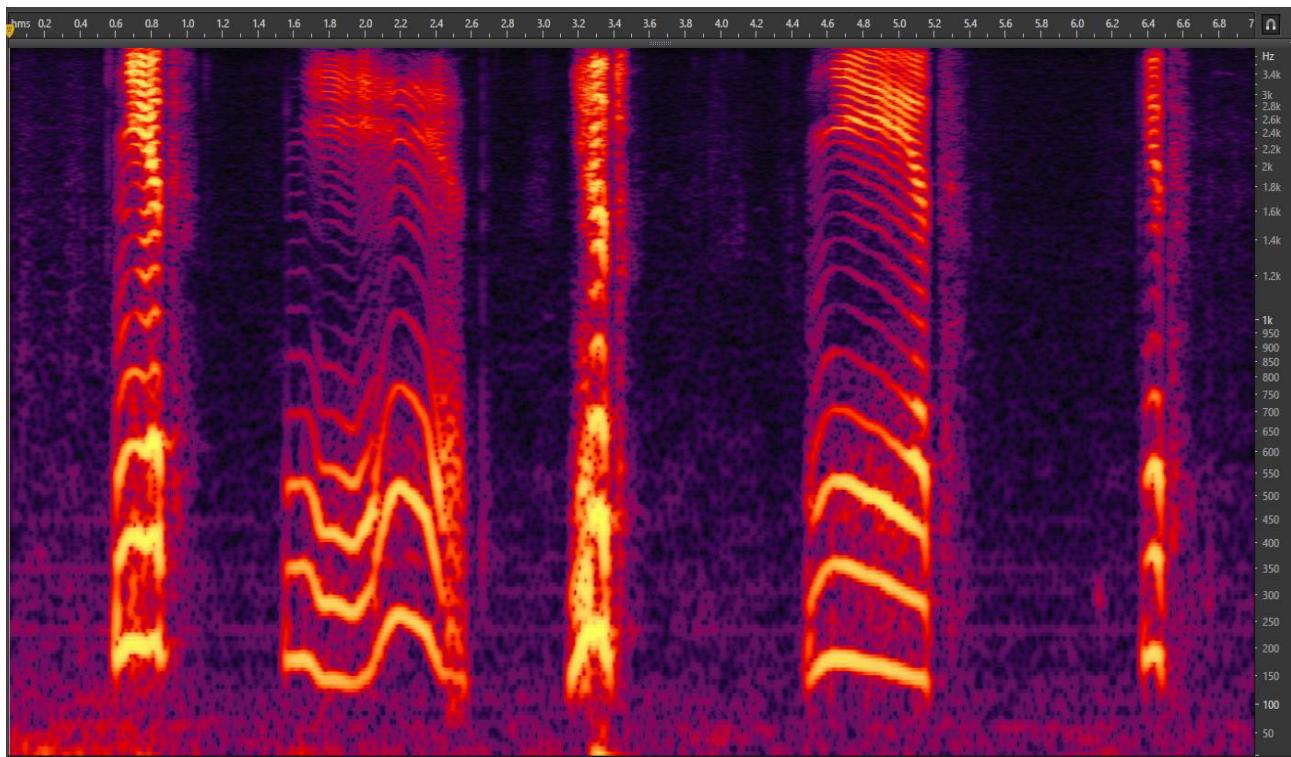


Рисунок 3.35 - Спектр речевого сигнала 1 “нет” в Adobe Audition

Изм.	Лист	№ докум.	Подпись	Дата
------	------	----------	---------	------

11070006.11.03.02.369.ПЗВКР

Лист
41

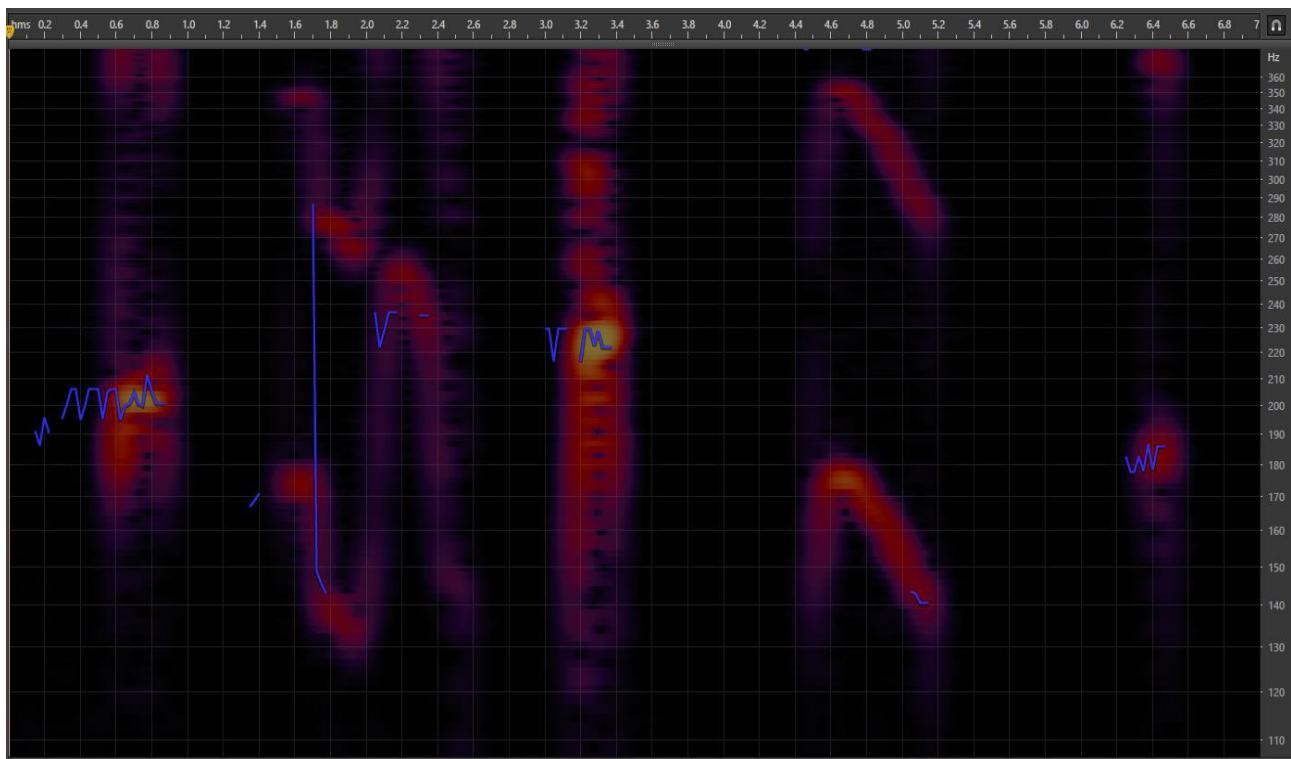


Рисунок 3.36 - Динамическое изменение частоты основного тона диктора речевого сигнала 1 “нет” в Adobe Audition

Аналогичные исследования динамики изменения частотного состава звуков речи были проведены для некоторого числа речевых сигналов.

В работе представлены:

- речевой сигнал 2 “без высшего образования” длиной $L = 15798$ отсчётов (сигнал был разделен на $k = 35$ отрезков длительностью $N = 400$ отсчётов).

Таблица 3.4 - Результаты оценивания основного тона речевого сигнала 2 “без высшего образования”

Номер отрезка речевого сигнала	$F_{\text{от}}$, Гц		
	амплитудный метод		спектральный метод
	вручную	автоматически	
I	2	3	4
1	182	187.5	156

Окончание таблицы 3.4

<i>1</i>	<i>2</i>	<i>3</i>	<i>4</i>
2	186	187.5	172
3	242.5	250	187.5
4	216	219	187.5
5	163	156	156
6	174	172	156
7	157	156.5	141

- речевой сигнал 3 “я тебя люблю” длиной $L = 8000$ отсчётов (сигнал был разделен на $k = 20$ отрезков длительностью $N = 400$ отсчётов).

- речевой сигнал 4 “ма” длиной $L = 48000$ отсчётов (сигнал был разделен на $k = 120$ отрезков длительностью $N = 400$ отсчётов).

Результаты исследований и сравнение с работой системы Adobe Audition представлены на рисунках 3.37-3.54.

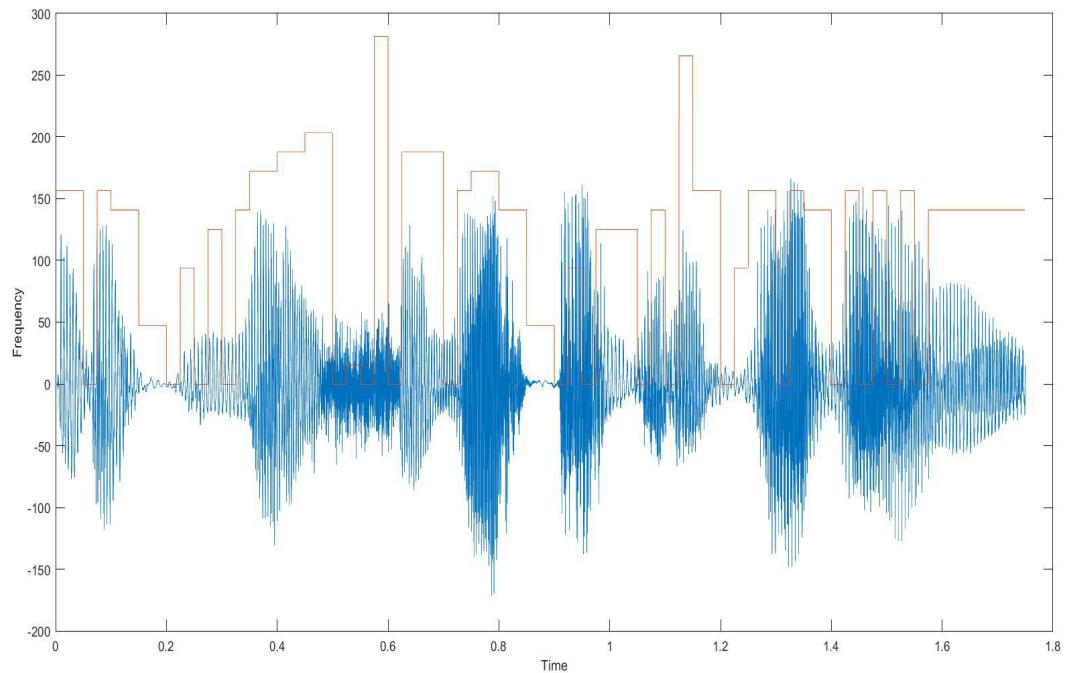


Рисунок 3.37 - Речевой сигнал 2 “без высшего образования” и динамическое изменение частоты основного тона диктора

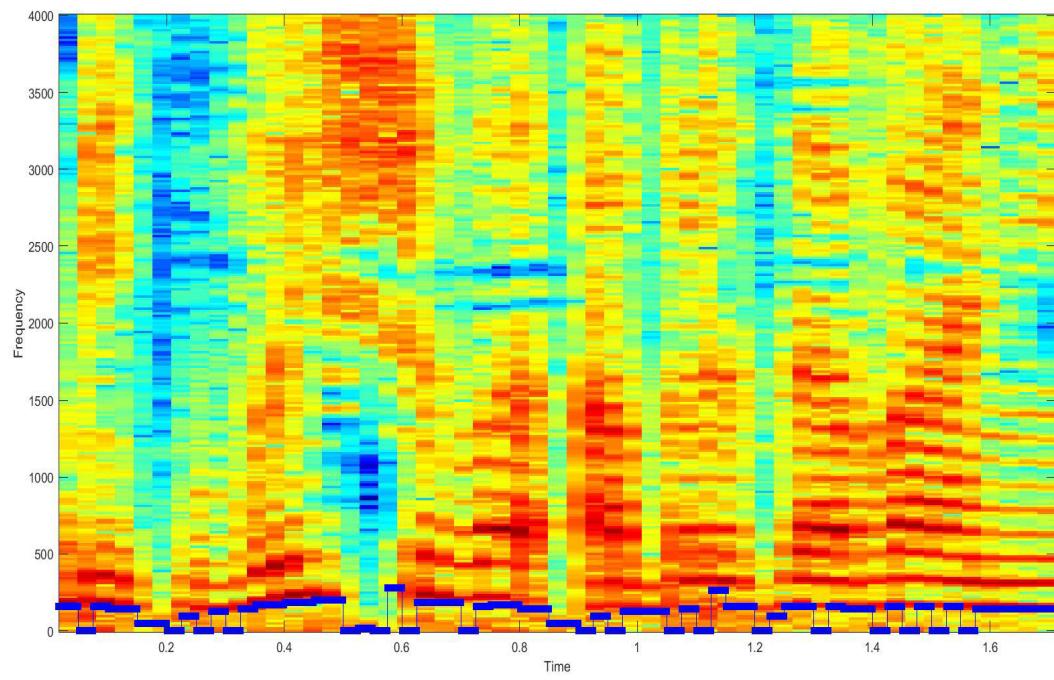


Рисунок 3.38 - Спектр речевого сигнала 2 “без высшего образования” и динамическое изменение частоты основного тона диктора (первый тон)

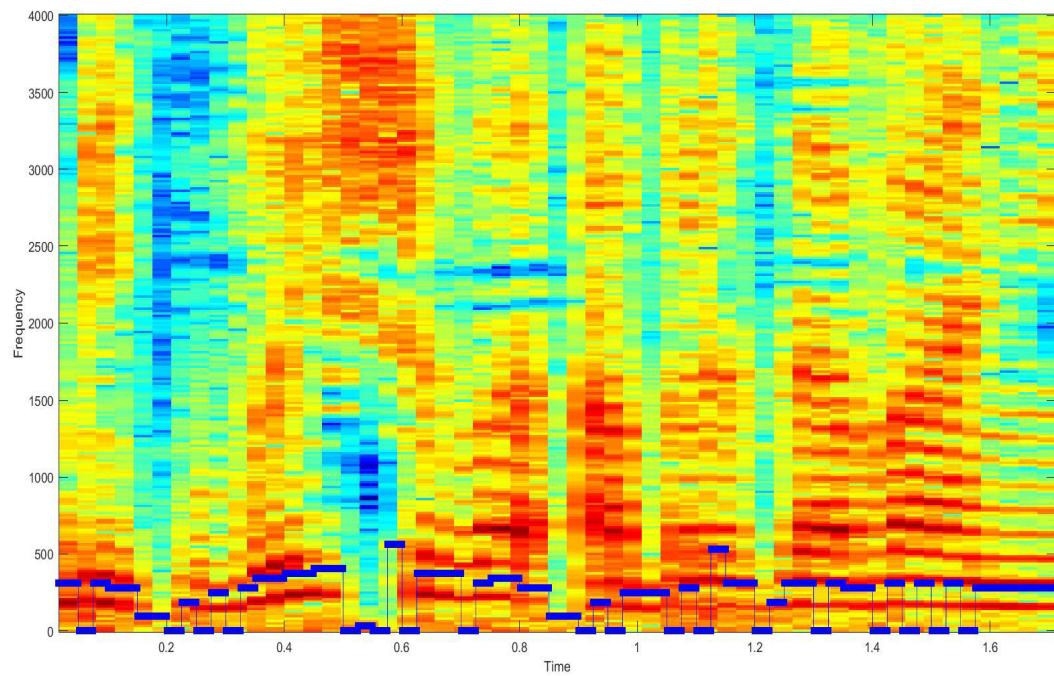


Рисунок 3.39 - Спектр речевого сигнала 2 “без высшего образования” и динамическое изменение частоты основного тона диктора (второй тон)

