

**О. Е. Елисеева**

# **Речевой интерфейс. Лабораторный практикум**

В 2-х частях

Часть 1

*Рекомендовано Учебно-методическим объединением вузов Республики Беларусь по образованию в области информатики и радиоэлектроники в качестве учебно-методического пособия для студентов учреждений, обеспечивающих получение высшего образования по специальности «Искусственный интеллект»*

Под редакцией профессора В. В. Голенкова

Минск БГУИР 2008

УДК 004.5 (075.8)  
ББК 32.973 я 7  
Е 51

Р е ц е н з е н т

доцент кафедры прикладной лингвистики БГУ, кандидат филологических наук А. И. Головня

**Елисеева, О. Е.**  
Е 51 Речевой интерфейс. Лабораторный практикум : учеб.-метод. пособие; под ред. проф. В. В. Голенкова : В 2 ч. Ч. 1 / О. Е. Елисеева. – Минск : БГУИР, 2008. – 44 с.  
ISBN 978-985-488-242-0 (ч. 1)

Настоящее пособие представляет собой описание методических рекомендаций к выполнению лабораторных работ по курсу «Речевой интерфейс». Первая часть пособия включает лабораторные работы, связанные с исследованием закономерностей русской речи и практическим изучением основных характеристик речевого сигнала. В пособие включены теоретические материалы, знакомящие студентов с основами работы со звуком в Microsoft Windows, характеристиками речевого сигнала, основами работы с системами обработки звука и речи. Содержание пособия соответствует типовой программе учебного курса «Речевой интерфейс» по специальности «Искусственный интеллект» (№ ТД-40-62/тип, 2003 г.).

Предназначено для студентов вузов, обучающихся по специальности «Искусственный интеллект» всех форм обучения, а также родственным специальностям. Может быть полезно специалистам в области разработки интеллектуальных интерфейсов и прикладных компьютерных систем различного назначения.

УДК 004.5 (075.8)  
ББК 32.973 я 7

ISBN 978-985-488-242-0 (ч. 1)

ISBN 978-985-488-247-5

© Елисеева О. Е., 2008

© УО «Белорусский государственный университет информатики и радиоэлектроники», 2008

## Содержание

ПРЕДИСЛОВИЕ .....	4
ВВЕДЕНИЕ.....	5
I. ОСНОВЫ РАБОТЫ СО ЗВУКОМ И РЕЧЬЮ В WINDOWS.....	6
А. КОНФИГУРАЦИЯ ЗВУКОВОГО УСТРОЙСТВА .....	6
Б. ВСТРОЕННЫЕ ПРИЛОЖЕНИЯ WINDOWS ДЛЯ РАБОТЫ СО ЗВУКОМ.....	9
<i>Регулятор громкости</i> .....	9
<i>Стандартное средство звукозаписи</i> .....	11
<i>Стандартные программы воспроизведения звука с компакт-диска</i> .....	12
В. УСТРАНЕНИЕ НЕПОЛАДОК ПРИ РАБОТЕ СО ЗВУКОМ .....	13
II. ОПИСАНИЕ РЕЧЕВОГО СИГНАЛА.....	14
ЛАБОРАТОРНАЯ РАБОТА №1 «ИЗУЧЕНИЕ ФОНЕТИЧЕСКИХ ЗАКОНОМЕРНОСТЕЙ РЕЧЕВОГО СИГНАЛА».....	16
1.1. ОБЩИЕ СВЕДЕНИЯ О ПРОГРАММЕ SOUND FORGE .....	16
1.2. ФОНЕТИЧЕСКИЙ СОСТАВ РУССКОЙ РЕЧИ .....	20
1.3. ПОРЯДОК ВЫПОЛНЕНИЯ ЛАБОРАТОРНОЙ РАБОТЫ №1 .....	21
1.4. ВОПРОСЫ ДЛЯ САМОПРОВЕРКИ .....	24
ЛАБОРАТОРНАЯ РАБОТА №2 «ИЗУЧЕНИЕ ПРОСОДИЧЕСКИХ ЗАКОНОМЕРНОСТЕЙ РЕЧЕВОГО СИГНАЛА».....	25
2.1. ПРОСОДИЧЕСКИЕ ХАРАКТЕРИСТИКИ РЕЧЕВОГО СИГНАЛА .....	25
2.2. ПОРЯДОК ВЫПОЛНЕНИЯ ЛАБОРАТОРНОЙ РАБОТЫ №2.....	26
2.3. ВОПРОСЫ ДЛЯ САМОПРОВЕРКИ .....	28
ЛАБОРАТОРНАЯ РАБОТА №3 «ИССЛЕДОВАНИЕ КОМПИЛЯЦИОННЫХ МЕТОДОВ СИНТЕЗА РЕЧЕВОГО СИГНАЛА» .....	29
3.1. МЕТОДЫ СИНТЕЗА РЕЧИ .....	29
3.2. ПРАВИЛА ТРАНСКРИБИРОВАНИЯ .....	30
3.3. ПОРЯДОК ВЫПОЛНЕНИЯ ЛАБОРАТОРНОЙ РАБОТЫ №3.....	31
3.4. ВОПРОСЫ ДЛЯ САМОПРОВЕРКИ .....	32
ЛАБОРАТОРНАЯ РАБОТА №4 «ПРАКТИЧЕСКОЕ ИЗУЧЕНИЕ ФОРМАНТНО-ПАРАМЕТРИЧЕСКОГО ОПИСАНИЯ РЕЧЕВОГО СИГНАЛА».....	33
4.1. ОПИСАНИЕ ПРОГРАММЫ PRAAT ДЛЯ ИССЛЕДОВАНИЯ СПЕКТРАЛЬНЫХ ХАРАКТЕРИСТИК РЕЧЕВОГО СИГНАЛА .....	33
4.2. ФОРМАНТНАЯ МОДЕЛЬ РЕЧЕВОГО СИГНАЛА .....	36
4.3. ПОРЯДОК ВЫПОЛНЕНИЯ ЛАБОРАТОРНОЙ РАБОТЫ №4.....	38
<i>Лабораторная работа №4. Часть 1</i> .....	38
<i>Лабораторная работа №4. Часть 2</i> .....	42
4.4. ВОПРОСЫ ДЛЯ САМОПРОВЕРКИ .....	44
ЛИТЕРАТУРА.....	44

## Предисловие

Данное учебно-методическое пособие является результатом обобщения многолетнего опыта проведения лабораторного практикума в рамках дисциплины «Речевой интерфейс» для студентов специальности «Искусственный интеллект» Учреждения образования «Белорусский государственный университет информатики и радиоэлектроники» при участии и поддержке специалистов Лаборатории распознавания и синтеза речи Объединенного института проблем информатики (ОИПИ) НАН Беларуси.

Для более эффективного изучения вопросов, связанных с разработкой речевых интерфейсов прикладных интеллектуальных систем, рекомендуется ознакомиться с теоретическим материалом, представленном в пособии [1] (*Лобанов Б.М..2006кн-Речев\_И\_ИС*). Там же подробно представлены цели и задачи дисциплины «Речевой интерфейс», ее место в учебном процессе.

Данное пособие предназначено для студентов вузов, обучающихся по специальности «Искусственный интеллект», родственным специальностям всех форм обучения, а также для специалистов в области вычислительной техники, желающих на практике получить представление о проблемах, связанных с разработкой речевых интерфейсов интеллектуальных систем.

Автор выражает искреннюю признательность рецензенту, всем сотрудникам Лаборатории распознавания и синтеза речи Объединенного института проблем информатики НАН Беларуси за помощь в постановке учебного курса и за предоставленные материалы, а также студентам специальности «Искусственный интеллект», которые принимали участие в сборе и подготовке материалов для данного пособия. Особую благодарность хочется выразить сотрудникам ОИПИ НАН Беларуси Б. М. Лобанову, Н. П. Дегтяреву, Г. В. Лосику, А. С. Рылову, Т. В. Левковской Л. И. Цирульнику за неоценимый вклад в развитие учебного курса «Речевой интерфейс», а также рецензенту: доценту кафедры прикладной лингвистики БГУ, канд. филол. наук А. И. Головня – за объективную критику и ценные замечания, позволившие существенно улучшить качество данного пособия.

## Введение

Современный мир немислим без использования компьютерной техники практически в любой сфере деятельности человека. Компьютеры используются везде, и каждый современный человек просто обязан знать, что такое компьютер и как с ним работать. Одной из самых естественных для человека форм взаимодействия является **речь**. Поэтому большую актуальность в настоящее время приобретает разработка **речевых интерфейсов** интеллектуальных систем. Это подтверждает и тот факт, что как современные компьютеры, так и соответствующие операционные системы снабжены практически всеми необходимыми средствами работы со звуком. Кроме того, в операционной системе Microsoft Windows XP и более новых версиях Windows XP имеются также встроенные возможности для работы с речью. Несмотря на это существующие в настоящее время разработки пока не удовлетворяют всем требованиям пользователей, так как синтезируемая компьютером речь не является достаточно естественной, а средства распознавания нуждаются в настройке (обучении) на каждого пользователя, т.е. не обладают достаточной степенью универсальности.

Разработка речевого интерфейса интеллектуальной системы сопряжена с решением огромного количества разнородных проблем [1] (*Лобанов Б.М., 2006 кн-Речев\_И\_ИС*), которые условно можно разделить на два основных класса:

- исследование закономерностей и характеристик речевого сигнала;
- программная реализация компонентов речевого интерфейса.

В связи с этим лабораторные работы, выполнение которых рекомендуется студентам, подразделяются на два аналогичных класса. В 1-й части учебно-методического пособия описаны лабораторные работы, основной **целью** которых является практическое изучение студентами закономерностей образования речи и характеристик речевого сигнала.

Все рассмотренные в пособии лабораторные работы выполняются в операционной среде Windows на персональном компьютере IBM PC типа Pentium, имеющем стандартный набор аудиоустройств: аудиокарту, динамики и микрофон (или телефонную гарнитуру).

# I. Основы работы со звуком и речью в Windows

Прежде чем приступить к выполнению лабораторного практикума, рекомендуется ознакомиться с основами работы со звуком и речью в *Microsoft Windows*.

Операционная система *Windows* содержит встроенную программную поддержку для широкого диапазона аппаратуры мультимедиа. Для работы с устройствами мультимедиа и специальными типами мультимедиа-файлов в *Windows* устанавливаются соответствующие драйверы устройств. Для проигрывания, например, звукового компакт-диска установлены драйвер звукового компакт-диска и драйвер накопителя на компакт-диске. Чаще всего операционная система *Windows* в процессе своей инсталляции просматривает установленное оборудование и автоматически устанавливает необходимые драйверы, если оборудование ей «знакомо» (так называемая система *PLUG AND PLAY*).

## A. Конфигурация звукового устройства

Для того чтобы иметь возможность работать со звуком, необходимо правильно настроить конфигурацию звукового устройства. В процессе установки операционной системы *Windows* мастер установки дает возможность пользователю настроить систему в диалоговом режиме. При этом пользователю не приходится заботиться о том, какие драйверы устанавливать и как их конфигурировать. Мастер установки делает все сам. Исключения составляют те случаи, когда звуковое устройство не поддерживается стандартным набором драйверов, поставляемых с *Windows*. В этом случае пользователь должен сам установить и настроить необходимые драйверы. Кроме того, в процессе работы пользователю может понадобиться изменять некоторые настройки звукового устройства. Все это можно сделать с помощью соответствующих средств конфигурации звукового устройства, входящих в поставку *Windows*.

Просмотреть или изменить текущие настройки звукового устройства можно с помощью диалога «Звуки и аудиоустройства» из панели управления *Windows* (рис. 1–5). Диалог состоит из закладок, позволяющих подобрать звуки *Windows* (рис. 1), которые будут сопровождать различные события, установить громкость воспроизведения звука (рис. 2), настроить звуковое оборудование (рис. 3), конфигурацию динамиков (рис. 4), проверить звучание каждого из них в отдельности, выбрать и настроить аудио-устройства (рис. 5), а также средства речевого интерфейса *Windows* (рис. 6).

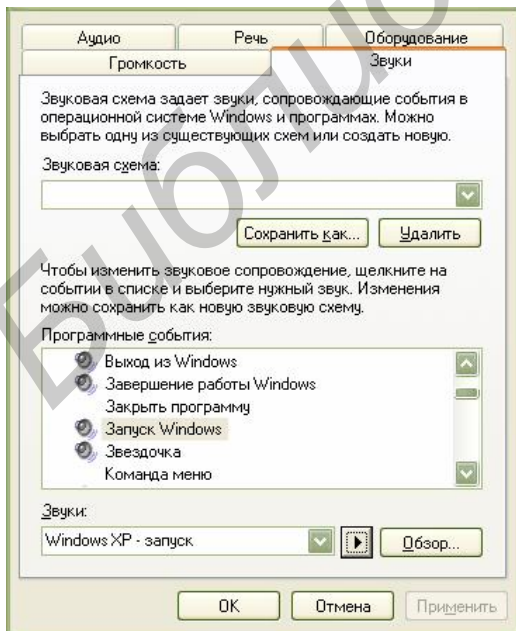
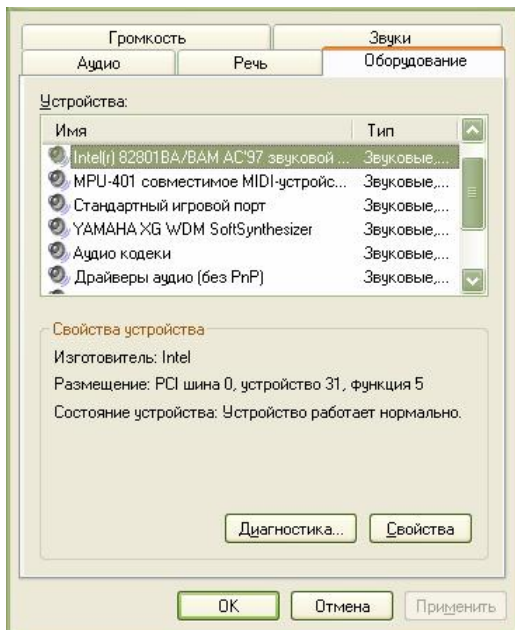