Изм.	Лист	№ докум.	Подпись	Дата	Лист
					11070006.11.03.02.369.ПЗВКР 44

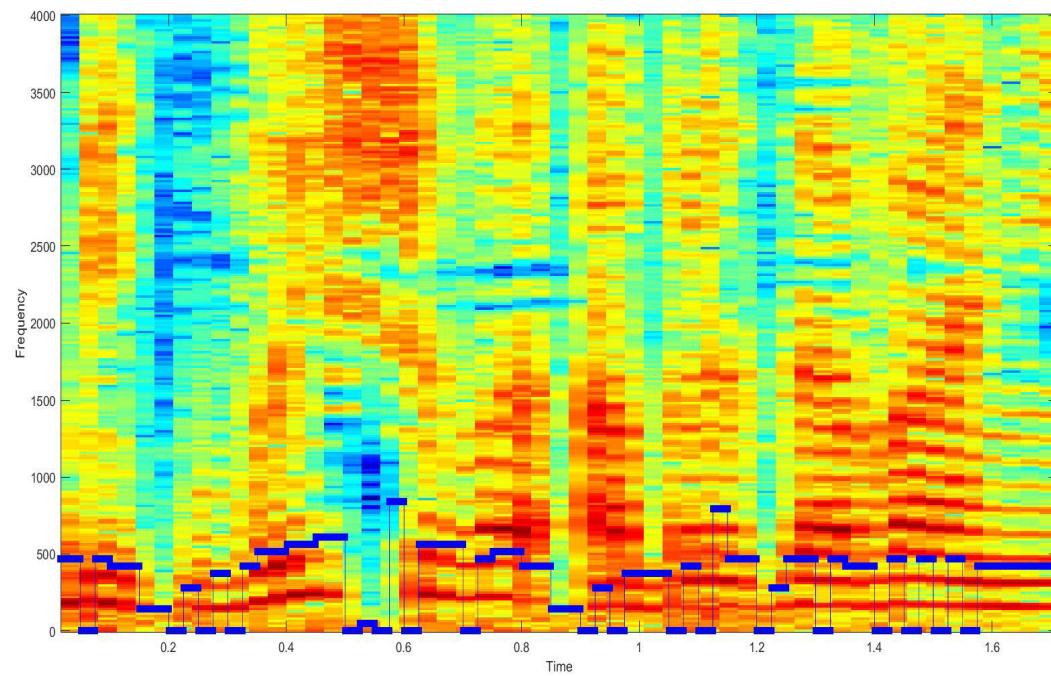


Рисунок 3.40 - Спектр речевого сигнала 2 “без высшего образования” и динамическое изменение частоты основного тона диктора (третий тон)

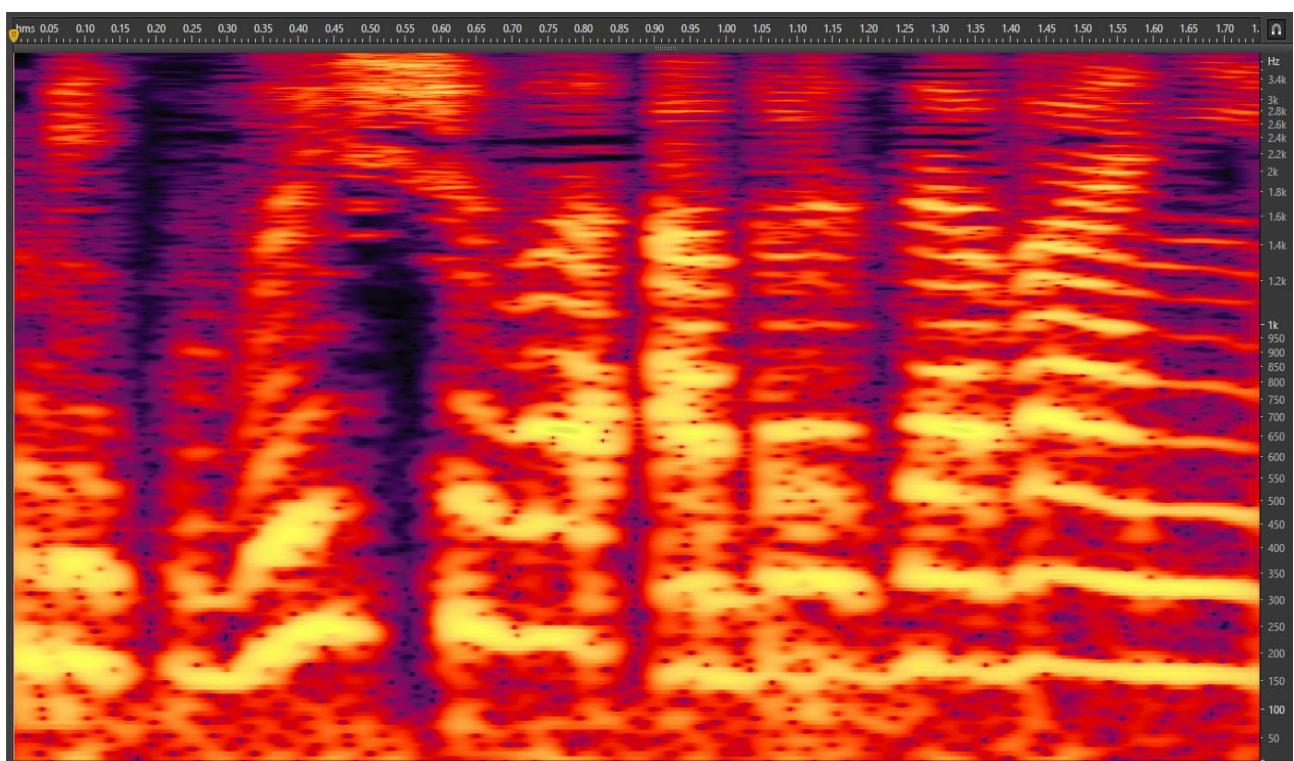


Рисунок 3.41 - Спектр речевого сигнала 2 “без высшего образования” в Adobe Audition

Изм.	Лист	№ докум.	Подпись	Дата	Лист
					45

11070006.11.03.02.369.ПЗВКР

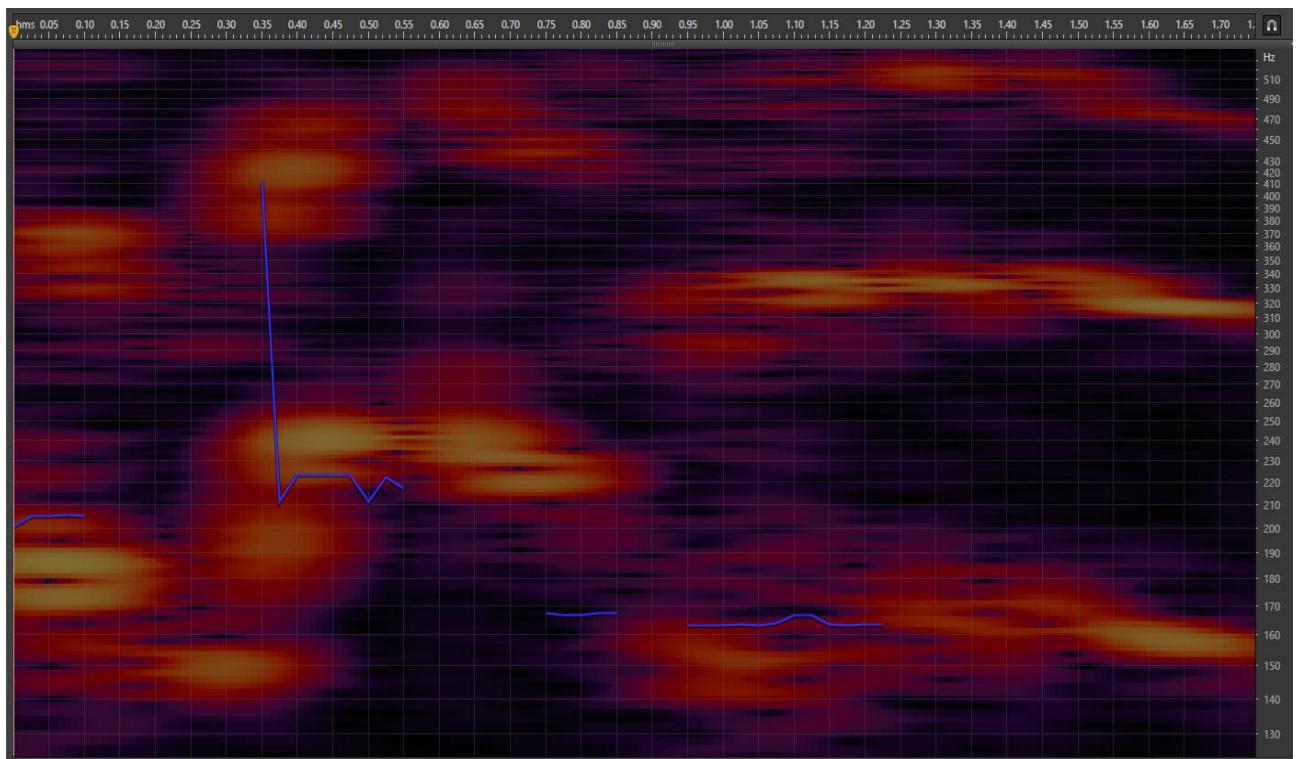


Рисунок 3.42 - Динамическое изменение частоты основного тона диктора речевого сигнала 2 “без высшего образования” в Adobe Audition

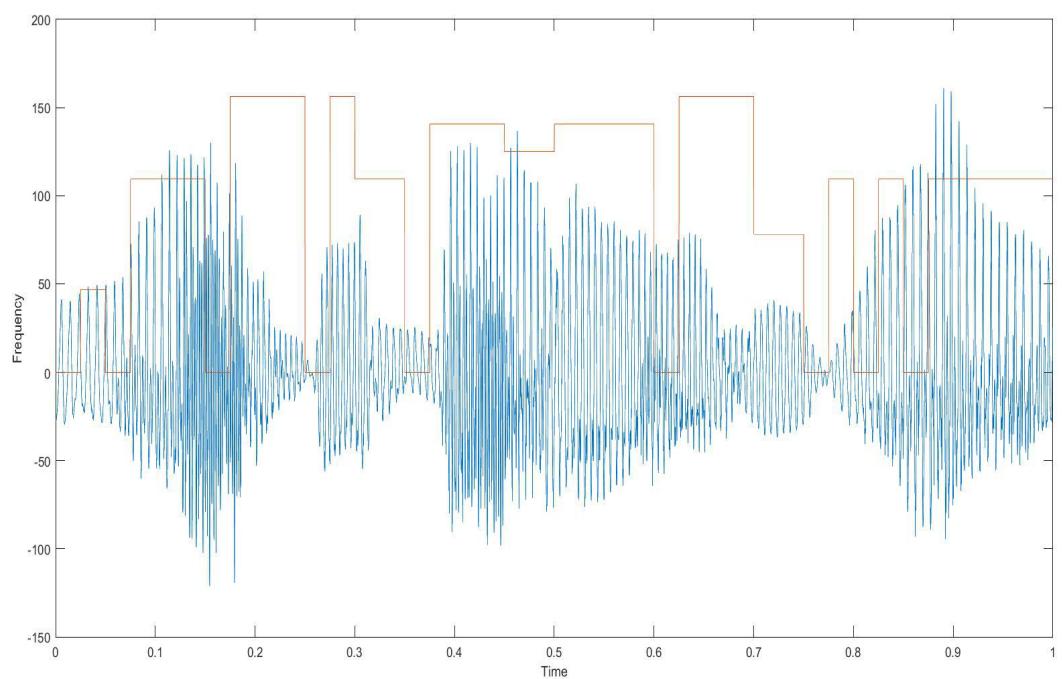


Рисунок 3.43 - Речевой сигнал 3 “я тебя люблю” и динамическое изменение частоты основного тона диктора

Изм.	Лист	№ докум.	Подпись	Дата
------	------	----------	---------	------

11070006.11.03.02.369.ПЗВКР

Лист
46

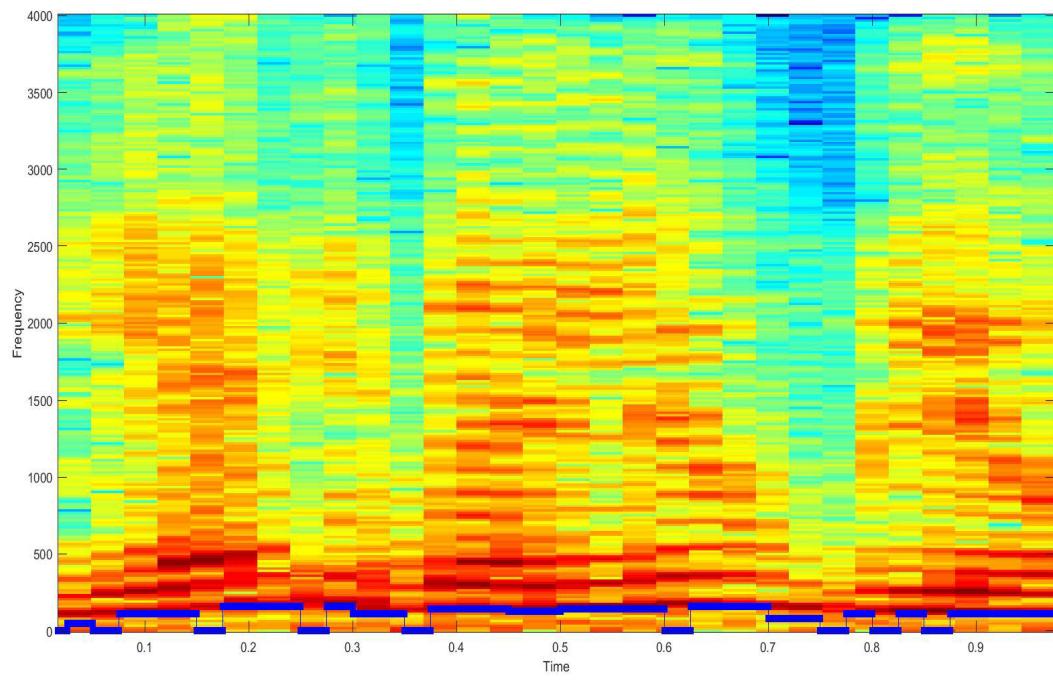


Рисунок 3.44 - Спектр речевого сигнала 3 “я тебя люблю” и динамическое изменение частоты основного тона диктора (первый тон)

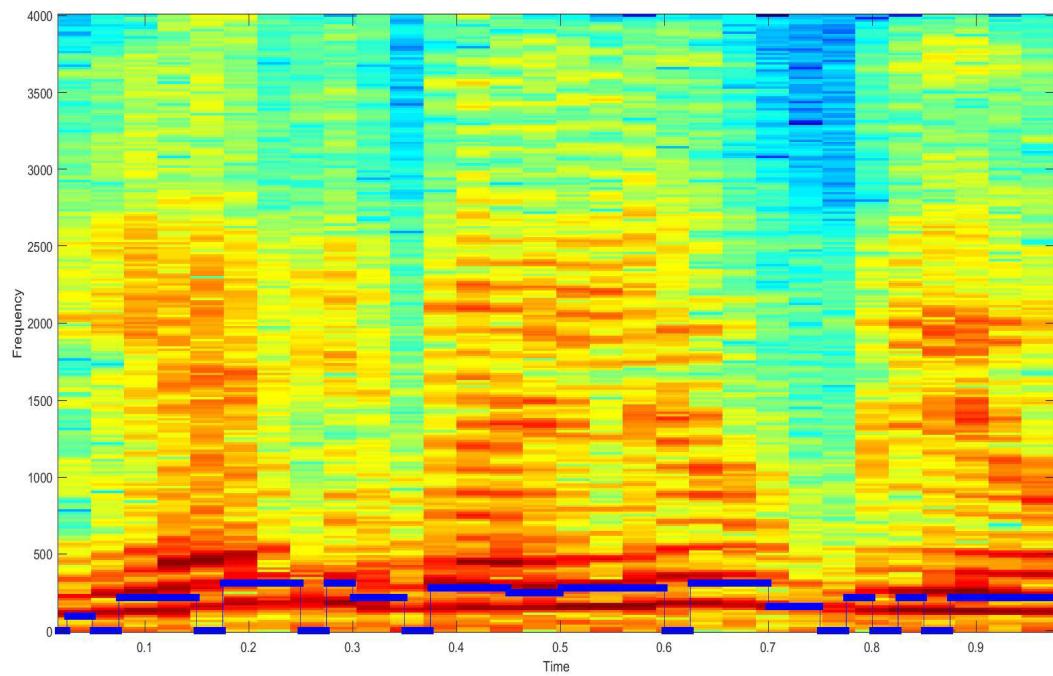


Рисунок 3.45 - Спектр речевого сигнала 3 “я тебя люблю” и динамическое изменение частоты основного тона диктора (второй тон)