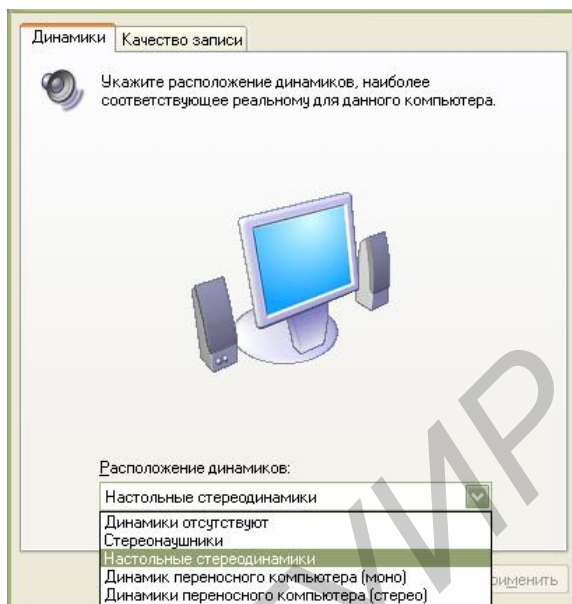
Рисунок 1. Подбор звуков в Windows



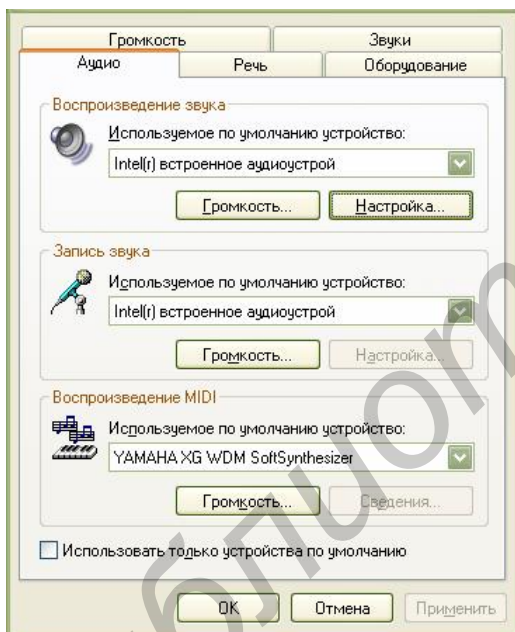
Рисунок 2. Настройка громкости воспроизведения звука



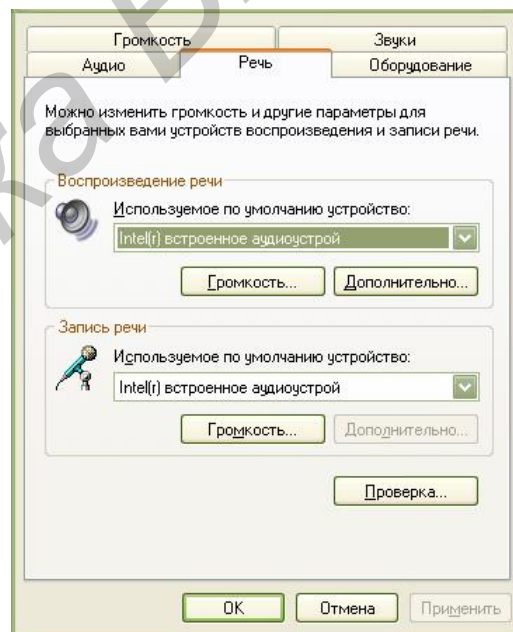
**Рисунок 3.** Выбор и настройка оборудования



**Рисунок 4.** Настройка динамиков



**Рисунок 5.** Настройка аудиоустройств



**Рисунок 6.** Настройка средств речевого интерфейса Windows

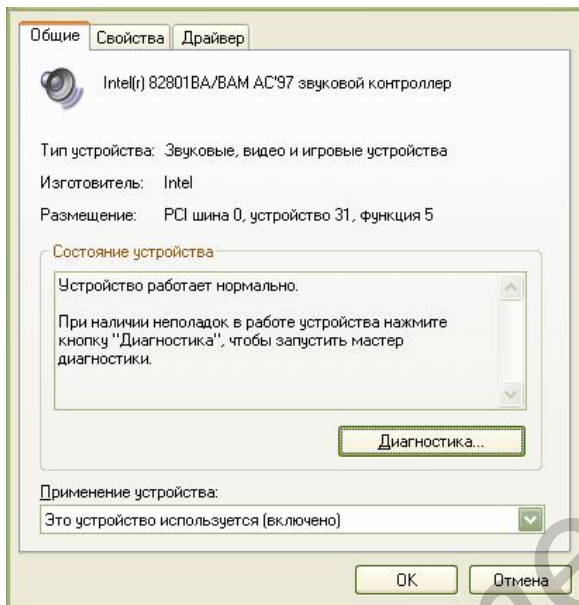
Рассмотрим закладки диалога «Звуки и аудиоустройства» более подробно.

Закладка «Звуки» (см. рис. 1) предоставляет пользователю широкие возможности для конфигурирования мультимедиа составляющей операционной системы. В первую очередь, пользователь может создавать свою собственную схему звуков, которая будет использоваться в процессе его работы с приложениями *Windows*. Благодаря звуковой схеме, работа в *Windows* сопровождается звуковыми сигналами разного типа, которые поставлены в соответствие разным системным событиям.

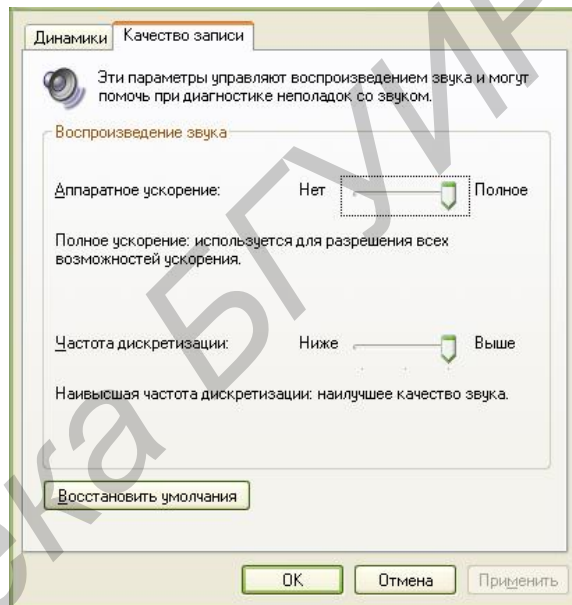
Для того чтобы определить звуковую схему, необходимо выбрать в раскрывающемся списке схем имя нужной схемы (или сохранить какую-нибудь уже существующую схему под новым

именем с помощью кнопки «*Save as...*»), а затем, при необходимости, приступить к редактированию списка соответствия звуков системным событиям. Прослушать полученный результат можно тут же при помощи находящейся поблизости кнопки воспроизведения.

На закладке «*Оборудование*» (см. рис. 3) выводится список всех мультимедийных устройств, установленных в системе, а также их кодеков и драйверов. Свойствами каждого из этих компонентов можно управлять с помощью кнопки *Свойства* (рис. 7). Выбрав, например, контроллер звукового устройства и просмотрев его свойства, можно обнаружить дополнительные настройки. Это в основном полная информация о типе устройства, его производителе и месте производства; о драйвере к этому устройству; об аудио- и миди-устройствах системы. Кроме того, можно обновить или просто переустановить драйвер, назначить использование дополнительных свойств, осуществить диагностику состояния устройства.



**Рисунок 7.** Свойства звукового устройства



**Рисунок 8.** Настройка качества воспроизведения

На закладке «*Аудио*» (см. рис. 5) пользователь может назначить различные аудио-устройства (в случае, если их у него несколько) для обработки различных видов сигналов: воспроизведения аудиофайлов; записи звука; воспроизведения MIDI-файлов.

Для каждого из этих устройств можно подобрать свои специфические настройки, начиная от громкости и частоты дискретизации (рис. 8) и заканчивая настройкой аппаратного ускорения и выбором типа и количества динамиков (см. рис. 4).

На закладке «*Речь*» (см. рис. 6) пользователю предоставляется возможность настроить громкость и другие параметры для устройств записи и воспроизведения речи. Кроме того, как сказано выше, в последних версиях *Microsoft Windows* (в частности, в *Windows XP*) имеются встроенные средства синтеза речи. В связи с этим в *Windows* имеются дополнительные возможности настройки свойств речи (рис. 9), где пользователю предоставляется возможность выбрать голос, который будет использоваться для синтеза речи, прослушать тестовую синтезированную выбранным голосом фразу, настроить параметры синтеза речи: скорость (темп), громкость, а также выбрать и настроить аудио-устройство, которое будет использоваться для воспроизведения речи.

Описанные речевые возможности *Windows* пока очень ограничены. В частности, для более качественного синтеза речи необходимо устанавливать дополнительные голоса (голосовые базы данных). При этом для синтеза английской речи и русской речи используются различные голоса. Для полноценного использования речевых возможностей необходимо



устанавливать специальные программы и утилиты для синтеза и распознавания речи, к которым, в частности, относятся всевозможные экранные чтецы и средства голосового управления типа «голосовая мышь».

Заметим также, что в специальных возможностях операционной системы *Windows* (рис. 10) имеется возможность выбрать и настроить средства визуального сопровождения звуковых сигналов, а также включить отображение субтитров для речевых сообщений.

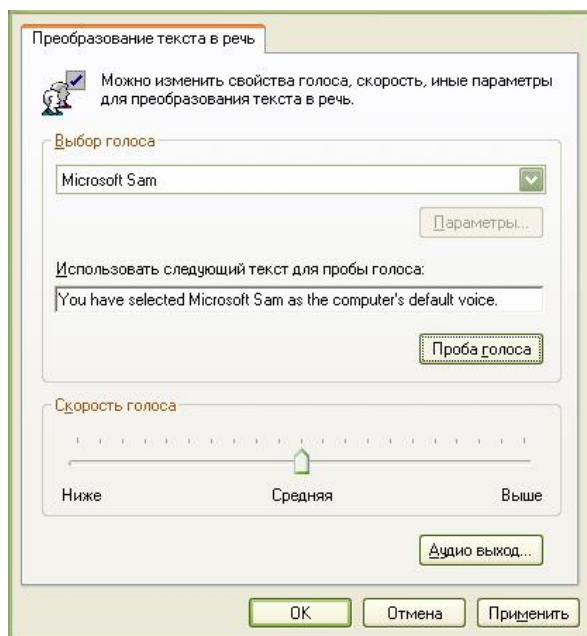


Рисунок 9. Настройка свойств речи в Windows XP

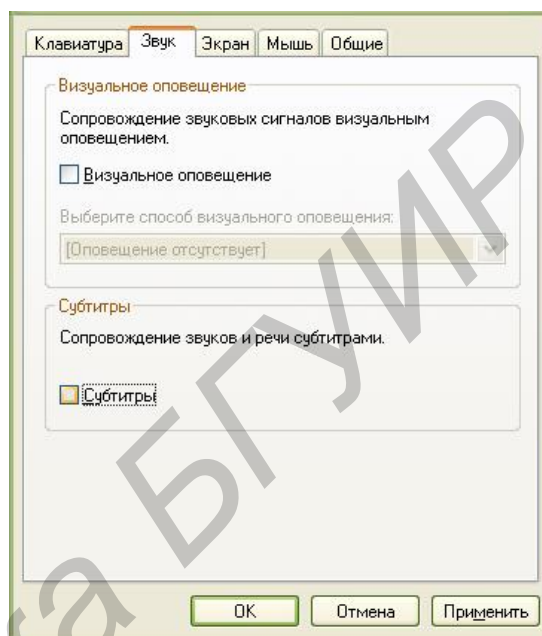


Рисунок 10. Специальные возможности Windows

## Б. Встроенные приложения Windows для работы со звуком

### Регулятор громкости

Регулятор громкости (рис. 11) предназначен для управления громкостью и балансом при воспроизведении звукозаписей на компьютере.

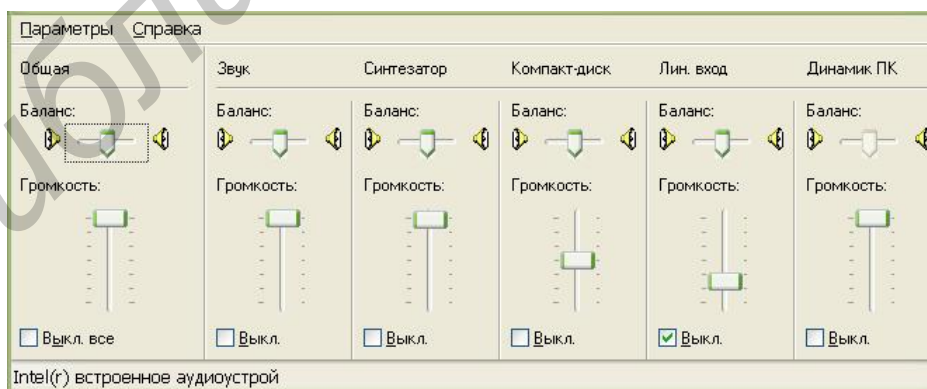



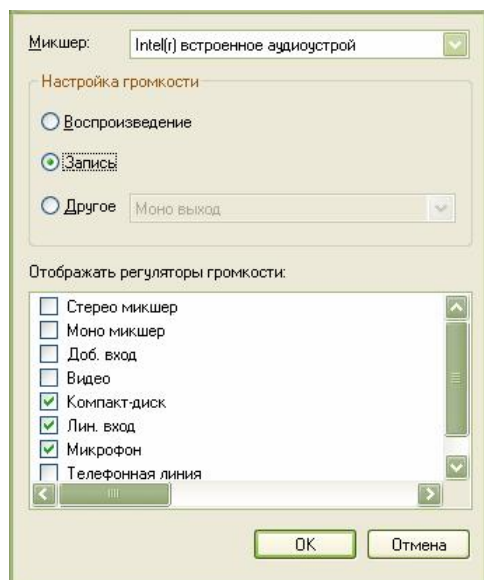
Рисунок 11. Регулятор громкости

Регулятор громкости может быть запущен различными способами:

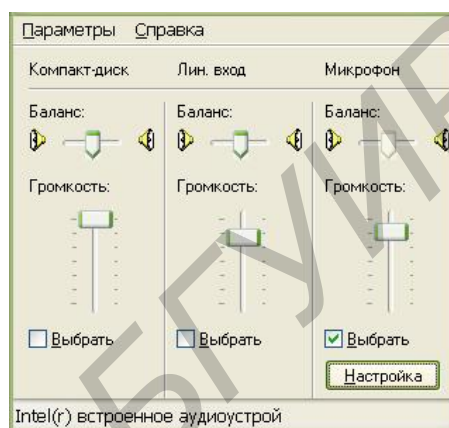
- из диалога «Звуки и аудиоустройства»;
- из группы программ *Стандартные* ® *Развлечения* ® *Громкость*;

- из панели задач *Windows*, используя ярлык для быстрого запуска регулятора громкости, значок которого имеет вид динамика .

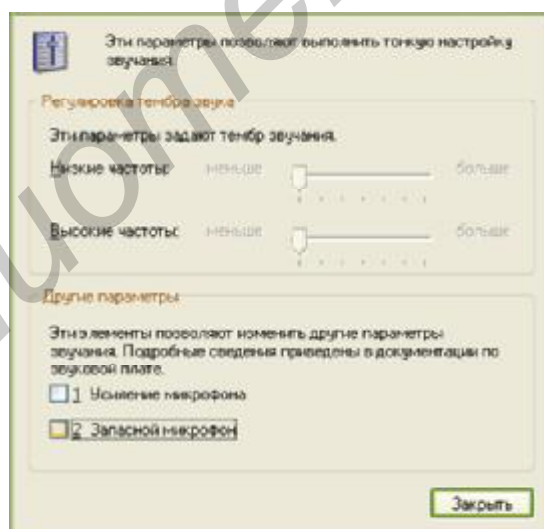
При настройке уровня воспроизведения необходимо перетащить ползунок регулятора *Громкость* вверх, чтобы увеличить громкость, и вниз, чтобы ее уменьшить. При перетаскивании ползунка регулятора *Баланс* происходит изменение баланса между правым и левым каналами динамиков.



**Рисунок 12.** Выбор настроек для записи звука



**Рисунок 13.** Настройка уровня записи



**Рисунок 14.** Дополнительные параметры микрофона

Если компьютер имеет несколько устройств (например MIDI или звукозаписи), для каждого устройства можно задать свою громкость. Для этого в меню *Параметры* необходимо выбрать команду *Свойства* (рис. 12), затем выбрать один из вариантов *Воспроизведение*, *Запись* (настройка уровня записи) или *Другое* и убедиться, что флажок рядом с данным устройством установлен. Для отключения воспроизведения звука под регулятором *Громкость* необходимо установить флажок *Выкл.*

Как показано на рис. 12 и 13, из окна регулятора громкости можно настроить уровень записи звука. Кроме того, для микрофона можно задать дополнительные настройки (рис. 14), используя кнопку *Настройка* в диалоге *Уровень записи* (см. рис. 13).

Для работы со звуком существует достаточно большое количество различных приложений. Наиболее развитые из них являются разработками различных производителей программного обеспечения. Например, программа *Sonic Foundry Sound Forge* (см. подразд. 1.1) содержит практически все необходимые средства для записи, воспроизведения и обработки звука на компьютере. Существуют также специализированные средства для работы, например, с музыкой, с помощью которых компьютер превращается в домашнюю музыкальную студию. Однако, даже не имея таких развитых средств, любой пользователь Windows имеет в своем распоряжении минимальный набор программ, обеспечивающих все необходимые возможности работы со звуком, к которым относятся запись и воспроизведение. Ниже кратко рассмотрены такие стандартные средства.

### Стандартное средство звукозаписи

Программа «*Звукозапись*» (*Sound Recorder*) (рис. 15) служит для записи, наложения, воспроизведения и редактирования звукозаписей. Кроме того, звукозаписи можно связывать с другими документами или вставлять в них.

Работа с программой «*Звукозапись*» требует наличия звуковой платы и подключенных к компьютеру динамиков. Для записи живого звука требуется также микрофон. Звукозапись можно воспроизвести как с помощью программы звукозаписи, так и с помощью универсального проигрывателя. Программа для записи звука запускается из группы программ *Стандартные ® Развлечения ® Звукозапись*.



Рисунок 15. Основной диалог программы «Звукозапись»

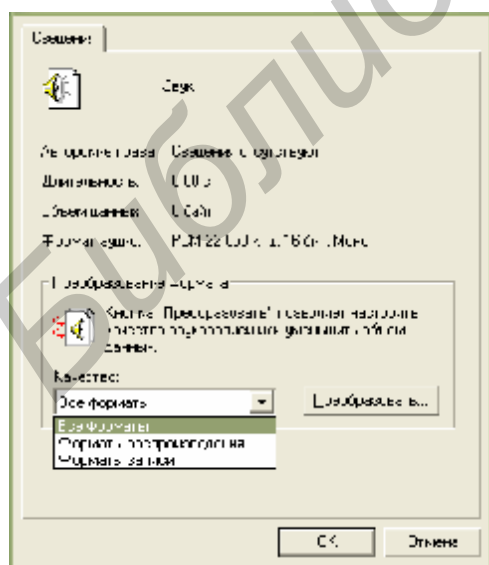


Рисунок 16. Свойства звукозаписи

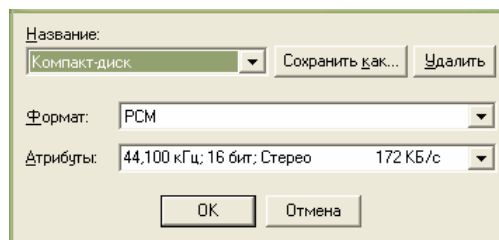
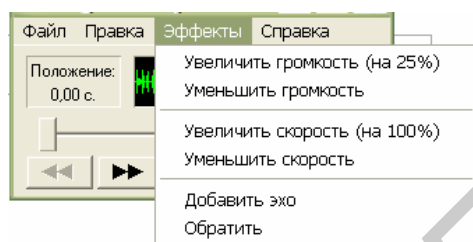


Рисунок 17. Настройка качества звукозаписи

В данной программе пользователь может записать любой звук, используя микрофон или линейный вход звуковой карты. Кроме того, в ней предусмотрены следующие дополнительные возможности: изменение скорости записи; регулирование громкости; настройка качества записи; определение формата звукозаписи; воспроизведение звукозаписи в обратном порядке; добавление эффекта эха.

Следует, однако, заметить, что эти изменения возможны только в том случае, если файл звукозаписи не сжат. Если в главном окне не отображается осциллограмма записи (зеленого цвета), это означает, что файл сжат, и поэтому применять специальные эффекты можно только после изменения качества звукозаписи. Для изменения качества звукозаписи в меню *Файл* необходимо выбрать команду *Свойства*, далее выбрать формат в группе *Преобразование формата* (рис. 16), нажать кнопку *Преобразовать* и указать формат и атрибуты звукозаписи (рис. 17). Далее для применения какого-либо преобразования звукозаписи необходимо использовать пункт меню *Эффекты* (рис. 18).



**Рисунок 18.** Выбор эффектов для преобразования звукозаписи

С помощью программы «*Звукозапись*» можно также осуществлять: вставку звука в существующий файл звукозаписи, причем записываемый звук заменяет существующий, начиная с выбранной позиции; вставку звукозаписи в документ; связывание файла звукозаписи с документом. При этом связанные данные автоматически обновляются при их изменении в исходном документе (если связывание не поддерживается, команда *Специальная вставка* отсутствует в меню *Правка*).

## Стандартные программы воспроизведения звука с компакт-диска

Одним из стандартных средств воспроизведения звука в операционной системе Windows является средство для воспроизведения компакт дисков *CD Player*. Как стандартный плеер, данное средство позволяет осуществить все возможные действия над аудио-файлами: воспроизводить, останавливать, прерывать исполнение, перемещаться вдоль каждого мультимедиа-трека, выбирать нужный трек из списка имеющихся и т.д. Программа *CD Player* загружается автоматически при установке любого аудио-диска в дисковод компьютера. Возможности данной программы позволяют пользователю посетить некоторые музыкальные сайты и скачать понравившиеся мелодии, не прибегая к помощи дополнительных Интернет-средств.

Следует отметить, что интерфейс практически всех средств работы со звуком на компьютере несколько напоминает набор средств управления обычным бытовым устройством для работы со звуком, например домашним магнитофоном. Поэтому набор кнопок на панели инструментов программного приложения оказывается знакомым любому пользователю. Отличительной особенностью программных средств работы со звуком является расширенный набор настроек. Например, диалог «*Настройки*» программы *CD Player* позволяет задавать опции самого плеера, устанавливать настройки каждого альбома и списков воспроизведения.

Еще одним поставляемым средством вместе с *Windows* для воспроизведения мультимедиа-файлов является проигрыватель *Windows Media*. Проигрывание музыкальных компакт-дисков при помощи проигрывателя *Windows Media* выполняется

так же просто, как и при использовании обычного проигрывателя для компакт-дисков *CD Player*. Достаточно просто вставить компакт-диск в дисковод. Проигрыватель *Windows Media* будет автоматически загружен, выведен на экран, а затем начнется воспроизведение первой записи на диске.

## В. Устранение неполадок при работе со звуком

В ходе работы с мультимедийной информацией могут возникнуть различные виды неполадок (рис. 19). Помощь в устранении неполадок при работе со звуком можно получить, используя *Центр справки и поддержки Windows*. После выбора обнаруженной неполадки пользователь может использовать *Мастер устранения неполадок* (рис. 20) и в диалоговом режиме попытаться исправить возникшую ошибку.



Рисунок 19. Неполадки при работе со звуком

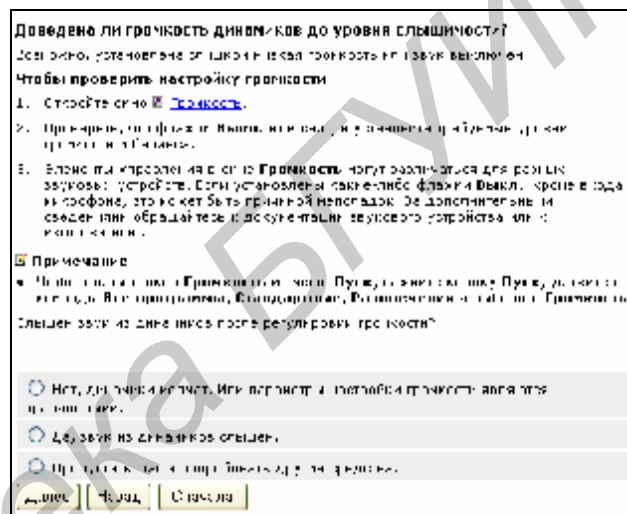


Рисунок 20. Устранение неполадок при работе со звуком

## II. Описание речевого сигнала

**Речь** – это сложное физическое явление. Она образуется в результате прохождения воздушных потоков через речевой тракт человека. С акустической точки зрения **звук** – это колебательные движения среды. Сложность речи заключается в том, что это не один колебательный процесс. Речевой сигнал состоит из множества **гармонических составляющих**, т.е. колебаний, которые периодически изменяются во времени по синусоидальным законам (рис. 21) и описываются следующим уравнением [1] (Лобанов Б.М..2006кн-Речев\_И\_ИС):

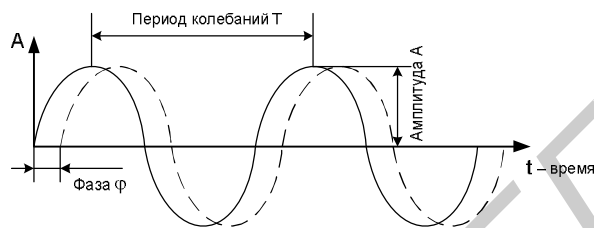
$$S(t) = A \cdot \cos(F \cdot t + \varphi),$$

где  $A$  – амплитуда колебаний;

$F$  – частота ( $F = 1 / T$ );

$t$  – время;

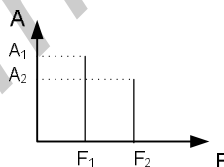
$\varphi$  – фаза.



**Рисунок 21.** Синусоидальное гармоническое колебание

Речевой сигнал, как и любой сложный сигнал, является результатом сложения нескольких гармонических составляющих, каждая из которых характеризуется своими значениями указанных параметров. Каждая такая составляющая называется **гармоникой**.

В результате разложения речевого сигнала на входящие в его состав гармонические колебания получается частотный **спектр** – амплитудно-частотная зависимость. Например, если для сигнала частота и амплитуда первой гармоники равны  $F_1$  и  $A_1$  соответственно, а второй –  $F_2$  и  $A_2$ , то получим амплитудно-частотную зависимость, представленную на рис. 22.



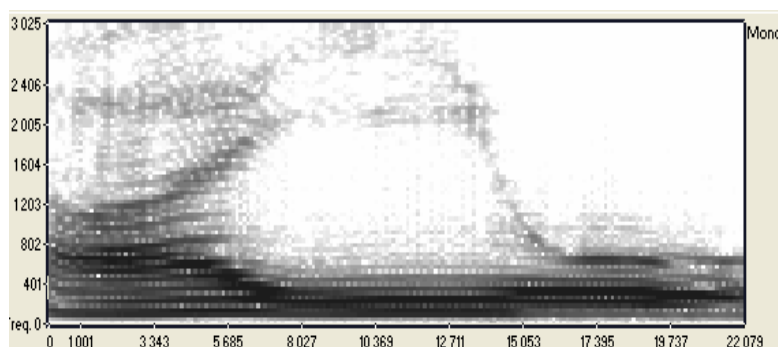
**Рисунок 22.** Представление сигнала в виде частотного спектра

Для более сложного сигнала амплитудно-частотная зависимость будет включать гораздо большее количество вертикальных отрезков с разными значениями частоты и амплитуды. В этом случае используется другое, более наглядное представление указанной зависимости в виде огибающей кривой линии.