Изм.	Лист	№ докум.	Подпись	Дата
------	------	----------	---------	------

11070006.11.03.02.369.ПЗВКР

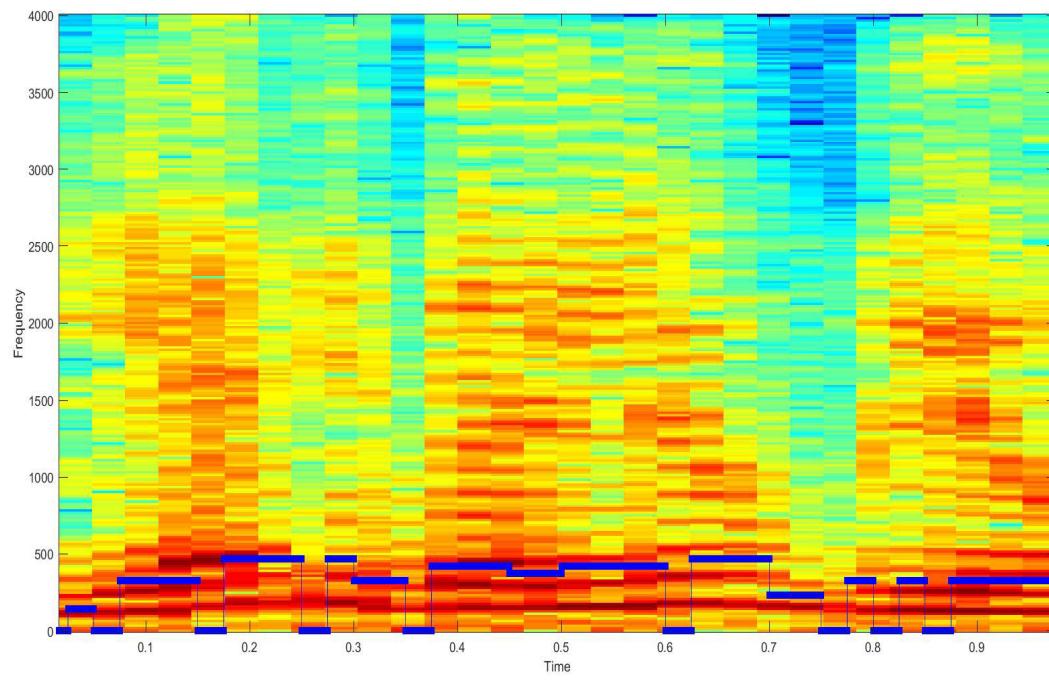


Рисунок 3.46 - Спектр речевого сигнала 3 “я тебя люблю” и динамическое изменение частоты основного тона диктора (третий тон)

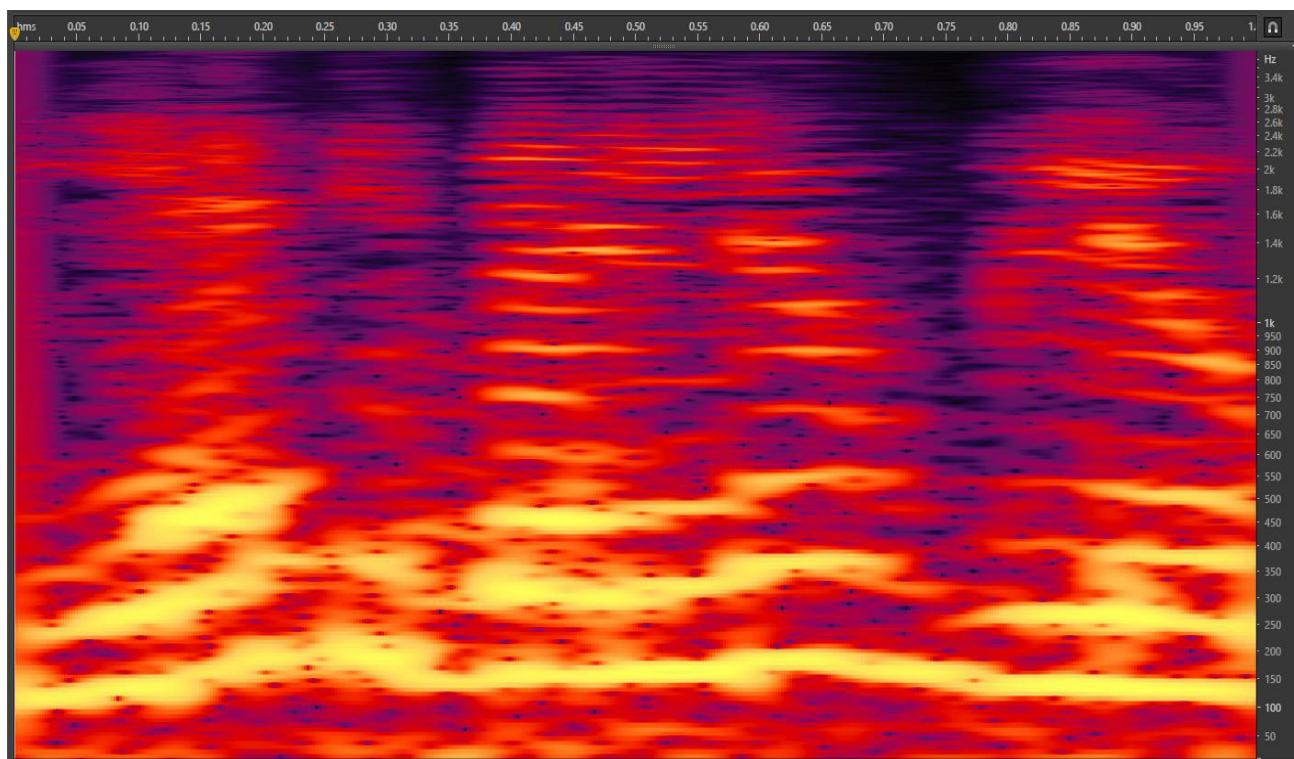


Рисунок 3.47 - Спектр речевого сигнала 3 “я тебя люблю” в Adobe Audition

Изм.	Лист	№ докум.	Подпись	Дата
------	------	----------	---------	------

11070006.11.03.02.369.ПЗВКР

Лист
48

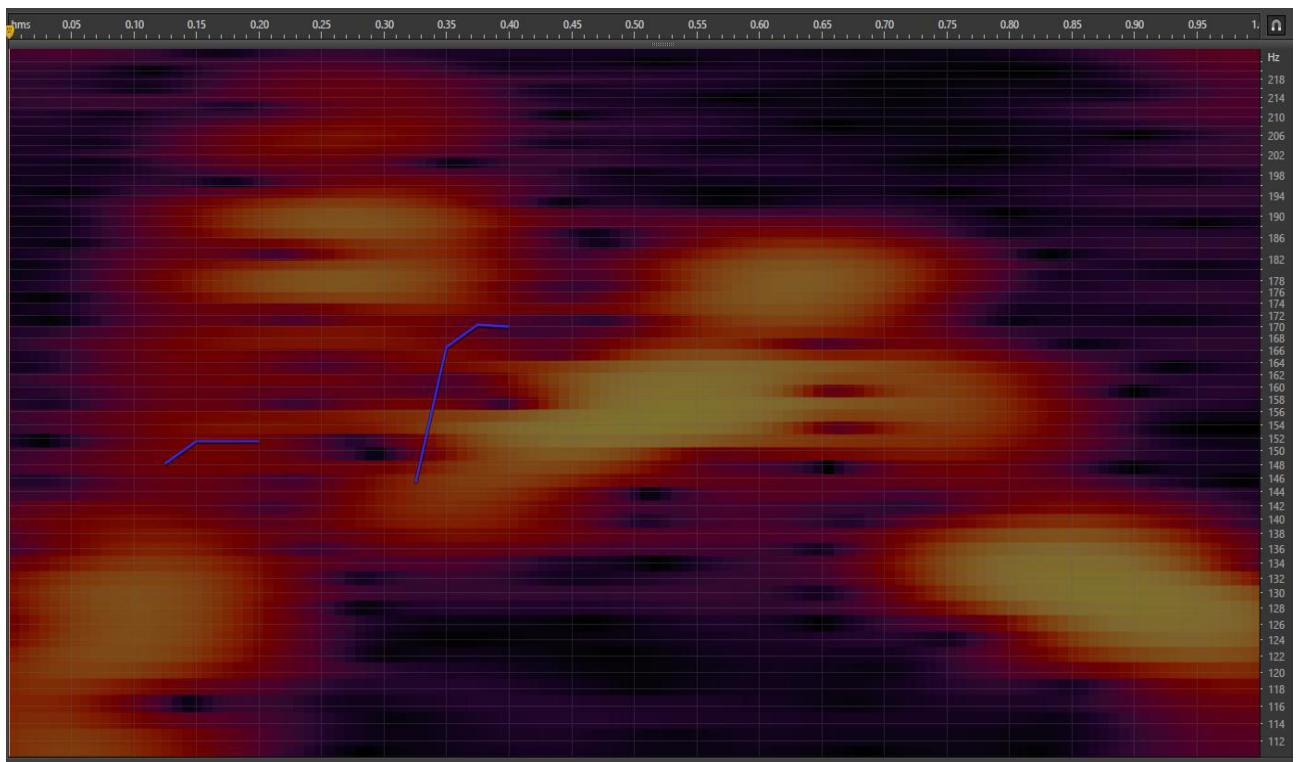


Рисунок 3.48 - Динамическое изменение частоты основного тона диктора речевого сигнала 3 “я тебя люблю” в Adobe Audition

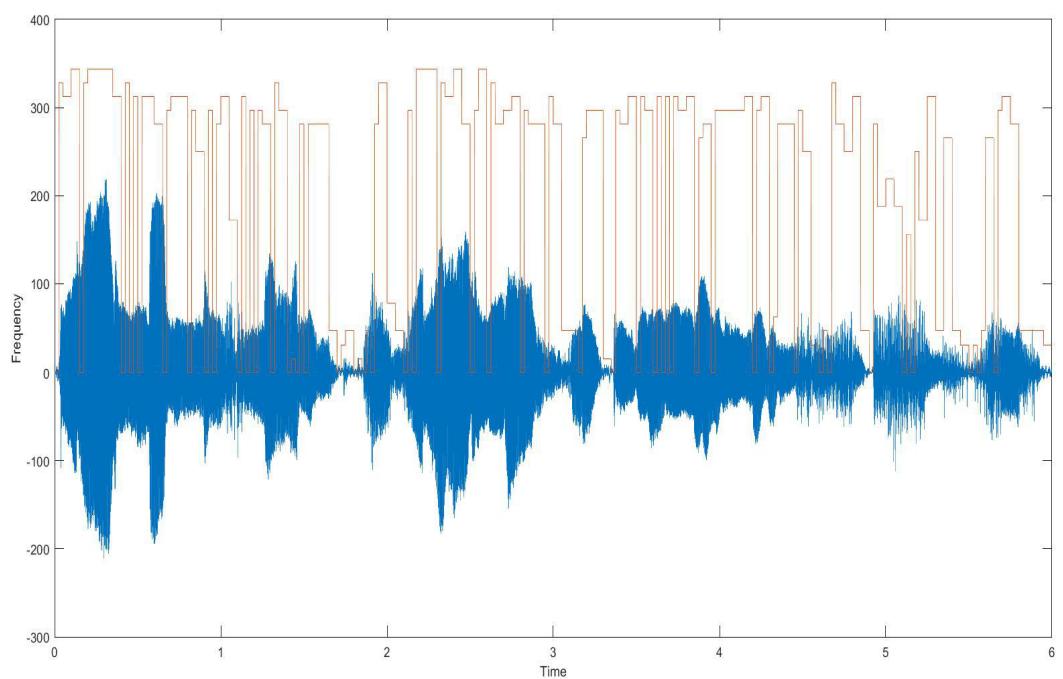


Рисунок 3.49 - Речевой сигнал 4 “ма” и динамическое изменение частоты основного тона диктора

Изм.	Лист	№ докум.	Подпись	Дата	Лист
					49

11070006.11.03.02.369.ПЗВКР

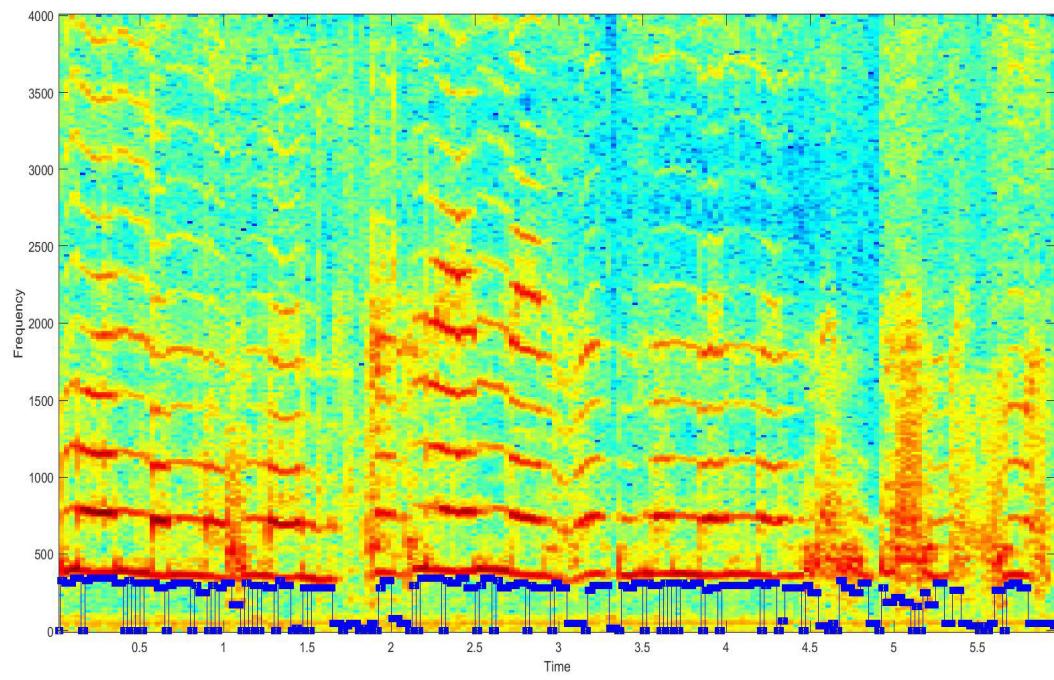


Рисунок 3.50 - Спектр речевого сигнала 4 “ма” и динамическое изменение частоты основного тона диктора (первый тон)

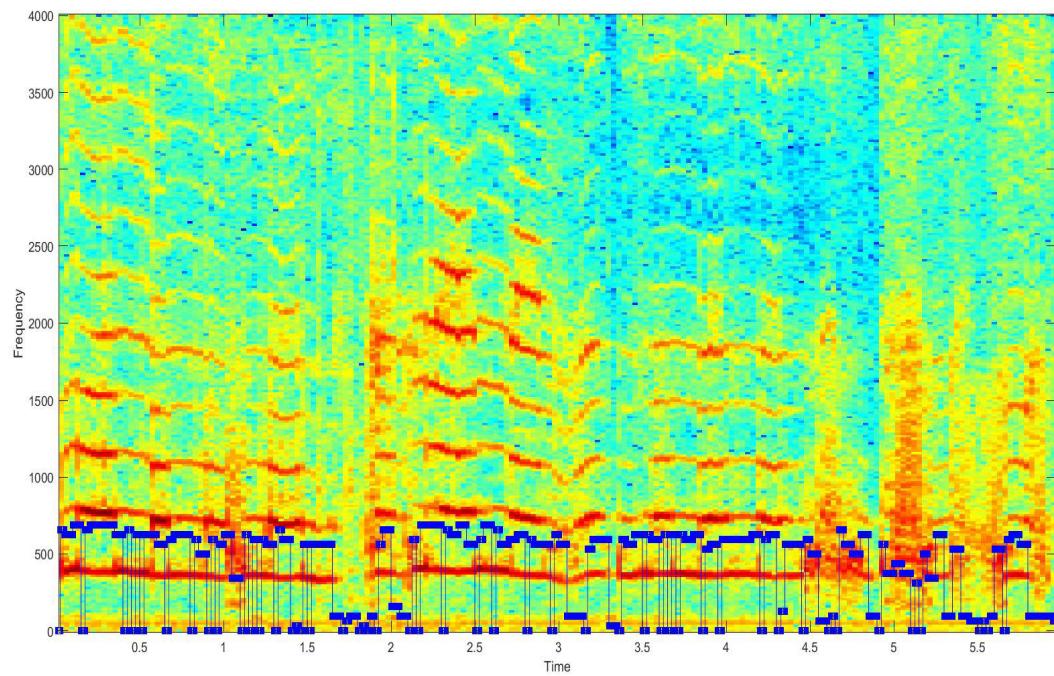


Рисунок 3.51 - Спектр речевого сигнала 4 “ма” и динамическое изменение частоты основного тона диктора (второй тон)