Спектр речевого сигнала выражает его частотный состав, т.е. самыми главными характеристиками речевого сигнала являются значения частот его гармонических составляющих. Именно в результате наложения этих частот и их восприятия органами слуха мы получаем возможность слышать богато украшенный звук – речевой сигнал. Для разложения речевого сигнала в спектр, как правило, используется преобразование Фурье [2, 3, 4, 5] (Рылов А.С.2003кн-Анализ\_Р; Рабинер Л.Р..1981кн-Цифро\_О\_Р\_С; Фант Г.1964кн-Акуст\_Т\_Р; Радзишевский А.2006кн-Основ\_А\_и\_Ц\_3). На практике такое разложение речевого сигнала представляется в виде так называемых **динамических спектрограмм** (сонограмм).

**Сонограмма** – это трехмерное изображение динамики артикуляторных процессов в речевом тракте. По оси абсцисс откладывается время, по оси ординат – значение частот, а степень

их зачернения соответствует энергии (амплитуде) частотных компонент в спектре сигнала (рис. 23).

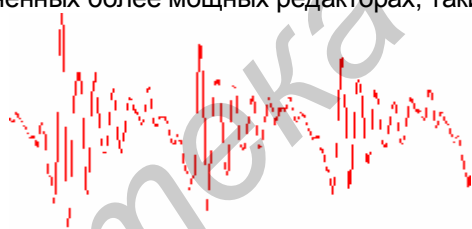


**Рисунок 23.** Трехмерное изображение динамической спектрограммы речевого сигнала

Благодаря описанным колебательным процессам происходит образование речевого сигнала. Для изучения речевого сигнала необходимо его представить в виде набора некоторых физических и математических составляющих, которые можно изобразить в виде графиков соответствующих функций. Основными видами изображения (описания) речевого сигнала являются:

- **спектр** – разложение амплитудно-частотных зависимостей;
- **осциллограмма** – амплитудно-временная зависимость.

Представление речевого сигнала в виде осциллограммы можно увидеть в любом редакторе звука, например, с помощью любых встроенных средств записи звука на компьютере, а также в широко распространенных более мощных редакторах, таких как Sound Forge (рис. 24).



**Рисунок 24.** Фрагмент осциллограммы фонемы [а]  
(спектр фонемы представлен на рис. 23)

С точки зрения лингвистики в речевом сигнале отражаются фонетические и просодические аспекты речи, с точки зрения акустики – акустические характеристики фонем и просодем, а информационная структура содержит дополнительную информацию о речевом сигнале, включающую его смысловое содержание, а также индивидуальные характеристики личности говорящего и передающей среды.

Для реализации речевого интерфейса необходимо предварительно тщательно изучить характеристики речевого сигнала и выявить закономерности. Полученные результаты в дальнейшем используются для моделирования (синтеза) речи, а также способствуют более качественному его распознаванию. Более подробно характеристики речевого сигнала описаны в [1] (Лобанов Б.М..2006кн-Речев\_И\_ИС).

# Лабораторная работа №1

## «Изучение фонетических закономерностей речевого сигнала»

**Цель работы** – изучение фонетических закономерностей речевого сигнала.

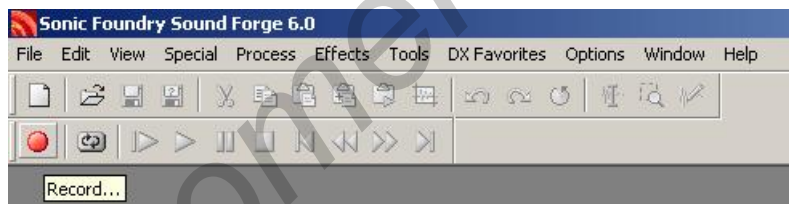
Основные **задачи** работы следующие:

1. Закрепление основных понятий и характеристик речевого сигнала.
2. Приобретение навыков работы с программой *Sonic Foundry Sound Forge* для изучения речевого сигнала.
3. Практическое исследование фонетических закономерностей речевого сигнала.

### 1.1. Общие сведения о программе Sound Forge

Для изучения фонетических закономерностей речевого сигнала предлагается использовать программу *Sonic Foundry Sound Forge* (<http://www.sonicfoundry.com>) (рис. 1.1), которая имеет развитое меню, наглядную панель инструментов и удобную для работы систему окон. Программа *Sound Forge* позволяет:

- записывать с микрофона речевой сигнал и воспроизводить его осциллограмму;
- сохранять записанный речевой сигнал в файл;
- воспроизводить осциллограмму (или ее часть) ранее записанного речевого сигнала;
- изменять при необходимости масштабы просмотра осциллограммы;
- вырезать ненужные участки речевого сигнала;
- вырезать и перемещать между собой различные участки сигнала;
- увеличивать и уменьшать громкость речевого сигнала;
- осуществлять различные преобразования сигнала.



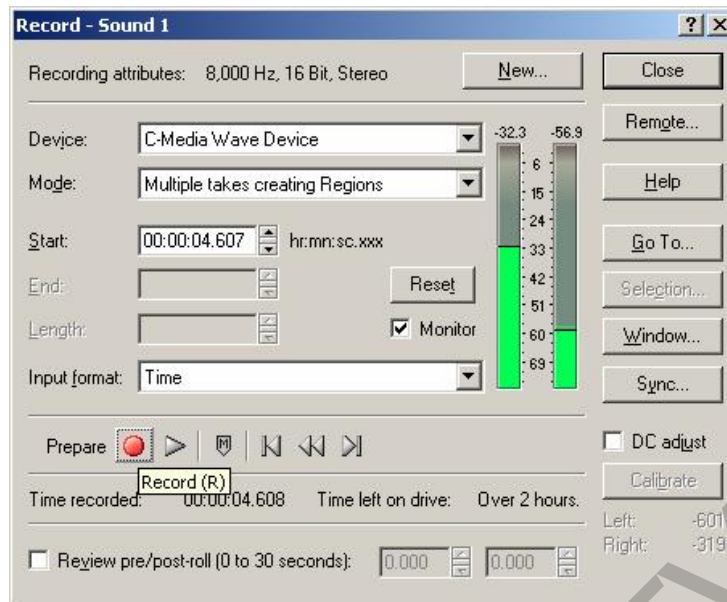
**Рисунок 1.1.** Общий вид внешнего интерфейса программы Sonic Foundry Sound Forge

Рассмотрим некоторые возможности программы *Sound Forge* более подробно.

Для того чтобы записать речь и сохранить ее в файл, необходимо подключить к компьютеру микрофон, настроить его с помощью средств *Microsoft Windows* (см. разд. I) и воспользоваться кнопкой *Record* в панели инструментов *Sound Forge* (рис. 1.1). После этого откроется диалоговое окно *Record* для записи звука (рис. 1.2), в котором можно установить все необходимые атрибуты звукозаписи (в частности, частоту дискретизации, разрядность, количество каналов записи), отследить и настроить с помощью световых индикаторов уровень записи, записать речь с микрофона, а также прослушать сделанную только что запись.

Для того чтобы записать сигнал, необходимо в рассматриваемом диалоговом окне нажать на кнопку *Record* (см. рис. 1.2) и начать произносить фразу в микрофон. В процессе записи звука почти все инструменты рассматриваемого диалога деактивируются. Для завершения записи достаточно нажать на кнопку *Stop* этого же диалогового окна.



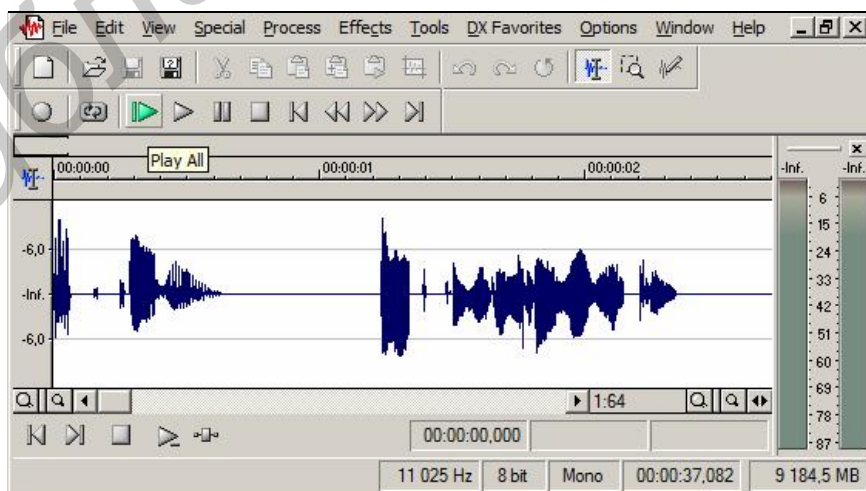


**Рисунок 1.2.** Диалоговое окно для записи звука в Sound Forge

После остановки звукозаписи ее сразу же можно прослушать, используя стандартную кнопку *Play* для воспроизведения звука, находящуюся справа от кнопки *Stop*, в этом же диалоговом окне. Если при воспроизведении окажется, что запись недостаточно качественная, то нажав на кнопку *New*, можно задать частоту дискретизации и разрядность для новой записи (рис. 1.3) и после этого записать речевой сигнал заново.



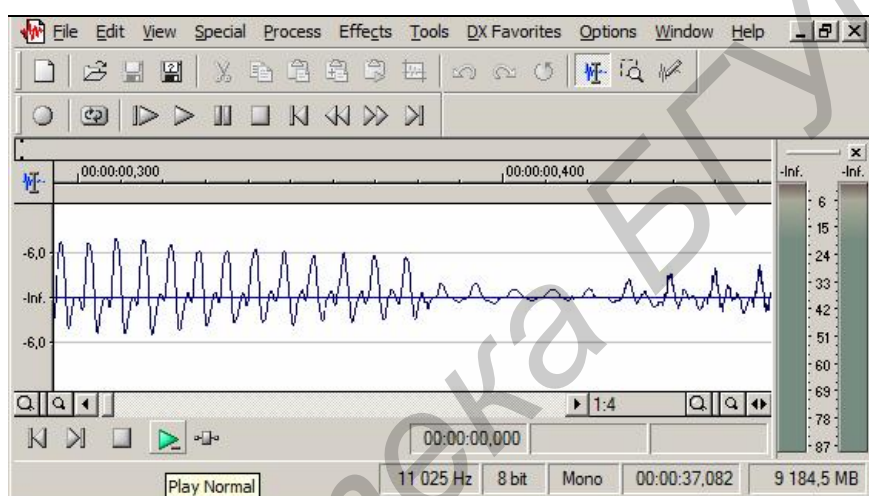
**Рисунок 1.3.** Установка частоты дискретизации и разрядности



**Рисунок 1.4.** Осциллограмма речевого сигнала в окне Sound Forge

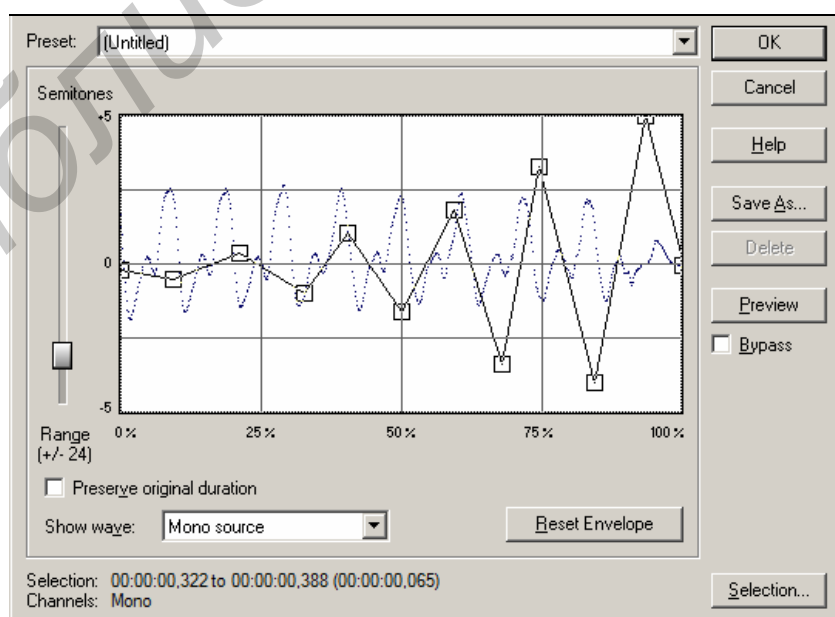
Для того чтобы получить возможность работать с записанным речевым сигналом далее и увидеть его осциллограмму, необходимо закрыть окно *Record*. После этого откроется дочернее окно *Sound Forge*, содержащее изображение осциллограммы только что сделанной записи (рис. 1.4).

Находясь в данном окне, можно исследовать полученную запись, а также осуществлять всевозможные преобразования записанного речевого сигнала. Прежде чем приступить к исследованиям, рекомендуется сохранить запись в звуковой файл в формате wav. Для этого используются стандартные средства сохранения файлов либо на панели инструментов, либо из пункта *File* главного меню программы *Sound Forge*. Следует также отметить, что с помощью стандартного набора инструментов можно открывать файлы с ранее сохраненными звукозаписями, создавать новые, а также вырезать, копировать, перемещать фрагменты записей, выделяя их с помощью мыши прямо на осциллограмме. Для удобства дальнейшей работы выведенное на экран изображение осциллограммы можно масштабировать с помощью инструмента *Zoom*. Тогда осциллограмма принимает вид, похожий на изображение, представленное на рис. 1.5.



**Рисунок 1.5.** Масштабированное изображение осциллограммы речевого сигнала

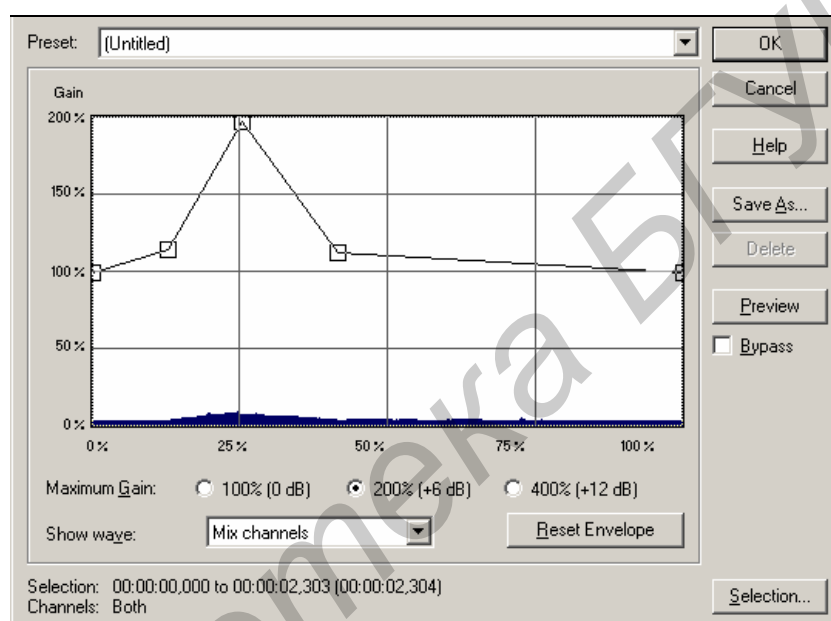
Для того чтобы прослушать записанный речевой сигнал, осциллограмма которого изображена в окне, можно использовать кнопки воспроизведения, промотывания и т.п. на панели инструментов, которые активизируются для этого окна.



**Рисунок 1.6.** Управление частотой основного тона речевого сигнала

В программе *Sound Forge* имеются различные способы преобразования записанного звука. В частности, для преобразования речевого сигнала иногда требуется использовать возможности изменения высоты тона аудиоданных, т.е. изменять частоту основного тона  $F_0$ . Для этого предназначена функция *Pitch Bend*, которую можно найти с помощью меню *Effect -> Pitch -> Bend* (рис. 1.6). Например, с помощью этой функции можно сделать так, что звук будет медленно повышаться от начала до конца записи.

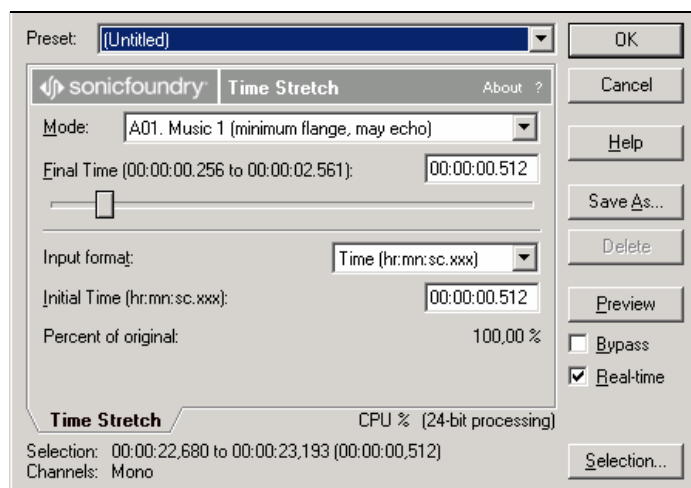
Обычно, когда изменяется высота тона аудиоданных, их длина также меняется. Если увеличить высоту тона, то и данные станут короче, и наоборот, – если уменьшить высоту тона, то данные станут длиннее. Именно это происходит при использовании функции *Pitch Bend*. Чтобы избежать этого, можно использовать функцию *Pitch Shift* программы *Sound Forge* для изменения высоты тона без изменения длины аудиоданных, которую можно найти с помощью меню *Effects -> Pitch -> Shift*.



**Рисунок 1.7.** Регулирование уровня громкости (амплитуды) речевого сигнала

Для регулирования уровня громкости (амплитуды) речевого сигнала и/или отдельных его фрагментов предназначена команда меню *Process -> Fade -> Graphic*, при обращении к которой открывается диалоговое окно *Graphic Fade* (рис. 1.7). В этом окне изображен график. На левой оси графика указываются значения амплитуды, которые могут варьироваться от 0 до 400 % (в зависимости от значения параметра *Maximum Gain*, находящегося под графиком). На графике изображена кривая, показывающая изменение громкости звуковых данных. Левый край этой кривой представляет начало выделенной области, а правый край – ее конец. Изменяя форму кривой, можно регулировать значение амплитуды различных частей выделенного фрагмента речевого сигнала.

Аналогичным образом регулируется панорамирование звука, т.е. то, как будет слышен звук в звуковом поле между двумя динамиками. Для изменения панорамирования используется команда меню *Process -> Pan/Expand*.



**Рисунок 1.8.** Регулировка длины аудиоданных

В процессе обработки звука при изменении длины аудиоданных высота тона тоже изменяется. Иными словами, если укоротить данные, то высота тона увеличится, а если удлинить данные, то высота тона уменьшится. Для того чтобы избежать изменения высоты тона, в программе *Sound Forge* имеется функция, позволяющая изменить длину аудиоданных без воздействия на высоту тона. Эта функция называется *Time Compress/Expand* и вызывается из меню *Process -> Time Stretch* (рис. 1.8).

Существуют также и другие возможности преобразования звука в программе *Sound Forge*. Однако для выполнения приведенных в данном пособии лабораторных работ описанных выше средств вполне достаточно.

## 1.2. Фонетический состав русской речи

**Фонема** – это минимальная значимая звуковая единица речи [1] (*Лобанов Б.М..2006кн-Речев\_И\_ИС*). Любое сообщение можно написать буквами, а можно – фонемами.

Акустические свойства каждой фонемы обусловлены артикуляторными особенностями ее образования в речевом тракте человека, к которым относятся:

- место образования;
- способ образования.

**Определение 1.1.** Под **местом образования** понимается сужение артикуляторного тракта, определяющее его артикуляторную статику и резонансные свойства. Место образования гласных фонем обусловлено положением тела языка (переднее / заднее, высокое / низкое) и губ (огубленное / неогубленное). Место образования согласных фонем определяет положение смычки или щели (губное, зубное, альвеолярное, нёбное, фарингальное), а также заднее или переднее положение тела языка (разделение на мягкие и твердые согласные).

**Определение 1.2.** От **способа образования** фонем зависят энергетические и динамические особенности артикуляции фонемы. Под способом образования фонемы понимается то, каким образом образуется звук при прохождении воздуха через артикуляторный тракт человека во время произнесения фонемы.

В соответствии с артикуляторными особенностями образования, фонемы различаются:

- 1) по способу образования;
- 2) по месту образования.

По способу образования фонемы подразделяются на следующие виды (более подробно см. [1] (Лобанов Б. М., 2006 кн-Речев\_И\_ИС)):

- гласные: [y], [o], [a], [э], [ы], [и];
- плавные: [й];
- боковые (латеральные): [л] (твёрдый) и [л'] (мягкий);
- носовые: [м], [м'], [н], [н'];
- дрожащие: [р], [р'];
- звонкие: [й], [л], [л'], [м], [м'], [н], [н'], [р], [р'], [з], [з'], [ж], [в], [в'], [б], [б'], [д], [д'], [г], [г'];
- глухие: [с], [с'], [ш], [ш'], [ф], [ф'], [х], [х'], [п], [п'], [т], [т'], [к], [к'];
- щелевые (фрикативные): [с], [с'], [ш], [ш'], [ф], [ф'], [х], [х'], [з], [з'], [ж], [в], [в'];
- взрывные: [п], [п'], [т], [т'], [к], [к'], [б], [б'], [д], [д'], [г], [г'];
- аффрикаты: [ц], [ч'].

Гласные отличаются по **месту образования** положением тела языка в координатах верхний / нижний – передний / задний и положением губ в координатах огубленный / неогубленный:

- верхний, задний, огубленный – [y];
- нижний, задний, огубленный – [o];
- нижний, задний, неогубленный – [a];
- нижний, передний, неогубленный – [э];
- верхний, задний, неогубленный – [ы];
- верхний, передний, неогубленный – [и].

Согласные по **месту образования** отличаются положением смычки или щели (губное, зубное, альвеолярное, заальвеолярное, нёбное), а также задним или передним положением тела языка (мягкие и твердые):

- губные твёрдые – [м], [в], [б], [ф], [п];
- зубные твёрдые – [н], [л], [д], [р], [з], [с], [т], [ц];
- заальвеолярные твёрдые – [ж], [ш];
- нёбные твёрдые – [г], [к], [х];
- губные мягкие – [м'], [в'], [б'], [ф'], [п'];
- зубные мягкие – [н'], [л'], [д'], [р'], [з'], [с'], [т'];
- заальвеолярные мягкие – [ш'], [ч'];
- нёбные мягкие – [й], [г'], [к'], [х'].

Полный перечень 42 фонем русского языка:

[y] [o] [a] [э] [и] [ы]

[й] [л] [л'] [м] [м'] [н] [н'] [р] [р'] [п] [п'] [к] [к'] [т] [т']

[в] [в'] [з] [з'] [ж] [б] [б'] [д] [д'] [г] [г']

[ф] [ф'] [с] [с'] [ш] [ш'] [х] [х'] [ц] [ч']

### 1.3. Порядок выполнения лабораторной работы №1

**Задание 1.1.** Повторите основные характеристики речевого сигнала и основы фонетики русской речи, классификацию фонем (см. подразд. 1.2).

**Задание 1.2.** Научитесь работать с программой *Sound Forge* для записи речевых сигналов и отрегулируйте уровни записи и воспроизведения этих сигналов (см. подразд. 1.1). Для этого выполните следующие описанные далее действия.

1. Запустите программу *Sound Forge*. Изучите содержание меню программы, основные команды, назначение кнопок на панели инструментов.
2. Научитесь записывать речевые сигналы с помощью программы *Sound Forge*. Для этого нажмите на панели инструментов кнопку с точкой *Record*. На экране появится окно

с заголовком *Record-Sound 1*. В этом окне нажмите кнопку *Record* и отчетливо произнесите в микрофон любую фразу. После этого нажмите кнопку *Stop* и закройте окно кнопкой *Close*. На экране появится осциллограмма произнесенной фразы. Прослушайте записанную фразу, нажав на кнопку *Play All* на панели инструментов.

- Отрегулируйте уровень громкости воспроизведения речевого сигнала с помощью стандартной программы системы *Windows* «*Регулятор громкости*» (см. разд. I). Для вызова этой программы 2 раза щелкните кнопкой мыши на значке *Громкость*



в правом углу панели задач *Windows*. Установите нормальный уровень громкости воспроизведения, прослушивая записанный речевой сигнал.

- Отрегулируйте уровень записи речевого сигнала с помощью программы «*Регулятор громкости*». Для этого из режима «*Громкость*» этой программы перейдите в режим «*Запись*» (в меню «*Параметры*» выберите команду «*Свойства*» и отметьте настройку уровня записи). Уровень записи будет установлен правильно, если при записи речевого сигнала с помощью программы *Sound Forge* амплитуда сигнала на осциллограмме не будет сильно превышать значение, отмеченное зеленой горизонтальной линией.

**Задание 1.3.** С помощью программы *Sound Forge* исследуйте влияние частоты дискретизации и разрядности речевого сигнала на его разборчивость и качество записи. Для этого выполните следующие описанные далее действия.

- Откройте в программе *Sound Forge* окно *Record* и нажмите кнопку *New...*
- Установленное значение частоты дискретизации (*Sample rate*) 16 000 Гц обеспечивает среднее качество записи речевого сигнала, типичное для микрофона. Запишите произвольную фразу, прослушайте и оцените разборчивость и качество звучания. Уменьшите значение частоты дискретизации до 4 000 Гц, произнесите ту же фразу и запишите новый речевой сигнал. Прослушайте речевой сигнал и отметьте заметное ухудшение качества и разборчивости речи.
- Установите значение частоты дискретизации 8 000 Гц, что соответствует телефонному качеству и разборчивости звучания. Запишите речевой сигнал при этом значении частоты дискретизации, прослушайте его и оцените качество звучания. Увеличьте значение частоты дискретизации до 32 000 Гц. Запишите и прослушайте речевой сигнал. Такая частота дискретизации обеспечивает высокое качество записи и звучания, но в то же время требуется значительно больший объем памяти для преобразования речевого сигнала.
- Установленное значение разрядности речевого сигнала *Bit-Depth* 16 бит соответствует высокому качеству звучания. Установите низкую разрядность 8 бит. Запишите в том же порядке фразы при разной частоте дискретизации, прослушайте речевой сигнал и сделайте выводы об изменении качества звучания и разборчивости речи.

**Задание 1.4.** С помощью программы *Sound Forge* исследуйте фонетические закономерности речевого сигнала.

**Задание 1.4.1.** Изучите **гласные** фонемы русской речи. Для этого выполните следующие описанные далее действия.

Запишите гласные звуки с помощью программы *Sound Forge* следующим образом.

- Откройте окно *Record*. В этом окне нажмите кнопку *Record* и отчетливо произнесите в микрофон гласные звуки: [a], [o], [y], [э], [ы], [и].
- Нажмите кнопку *Stop* и закройте окно кнопкой *Close*. На экране появится осциллограмма произнесенных звуков.
- Проверьте их звучание, нажав на кнопку *Play All* на панели инструментов. Сохраните речевой сигнал с гласными фонемами в файле *sample1.wav* с помощью команды *Save* из пункта меню *File*.

4. Найдите на осциллограмме речевого сигнала участки, соответствующие каждой гласной фонеме (участки выделяются нажатием правой кнопки мыши, озвучиваются – нажатием кнопки *Play* на панели инструментов).
5. Прослушайте, как звучит каждая гласная фонема в отдельности.
6. Исследуйте осциллограмму фонемы [а]. Для этого увеличьте масштаб осциллограммы до 1:2 с помощью кнопки *Zoom In*, находящейся в правом нижнем углу окна осциллограммы. Найдите область фонемы [а] и приблизительно зарисуйте общий вид осциллограммы в виде огибающего контура. После этого визуально определите периодичность сигнала фонемы [а] и выделите с помощью мыши один период. Прослушайте, как он звучит. Это будет похоже на щелчок. Теперь нажмите кнопку *Play Looped* (циклический повтор отмеченного сигнала), и вы услышите подобие произнесенного вами звука [а]. Для остановки нажмите кнопку *Stop*.
7. Таким же образом исследуйте и зарисуйте примерный вид осциллограмм всех остальных гласных фонем русского языка.

В результате проведенных исследований сделайте вывод о том, что гласные фонемы речи могут разборчиво звучать отдельно от других фонем, и более того, даже один повторяющийся период сигнала этих фонем звучит вполне приемлемо.

8. Запишите с помощью программы *Sound Forge* гласные «я», «ё», «ю», «е» и проведите аналогичные исследования, в результате которых покажите, какие фонемы образуют указанные буквы.

Завершите выполнение этого задания нажатием кнопки *Close Window*.