Изм.	Лист	№ докум.	Подпись	Дата
------	------	----------	---------	------

11070006.11.03.02.369.ПЗВКР

Лист
50

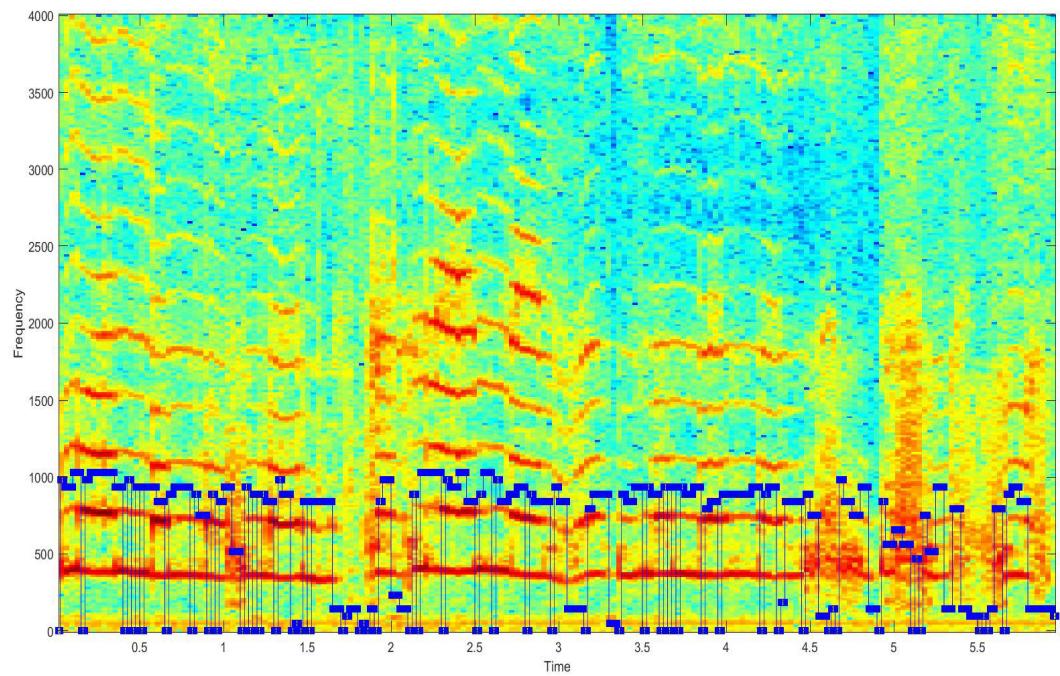


Рисунок 3.52 - Спектр речевого сигнала 4 “ма” и динамическое изменение частоты основного тона диктора (третий тон)

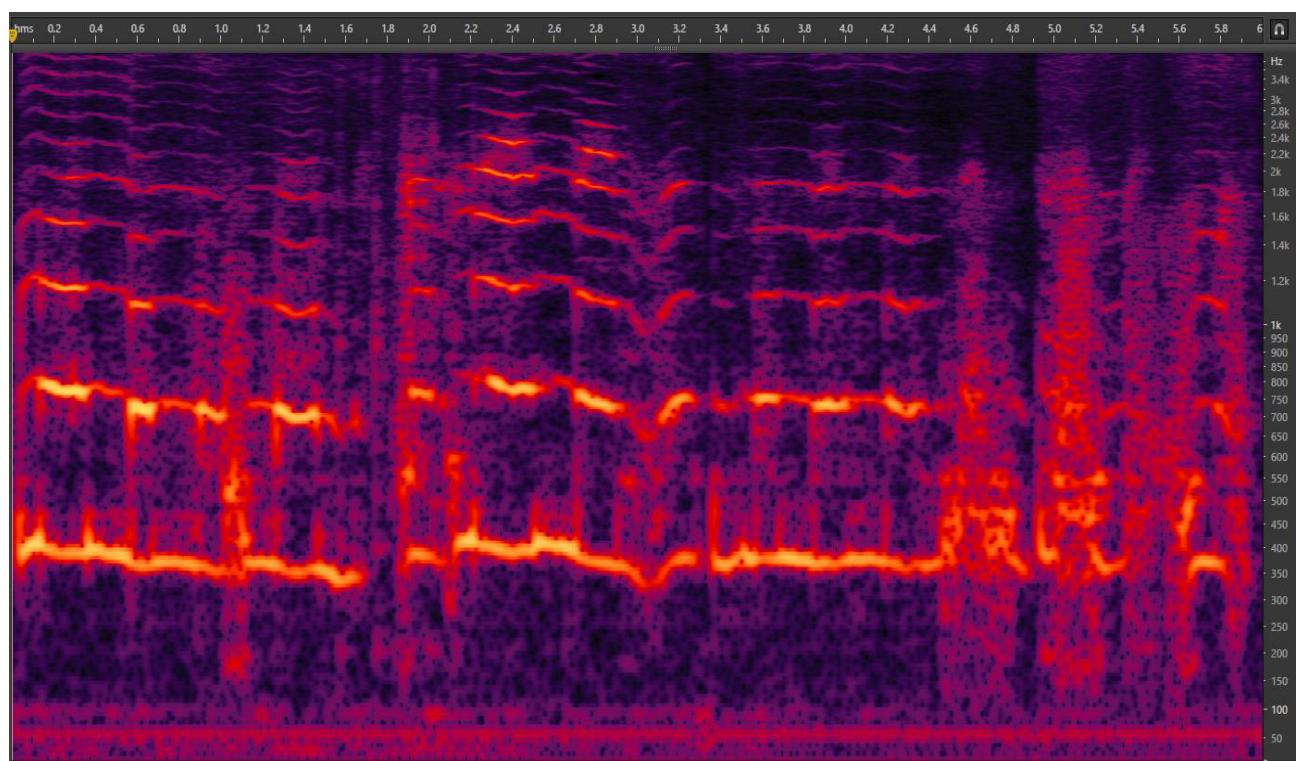


Рисунок 3.53 - Спектр речевого сигнала 4 “ма” в Adobe Audition

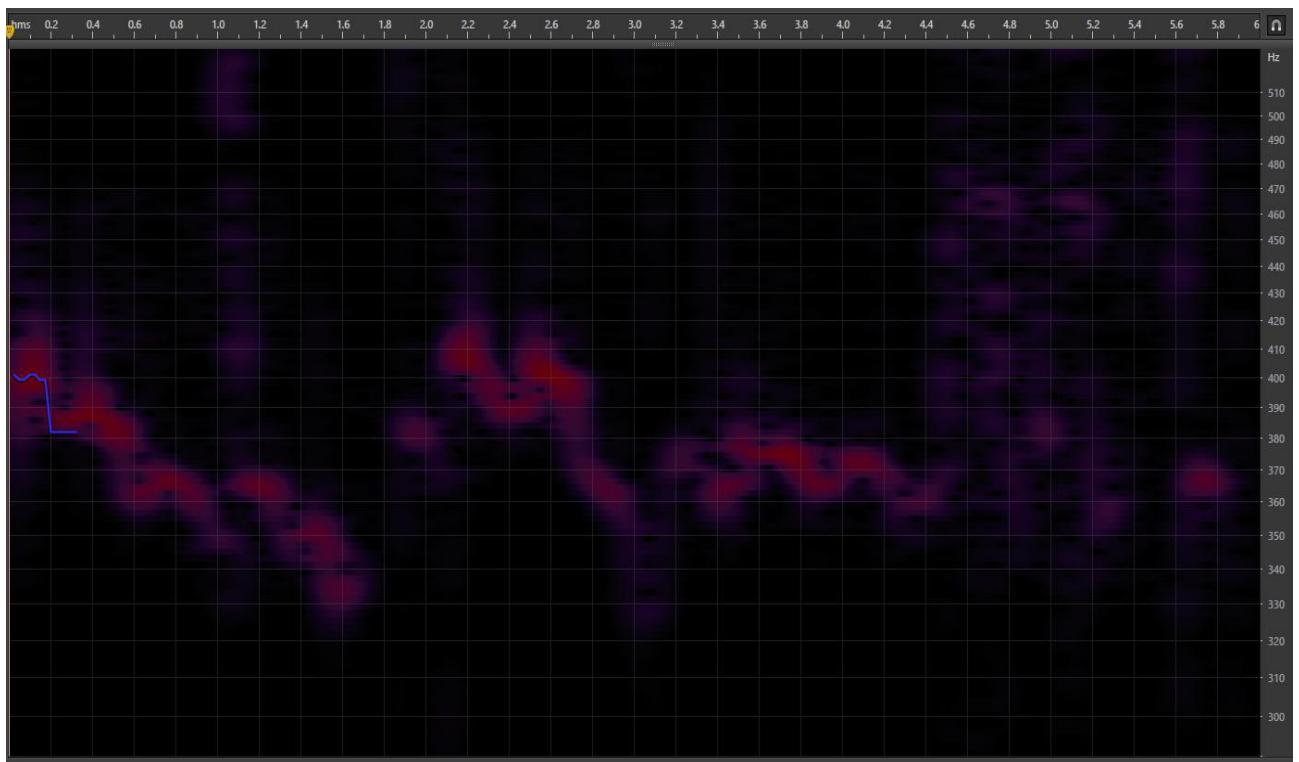


Рисунок 3.54 - Динамическое изменение частоты основного тона диктора речевого сигнала 4 “ма” в Adobe Audition

Из рисунков видно, что динамическое изменение частоты основного тона диктора с использованием алгоритма оценивания основного тона больше точнее, чем в Adobe Audition.

4 СИНТЕЗ ЗВУКОВ С ИСПОЛЬЗОВАНИЕМ СТРУКТУРЫ ФОНЕМ

Представленные в разделах 2 и 3 алгоритмы позволяют создать описание фонем с точки зрения динамического изменения частот формант на протяжении звучания фонемы, что может быть использовано в задаче синтеза речевых сигналов для обеспечения естественности.

В данном разделе представлен алгоритм синтеза речевых сигналов, состоящий из нескольких этапов.

Этап 1. Разделение сигнала на отрезки

Исходный сигнал разделяется на фонемы.

В качестве примера исходного сигнала на рисунке 4.1 приведен сигнал "шишка" длиной $L = 6400$ отсчётов.

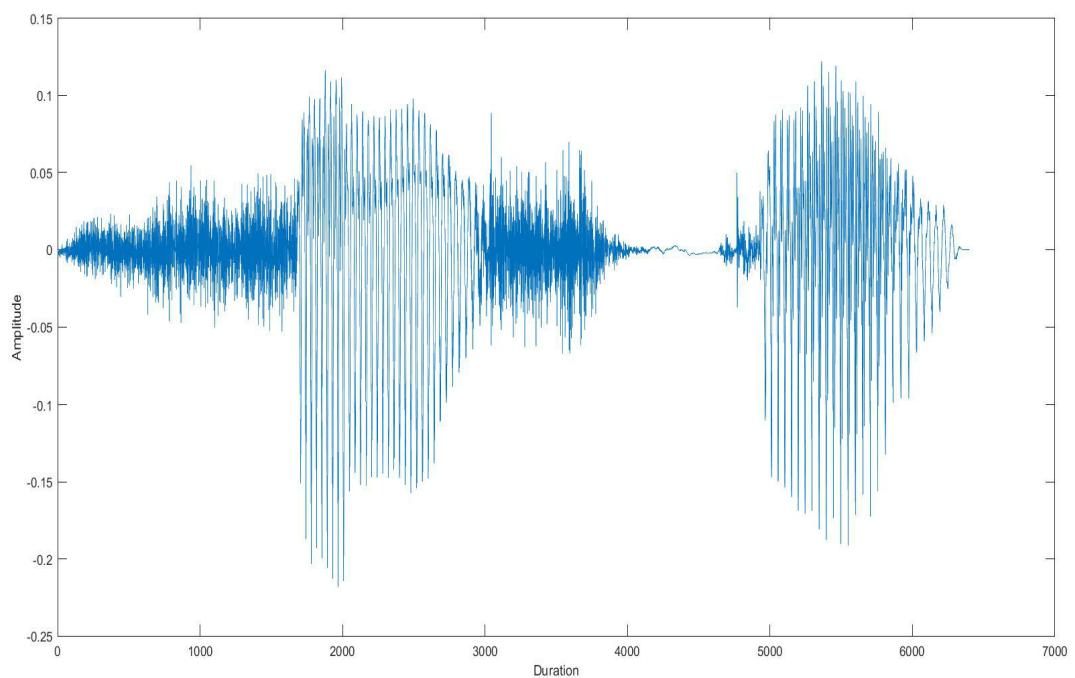


Рисунок 4.1 - Исходный сигнал "шишка" длиной $L = 6400$ отсчётов

Сигнал был разделен на фонемы "шиш" длительностью $N_1 = 4400$

Изм.	Лист	№ докум.	Подпись	Дата	Лист
					53

11070006.11.03.02.369.ПЗВКР

отсчётов (рисунок 4.2) и "ка" длительностью $N_2 = 2000$ отсчётов (рисунок 4.3).

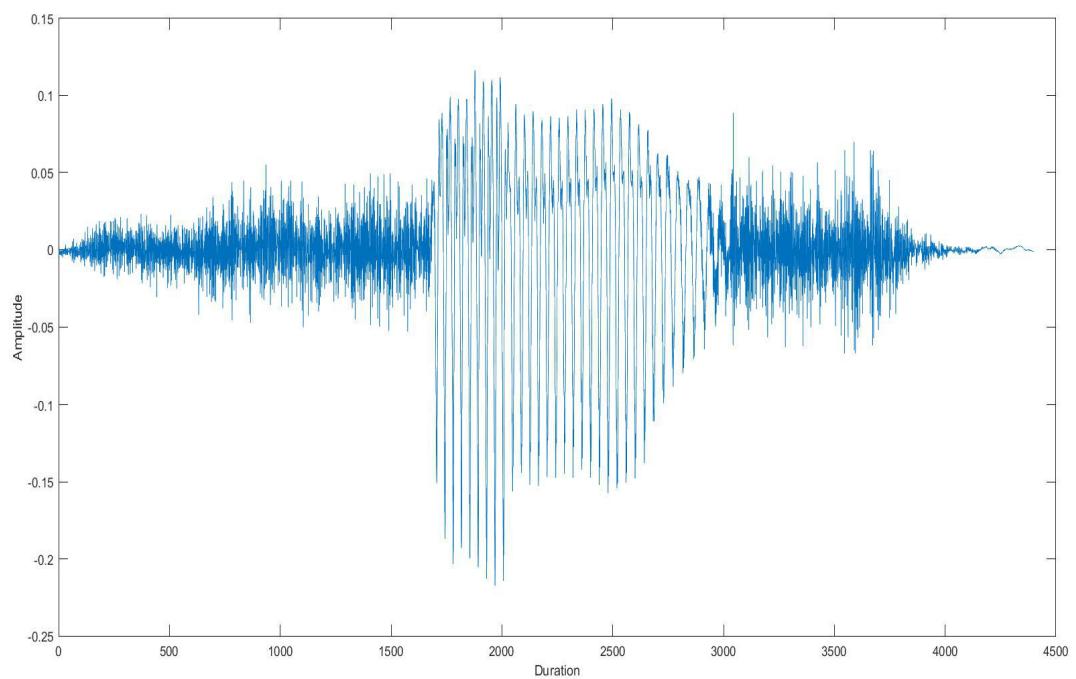


Рисунок 4.2 - Разделение сигнала на "шиш"

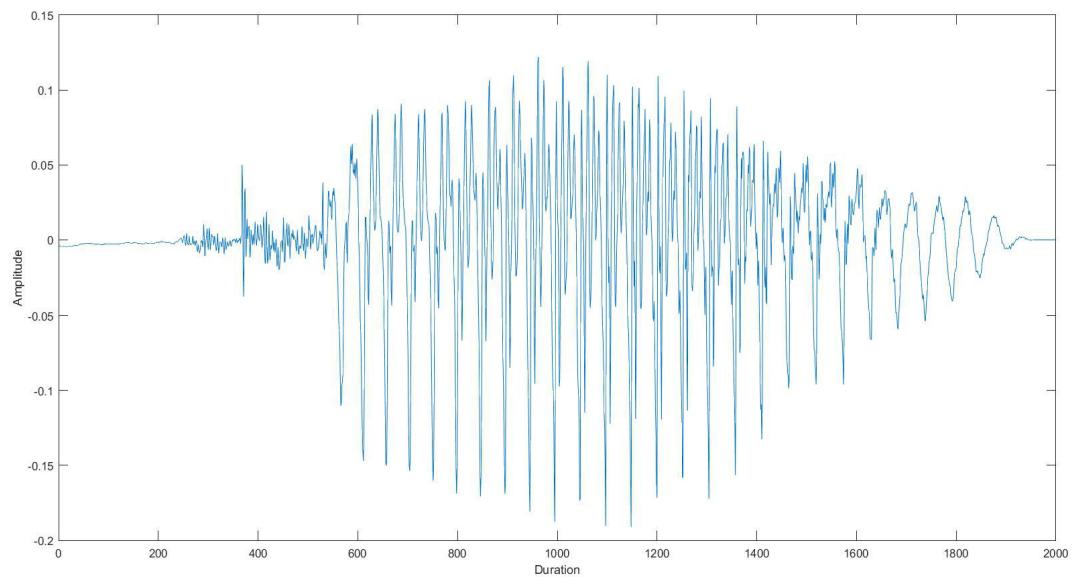


Рисунок 4.3 - Разделение сигнала на "ка"

Этап 2. Спектральный анализ и формантный синтез

Осуществляется частотный анализ методом ДПФ. Выбираются форманты

Изм.	Лист	№ докум.	Подпись	Дата
------	------	----------	---------	------

11070006.11.03.02.369.ПЗВКР

Лист
54

и определяются их параметры (частота, амплитуда, фаза) (таблицы 4.1, 4.2).