**Задание 1.4.2.** Изучите **шипящие** согласные фонемы русской речи. Для этого выполните следующие описанные далее действия.

1. Запишите с помощью программы *Sound Forge* следующие фонемы: [ф], [с], [ш], [х]. Сохраните этот речевой сигнал в файле *sample2.wav*.
2. На полученной осциллограмме речевого сигнала найдите участки, соответствующие каждой фонеме. Выделите с помощью мыши и отдельно прослушайте каждую фонему.
3. Выделите и отдельно прослушайте участки шипящих согласных фонем. В результате сделайте вывод о том, что шипящие согласные можно выделять и прослушивать отдельно.
4. Запишите с помощью программы *Sound Forge* следующие слоги: «фа», «са», «ша», «ха». Сохраните этот речевой сигнал в файле *sample3.wav*.
5. На полученной осциллограмме речевого сигнала найдите участки, соответствующие каждому слогу. Выделите и отдельно прослушайте слоги с шипящими согласными.
6. Выделите и отдельно прослушайте участки шипящих согласных фонем в слогах. В результате сделайте вывод о том, что шипящие согласные, в отличие от других согласных, можно выделять и прослушивать отдельно от гласных звуков. Однако обратите внимание на влияние гласных на их звучание, сравнив результаты этапов 1–3 с текущими результатами.
7. Увеличьте масштаб осциллограммы до 1:2 и найдите осциллограммы шипящих согласных фонем. Зарисуйте их примерный вид.
8. Теперь попытайтесь выделить отдельный период осциллограммы любой шипящей согласной. Нажмите кнопку *Play Looped* для циклического повторения звучания выделенного периода. Вы услышите звук, даже отдаленно не напоминающий шипящую согласную фонему.

Сделайте вывод о том, что в шипящих согласных не звучат отдельные периоды сигнала в циклическом режиме.

Завершите выполнение этого задания нажатием кнопки *Close Window*.

**Задание 1.4.3.** Исследование фонетических характеристик **взрывных, носовых** и всех остальных согласных русской речи.

1. Запишите с помощью программы *Sound Forge* следующие фонемы: [б], [д], [м], [н]. Сохраните этот речевой сигнал в файле *sample4.wav*.
2. На полученной осциллограмме речевого сигнала найдите участки, соответствующие каждой фонеме. Попробуйте выделить и отдельно прослушать участки каждой из согласных фонем.
3. Запишите с помощью программ *Sound Forge* следующие слоги: «ба», «да», «ма», «на» и аналогичные слоги с другими согласными. Сохраните этот речевой сигнал в файле *sample5.wav*.
4. На полученной осциллограмме речевого сигнала найдите участки, соответствующие слогам «ба» и «да». Попробуйте выделить и отдельно прослушать участки взрывных согласных фонем. Определенных звуков [б] и [д] вы не услышите, так как они как бы влились в соседний гласный звук [а].
5. В слогах «ма» и «на» попробуйте выделить и прослушать участки носовых согласных фонем. Результат будет тем же. Носовые согласные влились в соседний звук [а] и отдельно не звучат.
6. В результате сделайте вывод о том, что все согласные звуки (кроме шипящих) звучат только в контексте с гласными, т.е. происходит наложение артикуляции одного звука на другой (так называемый эффект **коартикуляции**).
7. Проанализируйте аналогичным образом все остальные согласные фонемы русской речи.
8. Увеличьте масштаб осциллограммы до 1:2 и найдите осциллограммы согласных фонем. Зарисуйте их примерный вид.
9. Теперь попытайтесь выделить отдельный период осциллограммы любой из этих согласных. Нажмите кнопку *Play Looped* для циклического повторения звучания выделенного периода.

Сделайте вывод о том, что во всех согласных фонемах (включая шипящие) не звучат отдельные периоды сигнала в циклическом режиме.

Завершите выполнение этого пункта нажатием кнопки *Close Window*.

**Задание 1.5.** Сравните полученные вами результаты с результатами своих однокурсников и проанализируйте влияние голосовых особенностей и особенностей произношения различных дикторов на внешний вид осциллограмм и звучание фонем.

**Задание 1.6.** В качестве отчета по работе продемонстрируйте примерные зарисовки осциллограмм изученных фонем и сделанные в процессе работы выводы.

**Примечание.** Постарайтесь сохранить полученные при выполнении данной работы результаты, для того чтобы их можно было использовать при выполнении следующих работ.

#### 1.4. Вопросы для самопроверки

1. Дайте определение понятия «фонетика». Каков фонетический состав русского языка?
2. Какими артикуляторными особенностями обусловлены акустические свойства фонем?
3. Объясните своими словами понятие места образования фонемы.
4. Объясните своими словами понятие способа образования фонемы.
5. Чем различаются фонемы в зависимости от артикуляторных особенностей их образования?
6. Изобразите в виде схемы классификацию фонем по способу их образования.
7. Перечислите гласные фонемы русского языка и опишите особенности их образования.
8. Изобразите в виде схемы классификацию фонем по месту их образования.
9. Изобразите в виде единой схемы классификацию фонем по способу и месту их образования одновременно.



## Лабораторная работа №2 «Изучение просодических закономерностей речевого сигнала»

**Цель работы** – изучение просодических закономерностей речевого сигнала.

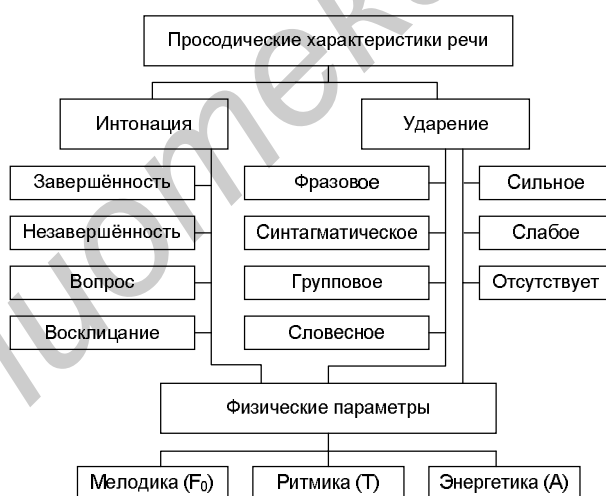
Основные **задачи** работы следующие:

1. Закрепление основных понятий и характеристик речевого сигнала.
2. Закрепление навыков работы с программой *Sound Forge*.
3. Практическое исследование просодических закономерностей речевого сигнала.

### 2.1. Просодические характеристики речевого сигнала

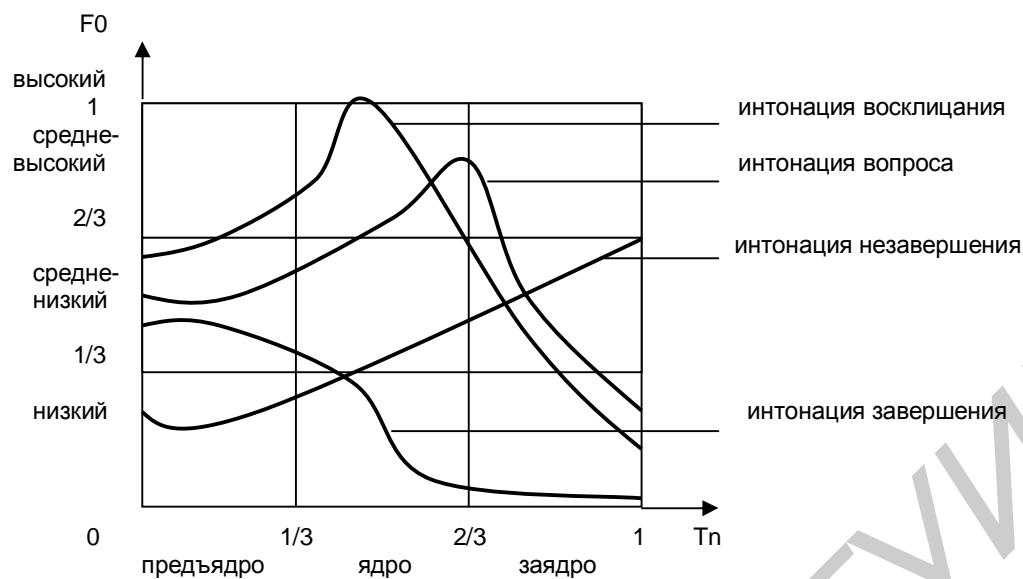
Просодика описывает **ударение** и **интонацию** в речи (рис. 2.1) [1] (*Лобанов Б.М..2006кн-Речев\_И\_ИС*).

Естественной речи присуще использование интонации, так как без этого речь будет звучать монотонно и неестественно. Речь без интонации очень трудно воспринимать, так как вместе с интонацией человек передает дополнительный смысл своих высказываний. Аналогично естественной чертой речи является также наличие ударений. Нам понятно использование ударений в отдельных словах, так как без ударений невозможно будет отличить одно слово от другого и понять смысл произносимых сообщений. Кроме словесных ударений в просодике речи рассматриваются также другие виды ударений, которые делают естественную речь более выразительной и в том числе помогают передавать ее интонацию.



**Рисунок 2.1.** Просодические характеристики речи

Из рис. 2.2 видно, что акцентная группа (АГ) делится на три составляющие: **предъядро**, **ядро** и **заядро**. Это так называемые элементы акцентной группы (ЭАГ). При этом ядро считается главным элементом АГ. Правила разбивки акцентной группы на элементы достаточно просты: **ядро** является главной ударной гласной слова, **заядро** – это все звуки справа от ядра, **предъядро** – все звуки слева от ядра. Получается, что заядро и предъядро могут содержать различное количество звуков вплоть до их отсутствия. Например, фонетический период: «*Ты поедешь сегодня? – Да!*» – содержит фразу «*Да!*», которая является при этом синтагмой, и акцентной группой одновременно. В этом случае единственный гласный звук придется делить на три части, потому что нужно выделить интонационный тип и определить ядро, заядро и предъядро.



**Рисунок 2.2.** Мелодические контуры четырех интонационных типов конечной акцентной группы

**Интонационный тип фразы** определяется положением мелодического максимума относительно главного ударного гласного. Причем положению максимума в центре гласного соответствует интонация побуждения или восклицания, сдвигу его влево – интонация завершения или утверждения, а сдвигу вправо – интонация незавершенности или вопроса. Другие параметры мелодической кривой (значение максимума и ее ширина) существенны при передаче эмоциональных оттенков речи.

## 2.2. Порядок выполнения лабораторной работы №2

**Задание 2.1.** Повторите основные понятия просодики. Вспомните основные фонетические особенности речевого сигнала.

**Задание 2.2.** Исследуйте просодические закономерности речевого сигнала.

**Задание 2.2.1.** Исследуйте явление ударения в просодике и понятие акцентной группы.

1. Запишите с помощью программы *Sound Forge* слово «Саша». Сохраните его в файле *sample2\_1.wav*.
2. Запишите с помощью программы *Sound Forge* слово «кушал». Сохраните его в файле *sample2\_2.wav*.
3. Запишите с помощью программы *Sound Forge* слово «кашу». Сохраните его в файле *sample2\_3.wav*.
4. Создайте новый файл и перенесите в него записанные ранее слова в следующих последовательностях (для этого каждое слово нужно вначале выделить и вырезать с помощью кнопки *Cut* на панели инструментов, а затем вставить в нужном месте осциллограммы нажатием кнопки *Paste*):

- «Саша», «кушал», «кашу»;
- «кушал», «Саша», «кашу»;
- «кушал», «кашу», «Саша»;
- «Саша», «кашу», «кушал»;
- «кашу», «Саша», «кушал»;
- «кашу», «кушал», «Саша».

5. Прослушайте полученные фразы и попытайтесь выявить интонационные различия в них. Сделайте вывод о том, что все они звучат с точки зрения интонации примерно одинаково и включают по 3 акцентные группы, что можно записать следующим образом: (+) (+) (+).
6. Сохраните результат в файле *sample2\_4.wav*.
7. Запишите с помощью программы *Sound Forge* фразу «Саша кушал кашу» целиком с ударением на первое слово. Сохраните ее в файле *sample2\_5.wav*.
8. Прослушайте фразу – она звучит как утверждение. Теперь рассматриваемая фраза состоит из одной акцентной группы вида (+ - -).
9. Найдите на осциллограмме речевого сигнала участок, соответствующий слову «Саша». Переставьте это слово в конец фразы.
10. Прослушайте измененную фразу – теперь она звучит почти как вопрос. При этом акцентная группа преобразовалась к виду (- - +).
11. Сохраните ее в файле *sample2\_6.wav*.
12. Проведите различные эксперименты, меняя местами слова фразы. Сравните полученные результаты с результатами пунктов 1–5.
13. Еще раз произнесите фразу, но с ударением не на первое слово, а на второе или третье. Снова поменяйте местами слова фразы, прослушайте, что у вас получилось, и сравните с предыдущими результатами.
14. Придумайте самостоятельно (или прочитайте из любого текста) несколько различных фраз и попытайтесь выделить в них акцентные группы. Обратите внимание на то, что более длинные и сложные фразы могут состоять из нескольких акцентных групп.

**Задание 2.2.2.** Исследуйте явление интонации в просодике.

1. Запишите с помощью программы *Sound Forge* фразу «Саша кушал кашу» с различными видами интонации:
  - простая завершенность (которая обозначается на письме с помощью точки ( . ));
  - незавершенность (( , ) и (-)) – для этого попробуйте продолжить фразу, например, словами «сегодня на обед»;
  - вопрос ((?));
  - восклицание ((!)).
 Сохраните записанные фразы либо в один файл, либо в разные файлы с расширением \* .wav.
2. Прослушайте записанные фразы и сравните их. Нарисуйте соответствующие каждой фразе интонационные контуры и расстановку ударений.
3. Проведите аналогичный эксперимент с целью анализа эстетической и эмоциональной функций просодики (например, прочитайте фразу с выражением и без, с чувством гордости, радости, огорчения и т.п.).
4. Придумайте самостоятельно (или прочитайте из любого текста) несколько различных фраз и проведите аналогичные исследования.

В результате проведенных исследований сделайте вывод о влиянии просодических характеристик (интонации и ударения) на речевой сигнал.

**Задание 2.3.** Научитесь работать с мелодическими контурами.

1. Запишите с помощью программы *Sound Forge* фразу «Студент писал реферат» монотонно, т.е. на одном тоне и без выделения какого-либо слова. Скопируйте записанную фразу либо после записанной, либо в новый аудиофайл и сохраните обе записи.
2. Выберите в программе *Sound Forge* режим *Effect -> Pitch -> Bend*. Подбирая мелодический контур изменения  $F_0$ , добейтесь звучания интонации завершенности для первой записи, соответствующей фразовому ударению:
  - на первое слово (ответ на вопрос «Кто писал реферат?»);
  - на второе слово (ответ на вопрос «Что делал студент?»);
  - на третье слово (ответ на вопрос «Что писал студент?»).

3. Запишите полученные звуковые файлы. Зарисуйте общий вид полученных интонационных контуров  $F_0$ .
4. Подбирая контур изменения  $F_0$ , добейтесь звучания интонации вопроса во фразе «Студент писал реферат» с ударениями на первое, второе и третье слово.
5. Запишите полученные звуковые файлы. Зарисуйте общий вид полученных интонационных контуров  $F_0$ .
6. Используя в программе *Sound Forge* режим *Process -> Fade -> Graphic*, увеличьте силу звука (амплитуду) ударных слов в записанных файлах. Добейтесь более отчётливо выраженного фразового ударения.
7. Запишите полученные звуковые файлы. Зарисуйте общий вид полученных контуров амплитуды  $A$ .
8. Используя в программе *Sound Forge* режим *Process -> Time Stretch*, увеличьте длительность ударных слов в записанных файлах. Добейтесь более отчётливо выраженного фразового ударения.
9. Проведите аналогичные преобразования для второй записи фразы «Студент писал реферат». При этом вместо режима *Effect -> Pitch -> Bend* используйте режим *Effects -> Pitch -> Shift*.
10. Запишите полученные звуковые файлы.
11. Сравните проведенные преобразования для обеих записей одной и той же фразы. Сделайте соответствующие выводы.

**Задание 2.4.** В качестве отчета по работе продемонстрируйте выполненные вами зарисовки и сделанные в процессе работы выводы.

### 2.3. Вопросы для самопроверки

1. Изобразите в виде схемы просодические характеристики речи.
2. Перечислите физические параметры речевого сигнала, с помощью которых описывается просодия речи.
3. Перечислите и поясните основные функции интонации в речи.
4. Поясните понятие ударения в речи.
5. Перечислите типы ударений в речи.
6. Приведите примеры использования в речи ударений различного типа.
7. Приведите все возможные варианты расстановки группового ударения для фразы «Саша кушал кашу».
8. Приведите все возможные варианты расстановки группового ударения для фразы «Сегодня очень хороший день».
9. Объясните понятие мелодического (интонационного) контура.
10. Перечислите и поясните на примере типы синтагм.
11. Чем определяется интонационный тип синтагмы?
12. Перечислите и поясните на примере типы акцентных групп.
13. Перечислите и поясните на примере основные интонационные типы конечных акцентных групп.
14. Изобразите на плоскости « $T - F_0$ » (время – частота основного тона) основные интонационные типы конечных акцентных групп в виде соответствующих мелодических контуров.
15. Перечислите элементы акцентных групп (ЭАГ) и приведите примеры разбиения акцентных групп на ЭАГ.

## Лабораторная работа №3

### «Исследование компиляционных методов синтеза речевого сигнала»

**Цель работы** – исследование компиляционных методов синтеза речевого сигнала и практическое моделирование синтеза речи с использованием указанных методов.

Основные **задачи** работы следующие:

1. Закрепление основных понятий и характеристик речевого сигнала.
2. Закрепление навыков работы с программой *Sound Forge*.
3. Практическое исследование основных проблем автоматического синтеза речи с использованием компиляционных методов.
4. Практическое моделирование синтеза речи с использованием компиляционных методов.

#### 3.1. Методы синтеза речи

Все существующие в настоящее время методы синтеза речевого сигнала делятся на два основных класса [1] (*Лобанов Б.М. 2006кн-Речев\_И\_ИС*):

- параметрические;
- компиляционные.

**Параметрические методы синтеза речевого сигнала** характеризуются тем, что при их реализации используются знания о том, каким образом осуществляется образование речи у человека и какими параметрами описывается речевой сигнал. Основываясь на этих знаниях, осуществляется моделирование акустических процессов речеобразования. В результате такого моделирования речевой сигнал генерируется подобно тому, как это делается в речевом тракте человека. Моделирование процессов речеобразования человека в системах автоматического синтеза речи осуществляется одним из двух способов:

- на основе знаний об артикуляторных процессах, происходящих в речевой системе человека;
- путем математического описания собственно акустических процессов.

**Компиляционные методы синтеза речевого сигнала** игнорируют физику процесса речеобразования. В них для синтеза речи используются отрезки естественной речевой волны, и из них в дальнейшем собирается (компилируется) речевой сигнал. В качестве отрезков речевой волны могут использоваться самые разные сегменты:

- микроволны (отрезок, вырезаемый из записей отдельных фонем или аллофонов);
- аллофоны;
- дифоны (всевозможные паросочетания звуков) и др.

Каждый из указанных классов методов синтеза речи обладает своими достоинствами и недостатками.

**Компиляционные методы** имеют следующие **достоинства**:

- синтезированная речь звучит близко к естественной;
- с помощью компиляционного метода можно «научить» компьютер разговаривать голосом любого человека (например своим голосом) – это качество используется в активно развивающемся в настоящее время направлении клонирования голоса [6, 7] (*Лобанов Б.М. 2002ст-Комп\_К\_П\_Г; Цирульник Л.И. 2006ст-Автом\_С\_К*).

**Недостатки компиляционных методов**:

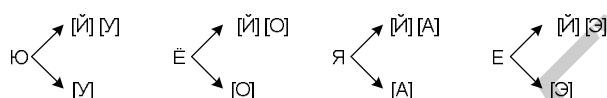
- трудоемкость разработки;
- необходимость довольно больших объемов памяти для хранения баз голосов.

## 3.2. Правила транскрибирования

Как указано выше, основной идеей компиляционных методов синтеза речи является использование сегментов естественной речевой волны для компиляции из них слитных потоков речи. Однако в процессе преобразования естественно-языкового текста в речь предварительно необходимо преобразовать этот текст в последовательность фонем, из которых фактически состоит речевой поток.

Для преобразования орфографического текста в фонемный используются правила транскрибирования. В каждом языке приняты свои правила произношения слов, которым соответствуют свои правила транскрибирования. Рассмотрим здесь некоторые правила транскрибирования, характерные для русской речи.

В русском языке 10 гласных букв: У, О, А, Э, Ы, И, Ю, Ё, Я, Е. Десяти гласным буквам ставятся в соответствие 6 гласных фонем: [У], [О], [А], [Э], [Ы], [И]. Получается, что 4 гласные буквы Ю, Ё, Я и Е как бы «лишние». Они превращаются либо в пары фонем, либо в отдельные фонемы следующим образом:



При этом используются следующие правила транскрибирования [8] (Печенева Т.А. 2004 *сост-Виды\_Я\_Р*):

1. «Преобразование гласной буквы Ю, Ё, Я или Е в пару фонем [Й] и [У], [О], [А] или [Э] соответственно происходит в тех случаях, когда эти буквы в слове стоят в одном из следующих положений: в начале транскрибируемого слова; после мягкого знака; после твердого знака; после гласной буквы». Например:  
вьюга -> [в й у г а], объём -> [а б й о м], яма -> [й а м а], поёт -> [п а й о т].
2. «В остальных случаях после согласных гласные Ю, Ё, Я или Е преобразуются в фонемы [У], [О], [А] или [Э] соответственно и при этом смягчают стоящие перед ними согласные». Например: тюк -> [т' у к], лёд -> [л' о т], пять -> [п' а т'], семь -> [с' э м'].

Из этих правил имеются исключения, в частности, в тех случаях, когда рассматриваемая гласная является безударной. В этом случае звук изменяется (редуцируется). Например, безударная буква Я может быть преобразована в звук [И]: тянуть -> [т' и н у т'].

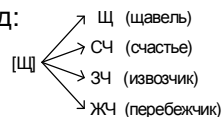
В русском языке есть правило, согласно которому безударная гласная буква О преобразуется в фонему [А], а безударная Е – в [И]. Например: молоко -> [м а л а к о], дерево -> [д' э р и в а].

Еще правило: «Буква И после согласных Ж и Ш преобразуется в звук [Ы]». Например: жить -> [ж ы т'], шить -> [ш ы т'].

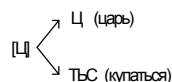
Как видно из формулировок, указанные правила хорошо описываются алгоритмически.

Имеются различные правила преобразования согласных. Рассмотрим некоторые из них:

1. Правило образования звука [Щ] имеет следующий вид:



2. Фонема [Ц] образуется одним из следующих двух способов:



3. Правило оглушения: звонкая согласная перед глухой согласной или в конце слова преобразуется в парную ей глухую. Например: лёд -> [л' о т], буг -> [б у к].
4. Правило озвончения: если глухая согласная стоит перед звонкой, то звук преобразуется в звонкий. Например: косьба -> [к а з' б а], просьба -> [п р о з' б а].
5. Для некоторых пар согласных работает явление уподобления. Например: сшить -> [ш ы т'].

6. Смягчение парных согласных: если после согласной стоит мягкая согласная, то и эта согласная также смягчается. Например: хвостик -> [х в о с' т' и к].

Здесь также есть исключения: слово «бог» преобразуется в транскрипцию [б о х], а не в [б о к], слово «что» преобразуется в [ш т о].

Все эти правила хорошо известны из фонетики [9] (*Зиндер Л.Р.1979кн-Общая\_Ф*) и легко поддаются алгоритмизации.

### 3.3. Порядок выполнения лабораторной работы №3

**Задание 3.1.** Повторите классификацию методов синтеза речевого сигнала. Повторите понятие системы автоматического синтеза речи. Вспомните или изучите обобщенную структуру системы синтеза речи [1] (*Лобанов Б.М..2006кн-Речев\_И\_ИС*). Изучите правила транскрибирования, используемые в русской речи.

**Задание 3.2.** Запишите несколько фраз естественного языка на бумаге, а затем прочитайте их в микрофон и запишите в файлы \*.wav в программе *Sound Forge*. Эти фразы будем считать эталонными. Они вам понадобятся для дальнейшего сравнительного анализа. При этом желательно предусмотреть наличие фраз следующего типа:

- фраза из одной синтагмы;
- фраза из нескольких синтагм;
- фразы с различным типом интонации;
- фразы с перечислениями, т.е. сложные предложения с запятыми, союзами и т.п.;
- общее количество фраз не должно быть слишком большим – примерно от 3 до 5.

**Задание 3.3.** Практическое исследование компиляционных методов синтеза речевого сигнала.

**Задание 3.3.1.** Подготовьте фонетическую базу для синтеза речи.

1. Для выполнения данного задания возьмите записанные вами при выполнении лабораторной работы №1 файлы с гласными и согласными фонемами (если они у вас не сохранились, то вам придется записать их заново).
2. Запишите на бумаге выбранные вами при выполнении задания 2 фразы естественного языка в виде фонемного текста, т.е. в виде последовательности фонем.

**Задание 3.3.2.** Сформируйте звуковые файлы на основе фонемных текстов.

1. Создайте новый пустой файл в программе *Sound Forge*.
2. На основе подготовленных ранее фонемных текстов сформируйте соответствующие фразы речи, компилируя их из подготовленных наборов фонем русского языка. Например, если у вас записано фонемное слово [с а ш а], то в пустой файл скопируйте сначала осциллограмму фонемы [с], затем – [а], [ш] и [а].
3. Таким же образом запишите в отдельные файлы все выбранные вами фразы.

**Задание 3.3.3.** Проанализируйте скомпилированную речь.

1. Прослушайте «синтезированные» вами речевые высказывания и сравните их с эталонными. Сравните их осциллограммы.
2. Сделайте выводы о степени разборчивости и естественности звучания. Отметьте в отчете основные положительные и отрицательные стороны компиляционного метода синтеза речи.

**Задание 3.3.4.** Попытайтесь улучшить качество звучания записанной речи путем дополнительной обработки осциллограммы.

1. Выберите какую-нибудь одну фразу.
2. Посмотрите еще раз внимательно на осциллограмму эталонной фразы и синтезированной. Попробуйте путем выбрасывания лишних участков осциллограммы,

склеивания, перетаскивания, увеличения или уменьшения амплитуды, длительности и других подобных действий добиться более качественного звучания.

3. Сделайте соответствующие выводы.

**Задание 3.3.5.** Попробуйте улучшить качество звучания записанной речи путем совершенствования фонетической базы.

1. Разбейте подготовленные вами фонемные тексты на отдельные слоги и запишите эти слоги аналогично тому, как были надиктованы отдельные фонемы.
2. Осуществите компиляцию речи для своих фраз с использованием заготовленных слогов.
3. Сравните полученные фразы с эталонными и с теми, которые были сформированы ранее. Сделайте соответствующие выводы. Отметьте, какие фонетические и просодические особенности речи следует учитывать при разработке синтезаторов речи.

**Задание 3.3.6\*.** Попробуйте улучшить естественность звучания записанной речи путем обработки контуров мелодики, ритмики и энергетики.

1. Вспомните просодические параметры речи.
2. Зарисуйте мелодические, ритмические и энергетические контуры синтезируемых фраз.
3. Попробуйте описать сформированные контуры путем перечисления приблизительных или относительных значений просодических параметров  $F_0(t)$ ,  $A(t)$ ,  $T(t)$ .
4. На основе описанных параметров попробуйте сделать синтезированные вами фразы более интонационно насыщенными, используя любые известные вам методы обработки просодических параметров. Сделайте соответствующие выводы.

**Задание 3.4.** В качестве отчета по работе представьте все рабочие материалы (естественно-языковые фразы, фонемные тексты, просодически размеченные тексты и т.д.) и сделанные в процессе работы выводы. Нарисуйте обобщенную структуру синтезатора речи, который использует компиляционный метод синтеза, промоделированный при выполнении данной работы. Данный вид синтезаторов иногда называют компиляторами.