Таблица 4.1 - Формантные частоты для "шиш"

Формантные частоты	Амплитуды	Фазы
211	16.77	-2.578
426	11.43	-0.966
637	4.66	-1.450
848	1.11	-0.614
1055	0.92	-0.423

Таблица 4.2 - Формантные частоты для "ка"

Формантные частоты	Амплитуды	Фазы
164	15.45	0.139
312.5	7.13	-2.001
473	2.43	-0.514
641	8.72	0.061
773	6.59	-2.729
926	2.91	-1.579

На рисунках 4.4, 4.5 показаны результаты кратковременного анализа фонем.

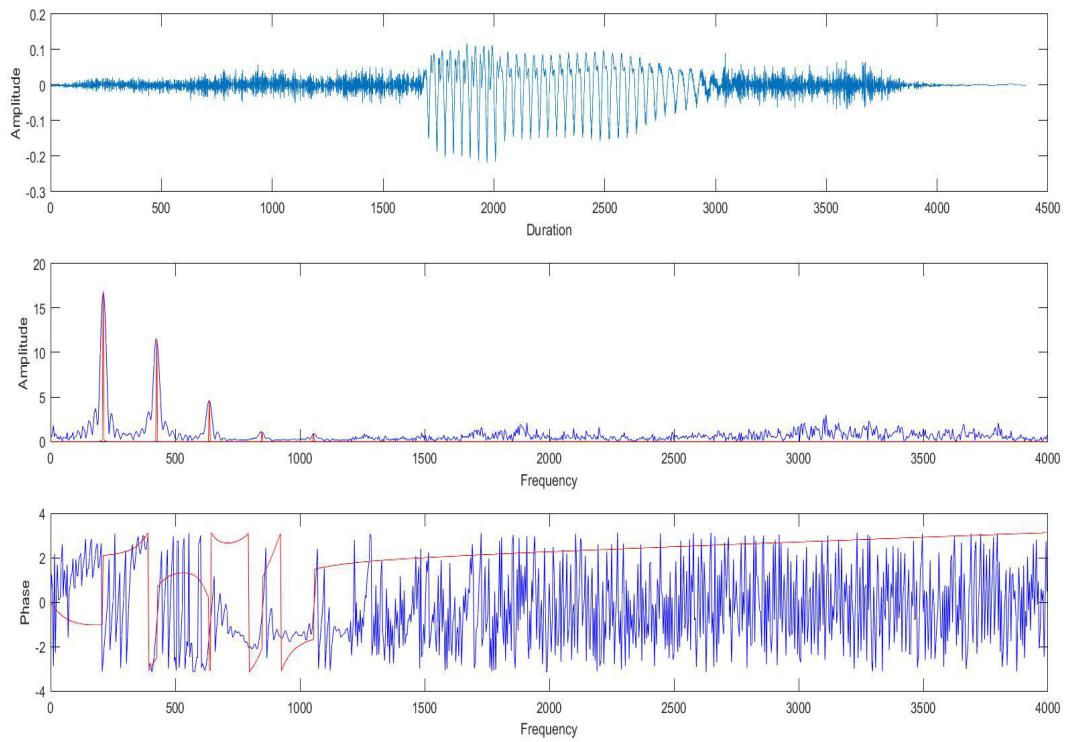


Рисунок 4.4 - Спектральный анализ и формантный синтез для сигнала "шиш"

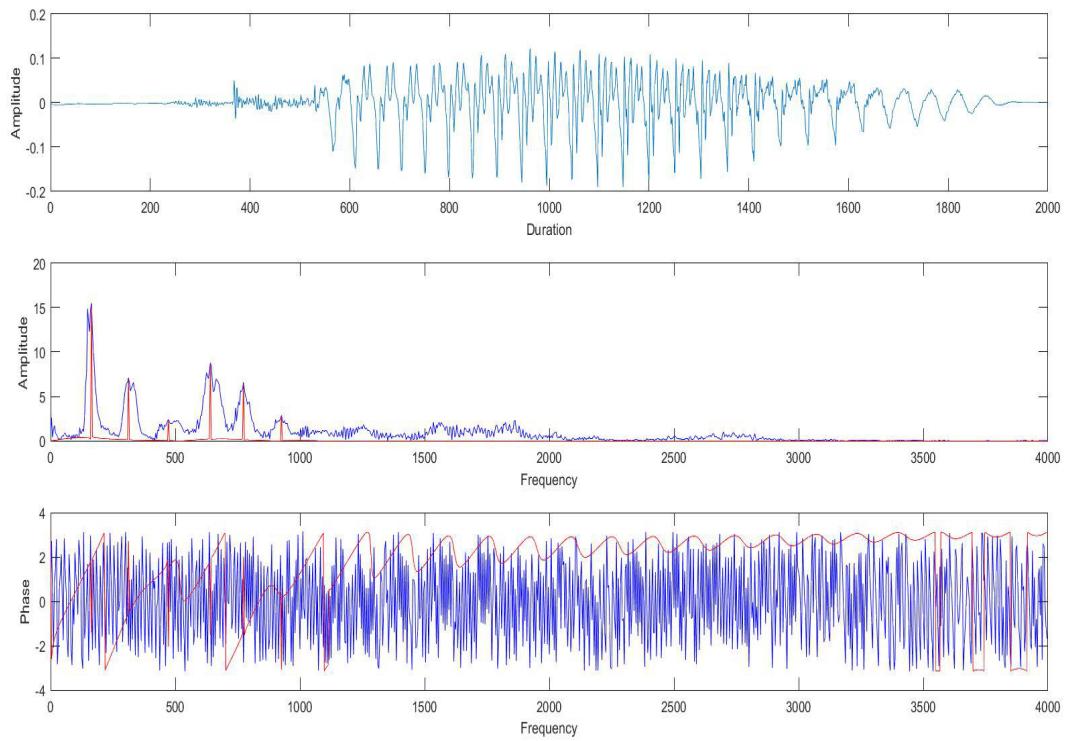


Рисунок 4.5 - Спектральный анализ и формантный синтез для сигнала "ка"

Изм.	Лист	№ докум.	Подпись	Дата
------	------	----------	---------	------

11070006.11.03.02.369.ПЗВКР

Лист
56

На данном этапе синтеза может использоваться как результат кратковременного анализа отрезка речевого сигнала (таблицы 4.1, 4.2), так и набор параметров, полученный в результате исследования динамики их изменения (разделы 2, 3).

Этап 3. Амплитудная модуляция

Частота амплитудной модуляции выбирается исходя из оценки основного тона конкретного сигнала, либо из диапазона основного тона, по выбору.

В приведенном примере частота амплитудной модуляции составляет $F_{c1} = 16$ Гц.

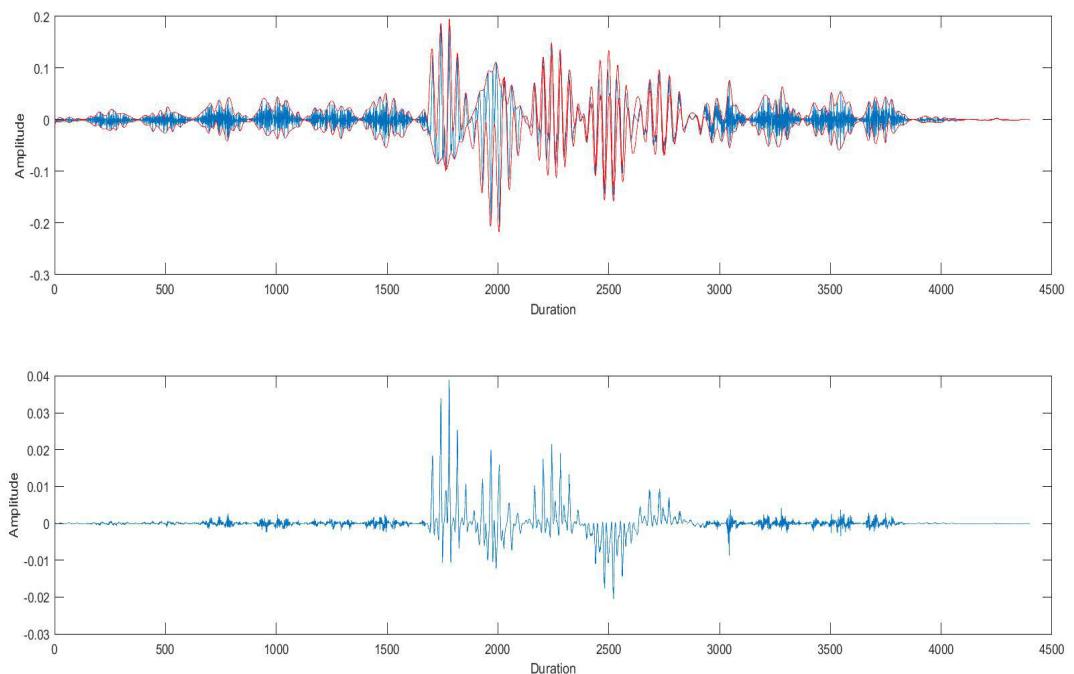


Рисунок 4.6 - Амплитудная модуляция для "шиш" и новый "шиш"

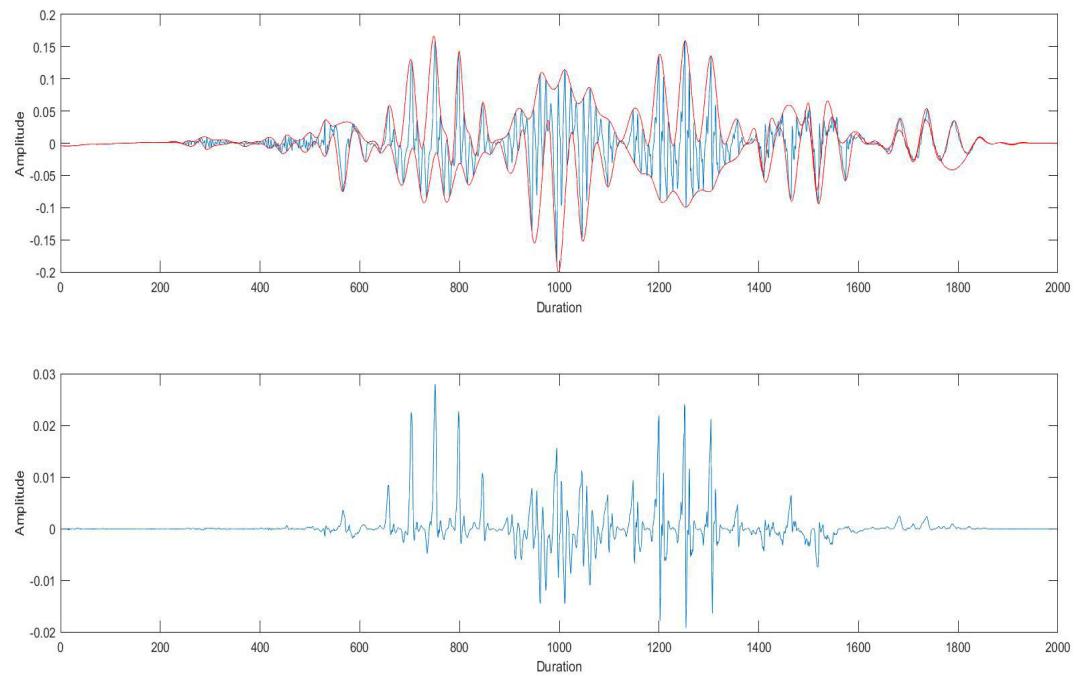


Рисунок 4.7 - Амплитудная модуляция для "ка" и новый "ка"

Аналогичные эксперименты были проведены для других речевых сигналов.

Сигнал "мама" длиной $L = 6000$ отсчётов.

Изм.	Лист	№ докум.	Подпись	Дата
------	------	----------	---------	------

11070006.11.03.02.369.ПЗВКР

Лист
58

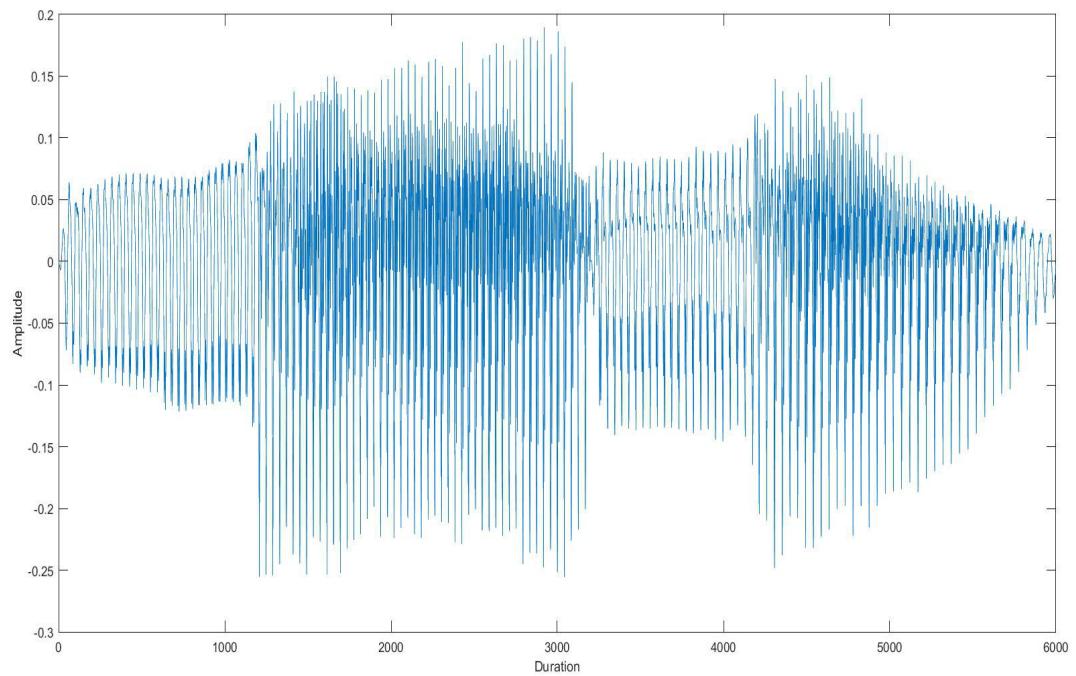


Рисунок 4.8 - Исходный сигнал "мама" длиной L = 6000 отсчётов

Сигнал был разделен на "ма" длительностью $N_1 = 3200$ отсчётов (рисунок 4.9) и "ма" длительностью $N_2 = 2800$ отсчётов (рисунок 4.10).

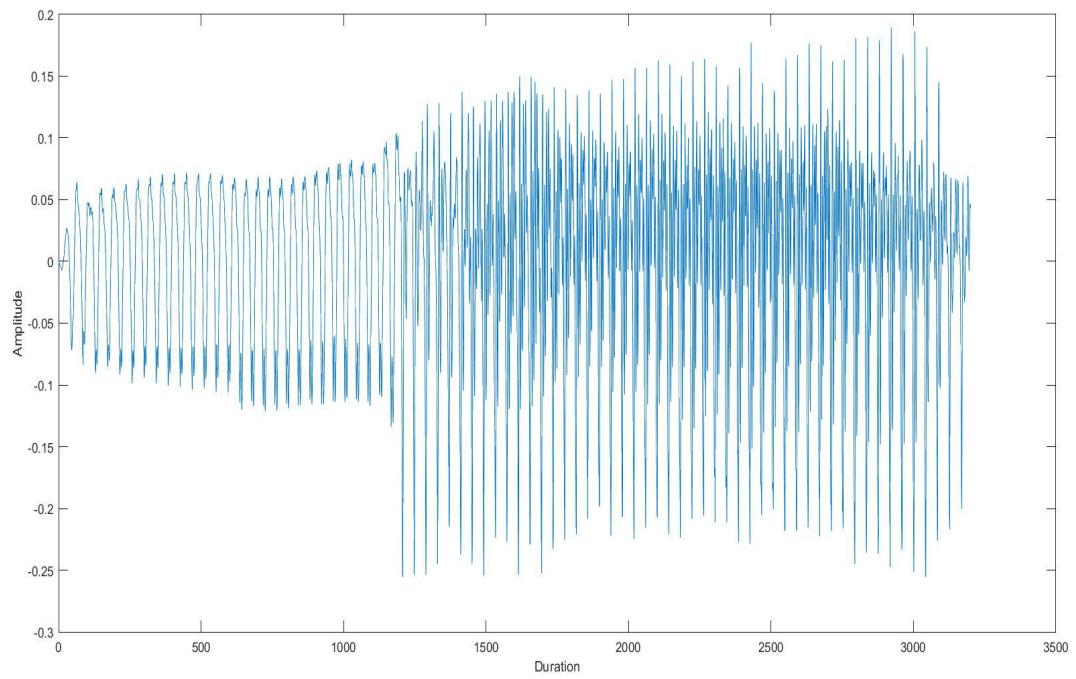


Рисунок 4.9 - Разделение сигнала на "ма"

Изм.	Лист	№ докум.	Подпись	Дата
------	------	----------	---------	------

11070006.11.03.02.369.ПЗВКР

Лист
59

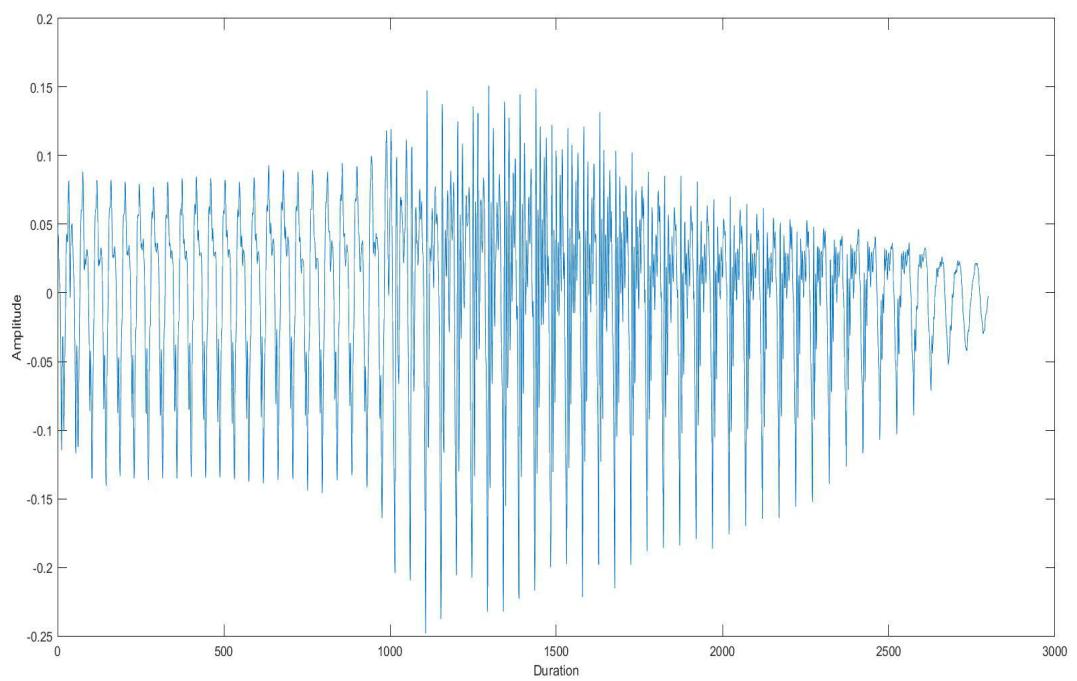


Рисунок 4.10 - Разделение сигнала на "ма"

Результат частотного анализа приведен в таблицах 4.1, 4.2 и на рисунках 4.11, 4.12.

Таблица 4.1 - Формантные частоты для "ма"

Формантные частоты	Амплитуды	Фазы
195.5	56.24	0.226
394.5	18.76	0.414
590	18.66	-1.636
789	19.63	3.064
988	15.12	2.332
1188	16.74	-1.094
1379	6.40	-1.411

Таблица 4.2 - Формантные частоты для "ма"

Формантные частоты	Амплитуды	Фазы
184	34.76	1.820
367	13.58	-0.111
512	10.14	-1.553
684	8.37	0.670
875	4.96	-2.624
1035	6.53	2.052
1191	7.34	-0.578
1332	8.26	-2.938

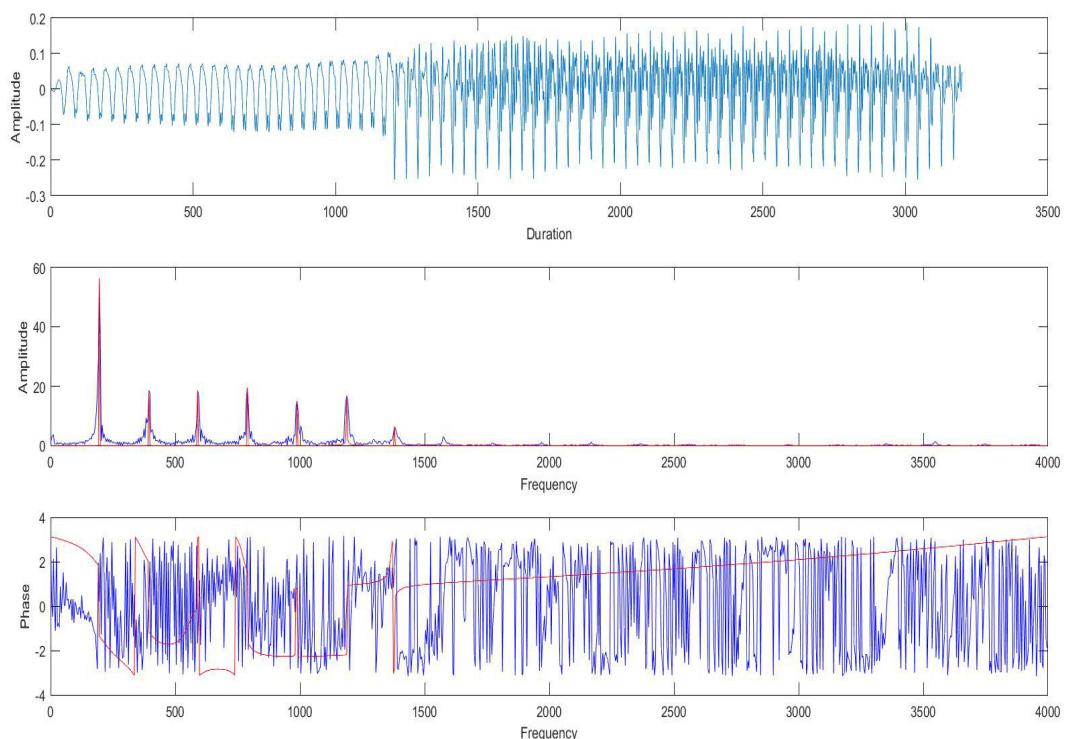


Рисунок 4.11 - Спектральный анализ и формантный синтез для сигнала "ма"

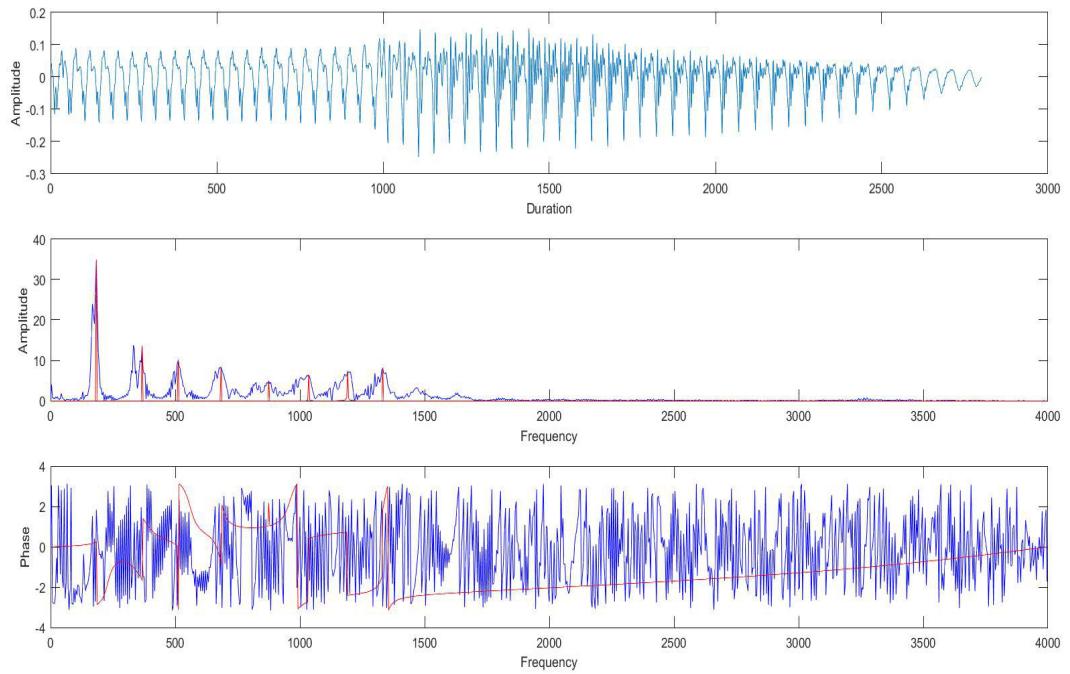


Рисунок 4.12 - Спектральный анализ и формантный синтез для сигнала "ма"

В этом примере частота амплитудной модуляции $F_{c2} = 176$.

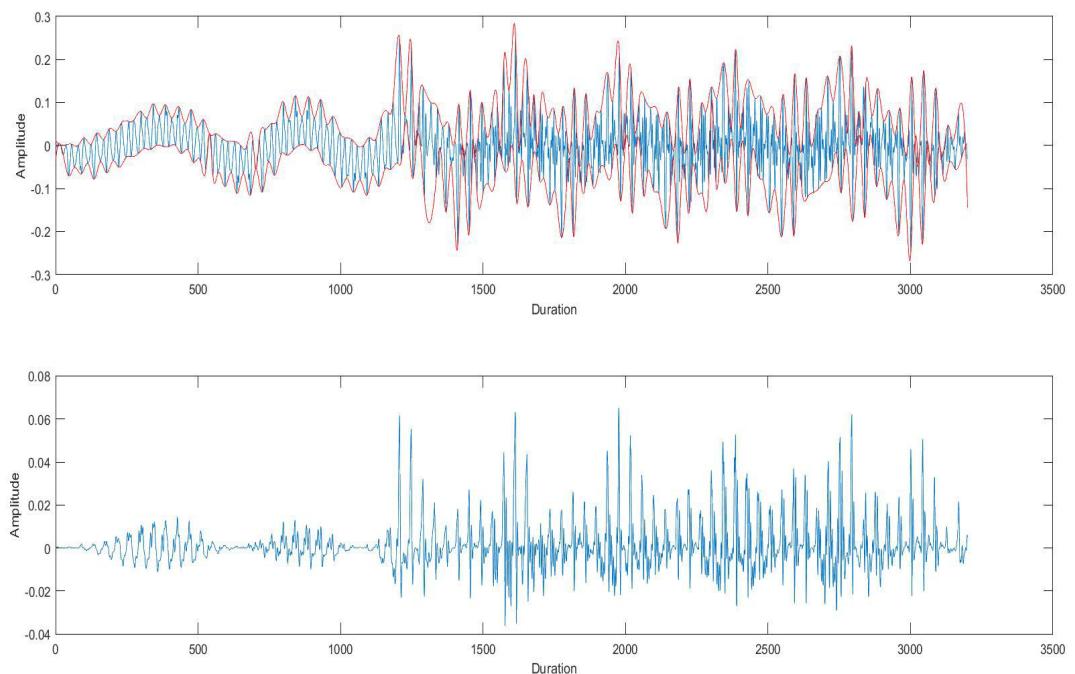


Рисунок 4.13 - Амплитудная модуляция для "ма" и новый "ма"

Изм.	Лист	№ докум.	Подпись	Дата
------	------	----------	---------	------

11070006.11.03.02.369.ПЗВКР

Лист

62

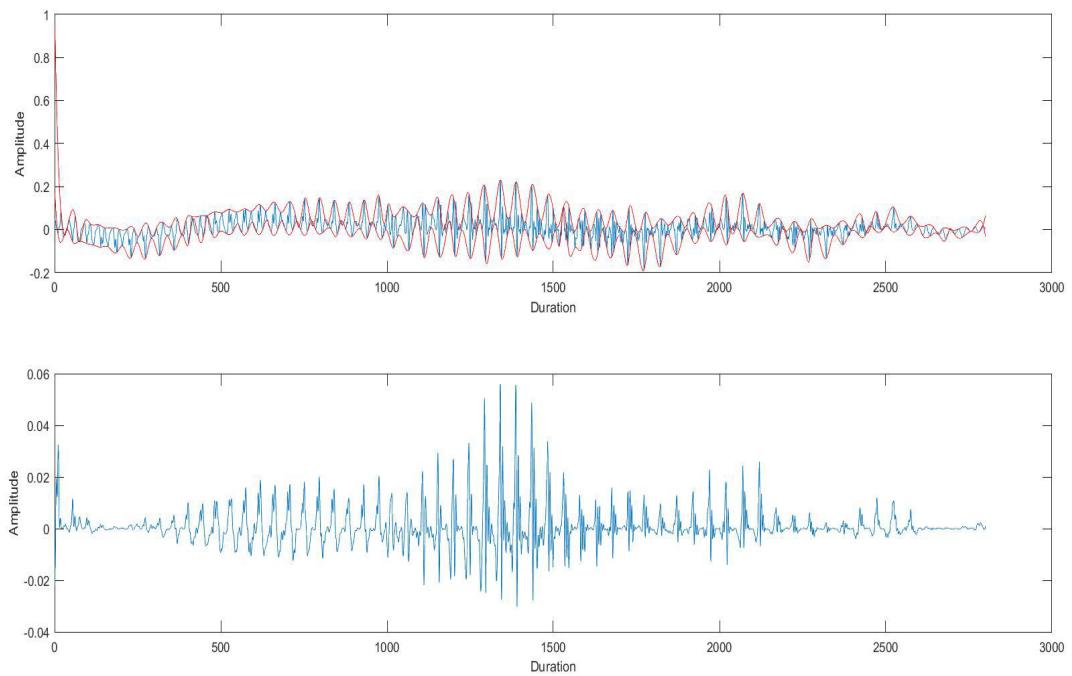


Рисунок 4.14 - Амплитудная модуляция для "ма" и новый "ма"

Сигнал "почта" длиной $L = 6800$ отсчётов.

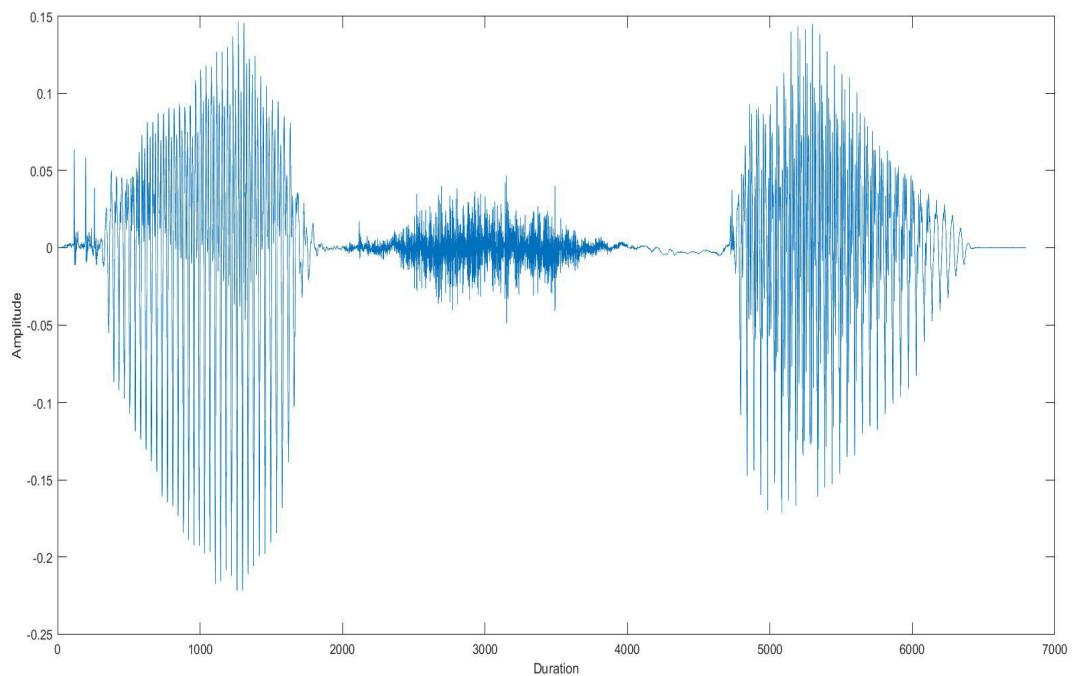


Рисунок 4.15 - Исходный сигнал "почта" длиной $L = 6800$ отсчётов

Изм.	Лист	№ докум.	Подпись	Дата
------	------	----------	---------	------

11070006.11.03.02.369.ПЗВКР

Лист

63

Сигнал был разделен на "пoch" длительностью $N1 = 3200$ отсчётов (рисунок 4.16) и "ta" длительностью $N2 = 2800$ отсчётов (рисунок 4.17).

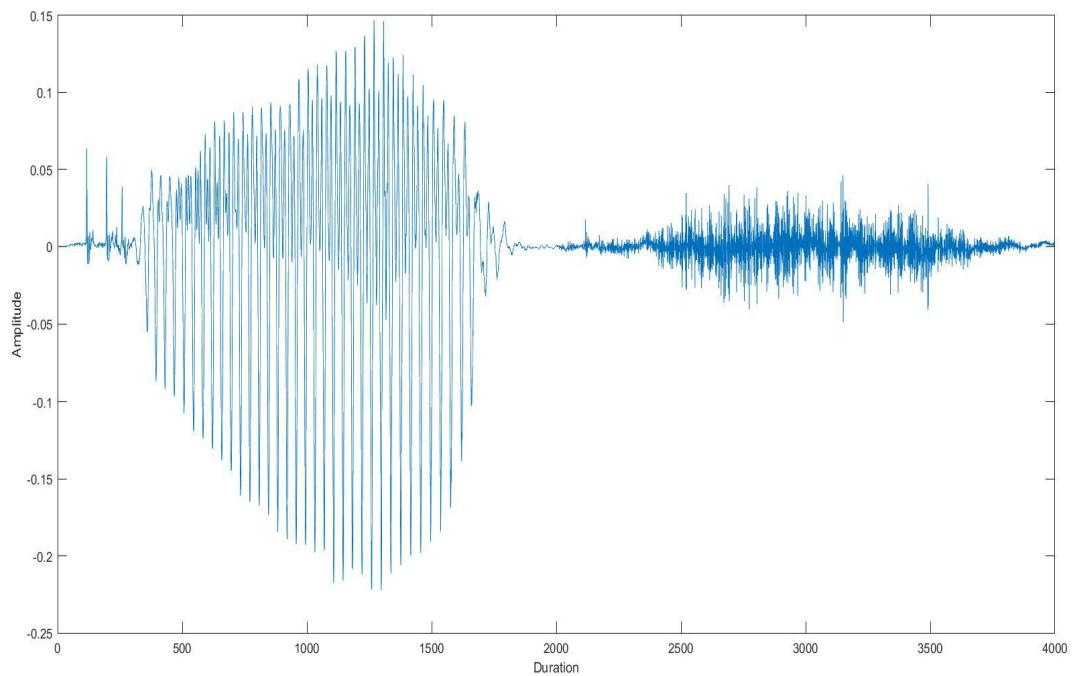


Рисунок 4.17 - Разделение сигнала на "пoch"

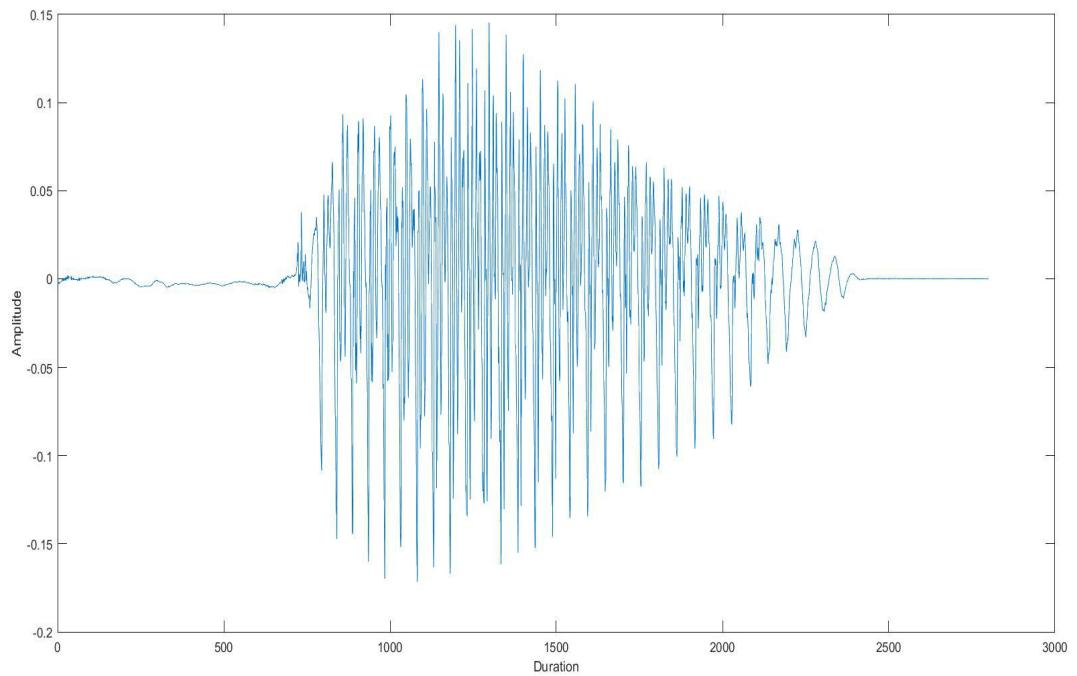


Рисунок 4.18 - Разделение сигнала на "та"

Изм.	Лист	№ докум.	Подпись	Дата
------	------	----------	---------	------

11070006.11.03.02.369.ПЗВКР

Лист

64

Результат частотного анализа приведен в таблицах 4.5, 4.6 и на рисунках 4.19, 4.20.

Таблица 4.5 - Формантные частоты для "пoch"

Формантные частоты	Амплитуды	Фазы
211	33.09	1.465
426	31.75	-2.872
640	6.93	-0.636

Таблица 4.6 - Формантные частоты для "та"

Формантные частоты	Амплитуды	Фазы
156	15.92	1.837
305	9.96	-1.467
500	4.23	1.949
633	10.88	1.871
762	6.41	-1.871

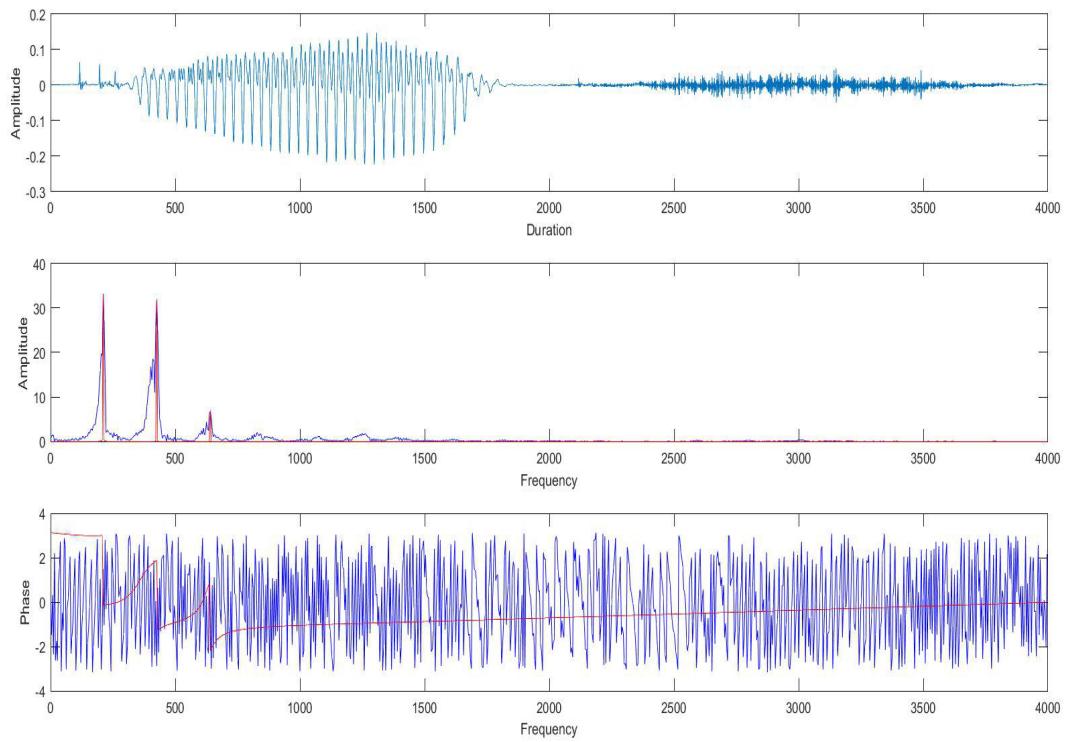


Рисунок 4.19 - Спектральный анализ и формантный синтез для сигнала "поч"

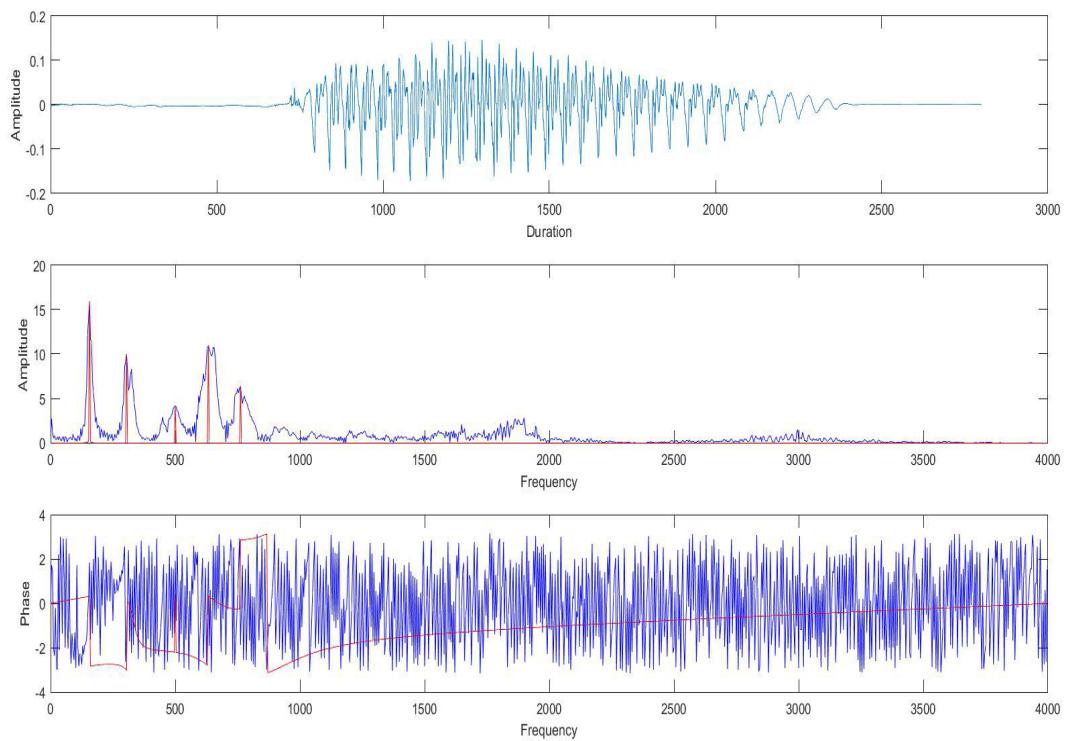


Рисунок 4.20 - Спектральный анализ и формантный синтез для сигнала "та"

Изм.	Лист	№ докум.	Подпись	Дата
------	------	----------	---------	------

11070006.11.03.02.369.ПЗВКР

Лист

66

В этом примере частота амплитудной модуляции $F_{c3} = 199$.

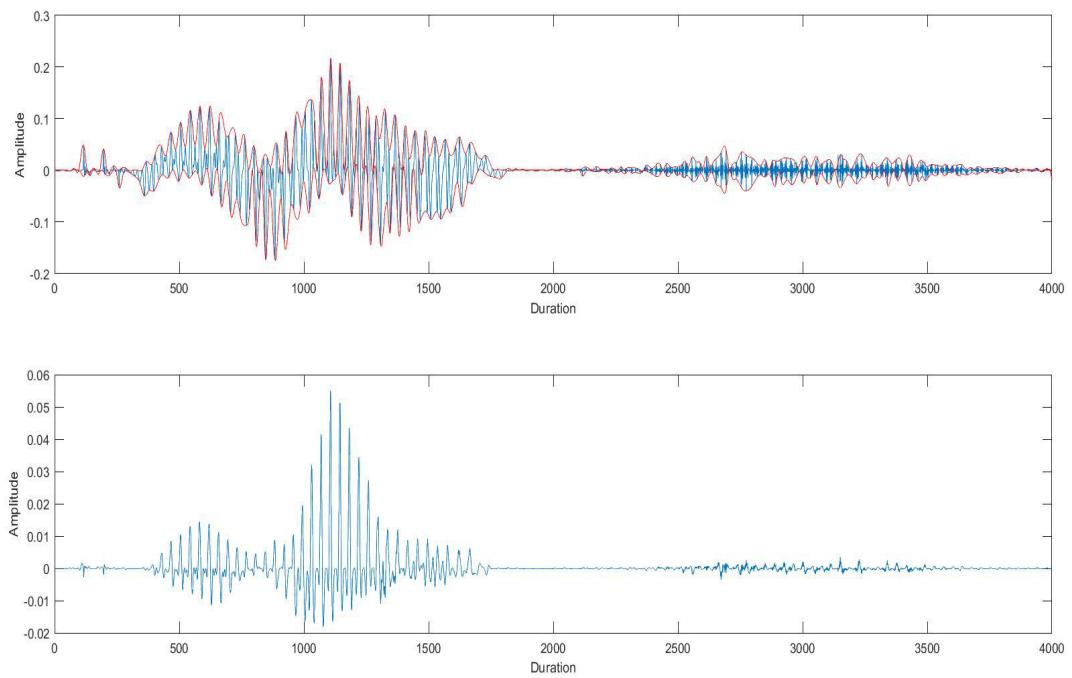


Рисунок 4.21 - Амплитудная модуляция для "пoch" и новый "пoch"

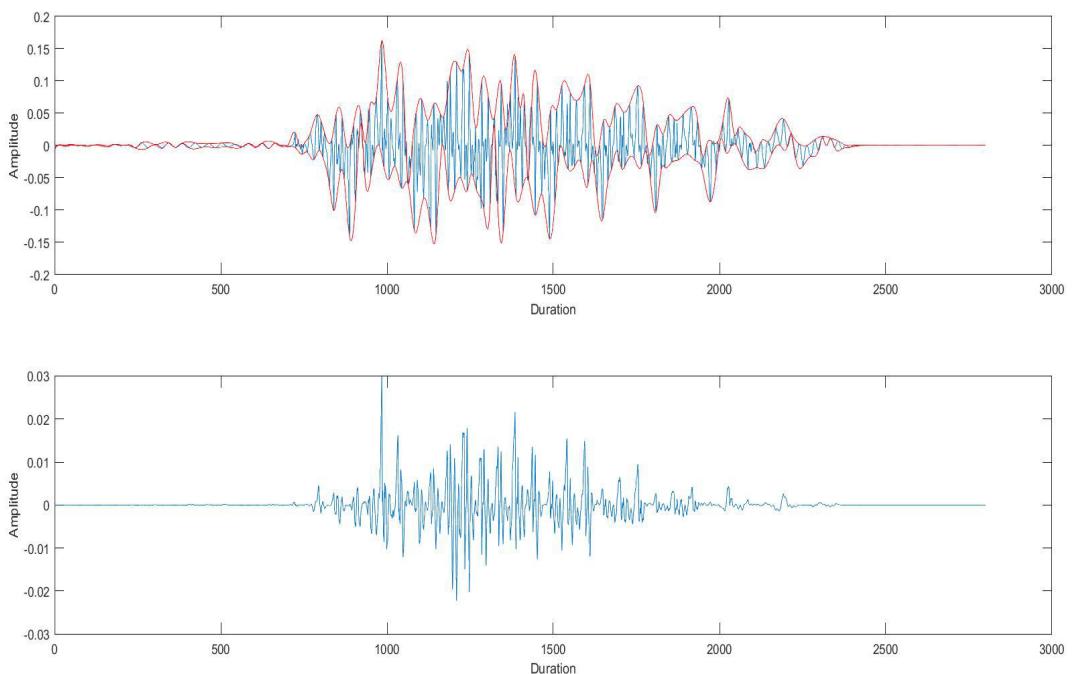


Рисунок 4.22 - Амплитудная модуляция для "та" и новый "та"

Изм.	Лист	№ докум.	Подпись	Дата
------	------	----------	---------	------

11070006.11.03.02.369.ПЗВКР

Лист

67

Воспроизведение синтезированных сигналов показало, что абсолютного сходства достигнуть не удалось, но синтезированный сигнал похож «на слух» на оригинал.

Изм.	Лист	№ докум.	Подпись	Дата
------	------	----------	---------	------

11070006.11.03.02.369.ПЗВКР

Лист

68

5 ЭКОНОМИЧЕСКАЯ ОЦЕНКА РЕЗУЛЬТАТОВ ИССЛЕДОВАНИЯ

Основной целью научно-исследовательской или опытно-конструкторской работы является проведение работ, направленных на анализ, проектирование или разработку каких-либо устройств. Результатом таких работ могут являться разработанный прототип прибора или программный продукт, выполняющий определенные функции, рекомендации по эксплуатации прибора или технологии, нормативные акты и т.д.

5.1 Планирование работ по исследованию

В реализации исследования заняты следующие специалисты:

- главный инженер или старший научный сотрудник (заведующий лабораторией), осуществляющий общее руководство исследованием;
- инженер I категории или младший научный сотрудник, проводящий разработку, исследование, необходимые расчеты, составляющий техническую документацию на исследование;
- экономист, дающий экономическую оценку исследования.

Расчет сроков проведения и трудоемкости представлен в таблице 5.1.

Таблица 5.1 - Планирование работ по исследованию

Наименование этапов работ	Исполнитель	Трудоемкость, час	Продолжи- тельность, дней
1	2	3	4
1. Подготовительный			
1.1. Сбор информации	Младший научный сотрудник	40	5

Окончание таблицы 5.1

<i>1</i>	<i>2</i>	<i>3</i>	<i>4</i>
1.2. Выработка идеи	Старший научный сотрудник	40 40	5 5
1.3. Определение объема исследовательских работ	Младший научный сотрудник	24	3
1.4. Формирование исследовательской работы	Младший научный сотрудник	16	2
1.5. Обработка и анализ информации	Младший научный сотрудник	72	9
Итого:		232	29
2. Основной (экономический анализ)			
2.1. Обоснование целесообразности работы	Старший научный сотрудник	24	3
2.2. Выполнение работы	Младший научный сотрудник	104	13
Итого:		128	16
3. Заключительный			
3.1. Технико-экономическое обоснование	Экономист	40	5
3.2. Оформление и утверждение документации	Младший научный сотрудник	40	5
Итого:		80	10

Результатом планирования является расчет трудоемкости исследования по часам и количеству дней. Определена численность штата производственного персонала, необходимого для проведения исследования.

5.2 Расчет расходов на оплату труда на исследование

Расчет расходов на оплату труда разработки исследования представлен в

Изм.	Лист	№ докум.	Подпись	Дата	11070006.11.03.02.369.ПЗВКР									Лист 70

таблице 5.2.

Таблица 5.2 - Расчет расходов на оплату труда

Должность исполнителей	Трудоемкость, час	Оклад, руб
Младший научный сотрудник	296	13000
Старший научный сотрудник	104	15000
Экономист	40	10000
Итого:	440	

Часовая тарифная ставка (χ_{TC}) рассчитывается по формуле:

$$\chi_{TC} = \frac{P}{F_{мес}} \quad (5.1)$$

где Р - оклад сотрудника; $F_{мес}$ - фонд рабочего времени месяца, составляет 176 часов (22 рабочих дня по 8 часов в день).

Расход на оплату труда (P_{OT}) находится следующим образом:

$$P_{OT} = \chi_{TC} * T_{сум} \quad (5.2)$$

где $T_{сум}$ - суммарная трудоемкость каждого из исполнителей.

Результаты расчетов сведены в таблицу 5.3.

Таблица 5.3 - Расчет расходов на оплату труда

Должность исполнителей	Трудоемкость, час	Оклад, руб	$\chi_{TC},$ руб/час	$P_{OT},$ руб
1	2	3	4	5
Младший научный сотрудник	296	13000	73,86	21862,56

Окончание таблицы 5.3

<i>1</i>	<i>2</i>	<i>3</i>	<i>4</i>	<i>5</i>
Старший научный сотрудник	104	15000	85,23	8863,92
Экономист	40	10000	56,82	2272,80
Итого:	440			32999,28

5.3 Расчет продолжительности исследования

Согласно расчетам, трудоемкость исследования составила 440 часов.

Продолжительность исследования составит:

$$T_{иссл} = T_{сум} / T_{РД} \quad (5.3)$$

где $T_{сум} = 440$ часов суммарная трудоемкость исследования; $T_{РД} = 8$ часов - продолжительность рабочего дня.

$$T_{иссл} = 440 / 8 = 55 \text{ дней.}$$

Продолжительность исследования составляет 55 дня, расчет производится без учета выходных и праздничных дней.

5.4 Расчет стоимости расходных материалов

В разделе стоимости расходных материалов учитываются расходы на приобретение основных материалов необходимых для проведения исследования, оформления соответствующей документации, а также учитывается стоимость картриджа. Расчет стоимости расходных материалов приведен в таблице 5.4.

Таблица 5.4 - Стоимость расходных материалов

Наименование расходных материалов	Цена за единицу, руб.	Количество, шт.	Сумма, руб.
Бумага	120	2	240
Канцтовары	160	1	160
Расходные материалы для принтера (картридж)	3500	1	3500
Ватман	10	10	100
Итого:			4000

Определили, что для проведения исследования затраты на приобретение расходных материалов потребуется 4000 рублей.

5.5 Расчет сметы расходов на исследование

С учетом часового тарифной ставки рассчитаем общие расходы на разработку и проведение исследования. В данную статью расходов включаются премиальные выплаты, районный коэффициент и страховые взносы. Для оценки затрат на исследование составляем смету на разработку и проведение исследования.

Произведем расчет расходов:

Премиальные выплаты рассчитываются по формуле:

$$ПВ = P_{OT} K_{ПВ} \quad (5.4)$$

где $K_{ПВ}$ - коэффициент премиальных выплат, составляет 20 %, в случае если премии не предусмотрены $K_{ПВ} = 1$.

$$ПВ = 32999,28 \cdot 0,2 = 6599,86 \text{ руб.}$$

Дополнительные затраты на проведение исследования можно определить,

Изм.	Лист	№ докум.	Подпись	Дата	Лист
					11070006.11.03.02.369.ПЗВКР 73

как:

$$З_{ДОП} = P_{OT} K \quad (5.5)$$

где К - коэффициент дополнительных затрат (К = 14%).

$$З_{ДОП} = 32999,28 \cdot 0,14 = 4619,90 \text{ руб.}$$

В заработной плате может быть предусмотрен районный коэффициент, которых характеризует доплату при работе в трудных условиях. Величина коэффициента определяется в зависимости от характера производства.

$$PK = P_{OT} K_{PB} \quad (5.6)$$

где K_{PB} - коэффициент районных выплат, для примера составляет 15 % от суммы.

$$PK = 32999,28 \cdot 0,14 = 4949,89 \text{ руб.}$$

Общие расходы на оплату труда вычисляются по формуле:

$$P_{общ} = P_{OT} + ПВ + PK + З_{ДОП} \quad (5.7)$$

где $З_{ДОП}$ - дополнительные затраты; ПВ - премиальные выплаты; РК - районный коэффициент; Р_{от} - основная заработка плата.

$$\Sigma P_{OT} = 32999,28 + 13199,72 + 4949,89 + 4619,90 = 55768,79 \text{ руб.}$$

Из таблицы 5.4 берется итоговая сумма стоимости расходных материалов по статье расходных материалов.

$$\Sigma P_{PM} = 4000 \text{ руб.}$$

Страховые взносы рассчитываются по формуле:

$$CB = P_{OT}^{0,3} \quad (5.8)$$

Изм.	Лист	№ докум.	Подпись	Дата

11070006.11.03.02.369.ПЗВКР

Лист

74

$$CB = 32999,28 \cdot 0,3 = 9899,78 \text{ руб.}$$

Амортизационные исчисления на использование компьютера вычисляются аналогично выражению (6). В данном примере они составляют 25% от стоимости компьютера.

$$AO = C_{PK}^{0,25} \quad (5.9)$$

$$AO = 30000 \cdot 0,25 = 7500 \text{ руб.}$$

Расходы на использование Интернета берутся из расчета месячной абонентской платы для предприятия. Пусть: Ринт = 1000 руб.

Административно-хозяйственные расходы составляют 50% от основной заработной платы (Рот).

$$P_{AX} = P_{OT}^{0,5} \quad (5.10)$$

$$P_{AX} = 32999,28 \cdot 0,5 = 16499,64 \text{ руб.}$$

Результаты расчета расходов сведем в таблицу. Смета расходов на разработку и проведение исследования представлена в таблице 5.5.

Таблица 5.5 - Смета расходов на разработку и проведение исследования

Наименование статей расходов	Сумма, руб.	Удельный вес статей, %
1	2	3
1. Стоимость расходных материалов	4000	4,23
2. Расходы на оплату труда	55768,79	
2.1. Основная заработная плата	32999,28	34,86
2.2. Дополнительные затраты	4619,90	4,88

Окончание таблицы 5.5

<i>1</i>	<i>2</i>	<i>3</i>
2.3. Премиальные выплаты	13199,72	13,94
2.4. Районный коэффициент	4949,89	5,23
3. Единый социальный налог	9899,78	10,46
4. Амортизационные исчисления на использование компьютера	7500	7,92
5. Расходы на использование Интернет	1000	1,06
6. Административно-хозяйственные расходы	16499,64	17,42
Итого:	94668,21	100

Результатом экономической оценки исследования является определение затрат на разработку и реализацию исследования:

- продолжительность исследовательских работ составила 55 дней;
- сметы расходов на исследование - 94668 рублей.

ЗАКЛЮЧЕНИЕ

В данной работе были исследованы амплитудный и спектральный методы оценивания основного тона речевого сигнала. Разработан алгоритм подготовки речевого сигнала к исследованию, который позволяет осуществлять селекцию вокализованных участков. Проведено сравнение амплитудного и спектрального методов. Определена частота основного тона речевого сигнала и проведена ее динамическая оценка. Динамическое изменение частоты основного тона диктора с использованием алгоритма оценивания основного тона больше точнее, чем в Adobe Audition. Воспроизведение синтезированных сигналов показало, что абсолютного сходства достигнуть не удалось, но синтезированный сигнал похож «на слух» на оригинал.

Цель работы достигнута, все задачи выполнены и получены следующие основные результаты:

1) исследование амплитудного и спектрального методов оценивания основного тона речевого сигнала показало их сравнимую близость. Различия в значениях частот незначительны и определяются разрешением частотной характеристики.

2) исследование динамического изменения основного тона речевого сигнала показало, что при естественном произнесении дикторами некоторых фраз, изменения частоты основного тона происходят в диапазоне порядка 70 - 90 Гц. Изменения частоты основного тона позволяют человеку произносить фразы с интонацией.

Результатом экономической оценки исследования является определение затрат на разработку и реализацию исследования:

- продолжительность исследовательских работ составила 55 дней;
- сметы расходов на исследование - 94668 рублей.

Изм.	Лист	№ докум.	Подпись	Дата
------	------	----------	---------	------

11070006.11.03.02.369.ПЗВКР

Лист

77

СПИСОК ИСПОЛЬЗОВАННЫХ ИСТОЧНИКОВ

1. Чистиков, П. Г. Методы и алгоритмы гибридного синтеза естественной русской речи на основе скрытых марковских моделей и метода Unit Selection [Текст] : дис. наук : 05.13.11 / П. Г. Чистиков. - Санкт-Петербург, 2013. - 134 с.
2. Синтез речи [Электронный ресурс] //ru.wikipedia.org/ Свободная энциклопедия Википедии. Режим доступа: https://ru.wikipedia.org/wiki/Синтез_речи (Дата обращения 09.04.18).
3. Siri [Электронный ресурс] //ru.wikipedia.org/ Свободная энциклопедия Википедии. Режим доступа - <https://ru.wikipedia.org/wiki/Siri> (Дата обращения 12.04.18).
4. Алиса (голосовой помощник) [Электронный ресурс] //ru.wikipedia.org/ Свободная энциклопедия Википедии. Режим доступа: [https://ru.wikipedia.org/wiki/Алиса_\(голосовой_помощник\)](https://ru.wikipedia.org/wiki/Алиса_(голосовой_помощник)) (Дата обращения 12.04.18).
5. Субтрактивный синтез [Электронный ресурс] //synthmusic.ru/ Студия звукозаписи SynthMusic. Режим доступа: <http://synthmusic.ru/articles/synthesistypes/subtractive/> (Дата обращения 17.04.18).
6. Формантный синтез [Электронный ресурс] //synthmusic.ru/ Студия звукозаписи SynthMusic. Режим доступа: <http://synthmusic.ru/articles/synthesistypes/formant/> (Дата обращения 17.04.18).
7. Частотная модуляция [Электронный ресурс] //synthmusic.ru/ Студия звукозаписи SynthMusic. Режим доступа: <http://synthmusic.ru/articles/synthesistypes/fm/> (Дата обращения 17.04.18).
8. Аналоговое моделирование [Электронный ресурс] //synthmusic.ru/ Студия звукозаписи SynthMusic. Режим доступа: <http://synthmusic.ru/articles/synthesistypes/analogmodeling/> (Дата обращения 17.04.18).

Изм.	Лист	№ докум.	Подпись	Дата
------	------	----------	---------	------

11070006.11.03.02.369.ПЗВКР

Лист

78

9. Векторный синтез [Электронный ресурс] //synthmusic.ru/ Студия звукозаписи SynthMusic. Режим доступа: <http://synthmusic.ru/articles/synthesistypes/vector/> (Дата обращения 19.04.18).

10. Линейно-арифметический синтез [Электронный ресурс] //synthmusic.ru/ Студия звукозаписи SynthMusic. Режим доступа: <http://synthmusic.ru/articles/synthesistypes/la/> (Дата обращения 22.04.18).

11. Ха Куанг Нинь, Прохоренко, Е. И. Выбор длины отрезка речевого сигнала при кратковременном анализе звуков речи [Текст] / Ха Куанг Нинь, Е. И. Прохоренко // Молодежный научный форум: Технические и математические науки. - М.: Общество с ограниченной ответственностью "Международный центр науки и образования", 11 (40), 2016. - с. 205-211.

12. Формантный синтез [Электронный ресурс] //www.ie.tusur.ru/ Электронные средства сбора, обработки и отображения информации. Режим доступа: http://www.ie.tusur.ru/books/COI/page_43.htm (Дата обращения 25.04.18).

13. Гапочкин, А. В. Определение основного тона речи с помощью вейвлет-преобразования и его применение [Текст] / А. В. Гапочкин // ISSN ONLINE: 2409-6652.

14. Голубинский, А. Н. Расчёт частоты основного тона речевого сигнала на основе полигармонической математической модели [Текст] / А.Н. Голубинский.

15. Болдышев, А.В. Методические рекомендации по выполнению технико-экономического обоснования выпускных квалификационных работ [Текст] / А.В. Болдышев - Белгород: Изд-во БелГУ, 2013. - 22 с.

Изм.	Лист	№ докум.	Подпись	Дата
------	------	----------	---------	------

11070006.11.03.02.369.ПЗВКР

Лист

79

Выпускная квалификационная работа выполнена мной совершенно самостоятельно. Все использованные в работе материалы и концепции из опубликованной научной литературы и других источников имеют ссылки на них.

«___» _____. ____ г.

(подпись)

(Ф.И.О.)