Сделайте общий вывод по работе, ответив на вопрос, какие особенности речи следует учитывать для того, чтобы получить более качественное и естественное звучание. Каких компонентов недостает для того, чтобы разработанный вами «синтезатор» стал полноценной системой автоматического синтеза речи с неограниченным словарем?

**Примечание.** Задания, помеченные звездочкой «\*», не являются основными и рекомендуются как дополнительные.

### 3.4. Вопросы для самопроверки

1. Перечислите и охарактеризуйте основные методы синтеза речевого сигнала.
2. Перечислите основные достоинства и недостатки параметрических методов синтеза речевого сигнала.
3. Перечислите основные достоинства и недостатки компиляционных методов синтеза речевого сигнала.
4. Перечислите основные этапы разработки компиляционного синтезатора речи.
5. Перечислите основные компоненты компиляционного синтезатора речи и попробуйте представить в виде схемы структуру такого синтезатора.
6. Попробуйте представить в виде схемы структуру синтезатора речи, использующего параметрические методы синтеза. Найдите отличия такого синтезатора от компиляционного.
7. Что такое правила транскрибирования и как они используются в синтезаторах речи?
8. Перечислите основные правила транскрибирования гласных в русской речи.
9. Перечислите основные правила транскрибирования согласных в русской речи.
10. Приведите примеры правил транскрибирования для белорусской речи, которые отличаются от правил транскрибирования русской речи.
11. Приведите примеры правил транскрибирования для речи какого-либо из иностранных языков.



## Лабораторная работа №4

### «Практическое изучение формантно-параметрического описания речевого сигнала»

**Цель работы** – практическое изучение формантно-параметрического описания речевого сигнала.

Основные **задачи** работы следующие:

1. Ознакомление с акустическими характеристиками фонем и основами формантного анализа речевых сигналов.
2. Получение навыков работы с программой *Praat* для исследования характеристик речевого сигнала.
3. Исследование с помощью программы *Praat* формантных характеристик фонем, эффекта коартикуляции фонем.

#### 4.1. Описание программы Praat для исследования спектральных характеристик речевого сигнала

Для изучения формантно-параметрического описания речевого сигнала предлагается использовать программу *Praat* (<http://www.praat.org/>), которая позволяет:

- записывать речевой сигнал и воспроизводить его;
- загружать ранее записанный речевой сигнал;
- прослушивать речевой сигнал;
- анализировать осциллограмму речевого сигнала;
- раскладывать речевой сигнал в спектр;
- анализировать частотные (в том числе формантные) и амплитудно-временные характеристики речевого сигнала;
- изменять при необходимости масштабы просмотра;
- фиксировать в графическом формате все необходимые для последующего использования данные о речевом сигнале и т.д.

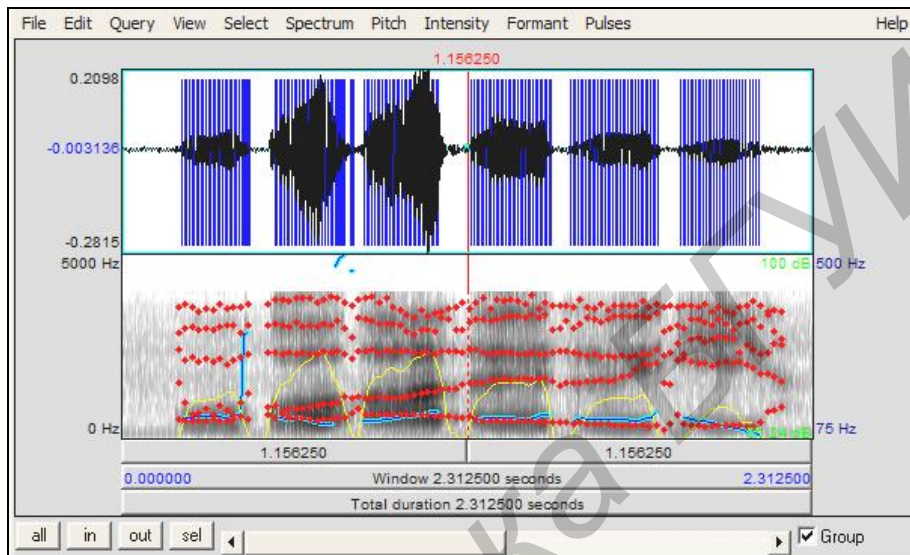
Для загрузки программы *Praat* необходимо запустить соответствующее приложение *praat.exe*. При этом откроются два основных окна программы: *Praat Objects* и *Praat Picture*. Для выполнения данной лабораторной работы потребуется только окно *Praat Objects*. Поэтому для удобства работы рекомендуется развернуть его на полный экран, а окно *Praat Picture* можно закрыть. В процессе выполнения работы будет использоваться еще одно окно – *Sound <file\_name>*. Поэтому ниже будет описана работа только в этих двух окнах.

Для работы с программой в каждом из его окон существует соответствующее главное меню, содержащее все необходимые функции для анализа речевого сигнала и отображения необходимой информации в графическом и текстовом виде. Для удобства работы в программе для большинства наиболее важных функций можно использовать комбинации клавиш. Основной комбинацией клавиш для окна *Praat Objects* является *Ctrl+O*, которая используется для открытия файла речевого сигнала, записанного в формате *\*.wav*. Основные функциональные клавиши для окна *Sound* следующие:

- *Ctrl+Z* – для отмены последнего действия;
- *Ctrl+X* – для вырезания выделенного фрагмента речевого сигнала;
- *Ctrl+C* – для копирования выделенного фрагмента;
- *Ctrl+V* – для вставки выделенного фрагмента;
- *Ctrl+A* – показать (уместить) весь сигнал/спектр на экране;
- *Ctrl+I* – увеличение на 10 %;

- *Ctrl+O* – уменьшение на 10 %;
- *Ctrl+N* – увеличение в рамках выделенного фрагмента;
- *Shift+Tab* – проиграть весь сигнал;
- *Ctrl+W* – закрыть окно.

В окне *Sound*, аналогично программе *Sound Forge*, отображается осциллограмма загруженного или записанного речевого сигнала, в нижней части окна – динамическая сонограмма этого же сигнала (рис. 4.1). В этом же окне с помощью мыши можно выделять любые фрагменты речевого сигнала и с использованием комбинации клавиш *Ctrl+N* масштабировать выделенную часть на все окно для его более детального анализа.

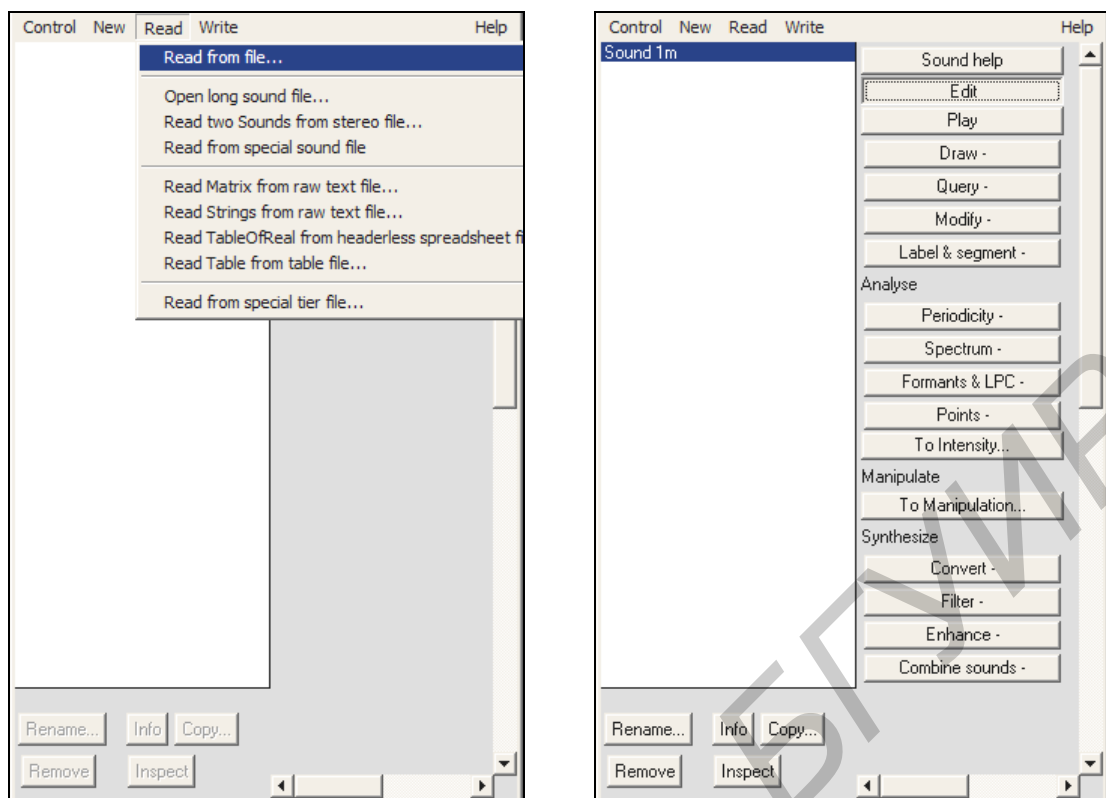


**Рисунок 4.1.** Окно Sound программы Praat

Выделять фрагменты сигнала/спектра можно нажатием левой кнопки мыши и перетягиванием курсора на необходимый интервал. После этого выделенный фрагмент окажется обрамленным широкой линией розового цвета в форме прямоугольника со скругленными углами. Выделенный фрагмент можно прослушать, нажав одну из длинных кнопок (с обозначением длительности звучания) прямо под выделенным фрагментом внизу окна. Как указывалось выше, размер выделенного фрагмента можно масштабировать для лучшего просмотра. Кроме того, его можно вырезать, копировать и перемещать, используя указанные выше комбинации клавиш. Для большей информативности рекомендуется разворачивать все рабочие окна на полный экран.

Для того чтобы получить изображения осциллограммы и сонограммы речевого сигнала, записанного в файле *\*.wav*, необходимо выбрать из главного меню окна *Praat Object* пункты *Read -> Read from file*, найти необходимый файл на диске, а затем, выделив его в списке объектов, использовать кнопку *Edit* на появляющейся панели инструментов (рис. 4.2).

Для записи нового звукового фрагмента необходимо выбрать из меню *New* пункт *Record Sound...*. В открывшемся окне необходимо выбрать частоту дискретизации 8000 и, нажав кнопку *Record*, произнести необходимую фразу. После этого нажать кнопку *Stop*. Затем ввести имя фрагмента в поле, находящееся справа от кнопки *Left to list*, и нажать ее.



**Рисунок 4.2.** Загрузка записанного речевого сигнала в программе Praat

Для анализа формантно-параметрических характеристик речевого сигнала необходимо использовать соответствующие пункты меню программы *Praat*. Как показано на рис. 4.1, на сонограмме речевого сигнала отображается дополнительная информация о значениях его характеристик. В частности, с помощью точек красного цвета выделяются значения формантных составляющих речи (подразд. 4.2), тонкая контурная линия желтого цвета указывает изменения интенсивности сигнала, а с помощью более толстых линий голубого цвета изображается интонационный контур (см. подразд. 2.1), т.е. изменение частоты основного тона  $F_0$ . Отображение всех перечисленных характеристик может быть как установлено, так и отменено с помощью пунктов меню *Pitch* -> *Show pitch*; *Intensity* -> *Show intensity*; *Formant* -> *Show formants* соответственно. Кроме того, имеется возможность настройки для расчета отображения указанных характеристик.

Для того чтобы получить числовые значения формант отдельных сегментов речевого сигнала, необходимо установить указатель мыши в анализируемую точку и либо использовать один из пунктов меню: *Formant* -> *Get first formant* - для вывода значения первой форманты, *Formant* -> *Get second formant* - второй форманты, *Formant* -> *Get third formant* - третьей форманты, *Formant* -> *Get fourth formant* - четвертой форманты, либо просто нажать одну из клавиш *F1*, *F2*, *F3* или *F4* соответственно (рис. 4.3). После этого в отдельном окне *Info* будет выведено соответствующее числовое значение. Для получения числовых значений полосы формант используются пункты меню *Formant* -> *Get first bandwidth* - для вывода значения первой форманты, *Formant* -> *Get second bandwidth* - второй форманты, *Formant* -> *Get third bandwidth* - третьей форманты, *Formant* -> *Get fourth bandwidth* - четвертой форманты.

Для анализа формантных характеристик речевого сигнала следует также установить параметры расчета формант, как указано на рис. 4.4.

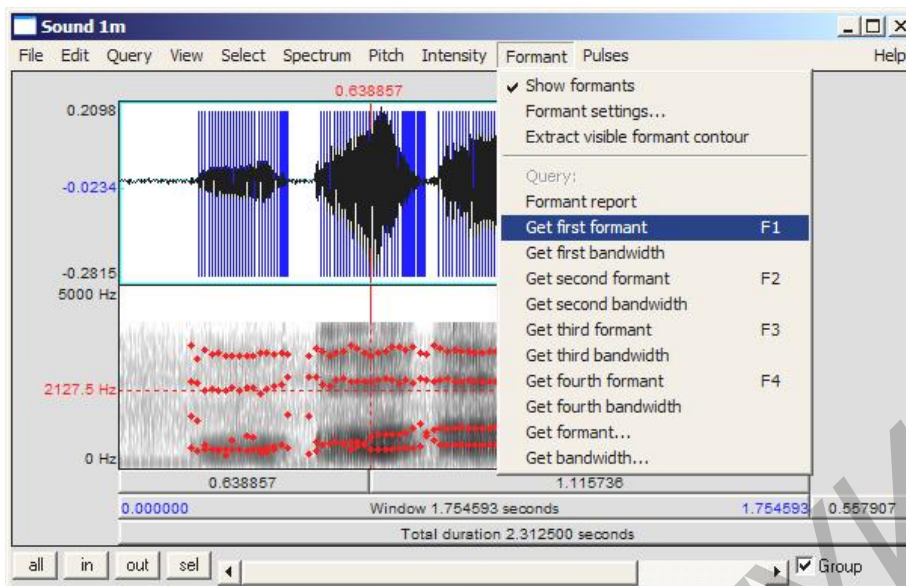


Рисунок 4.3. Получение числовых значений формант речевого сигнала в программе Praat

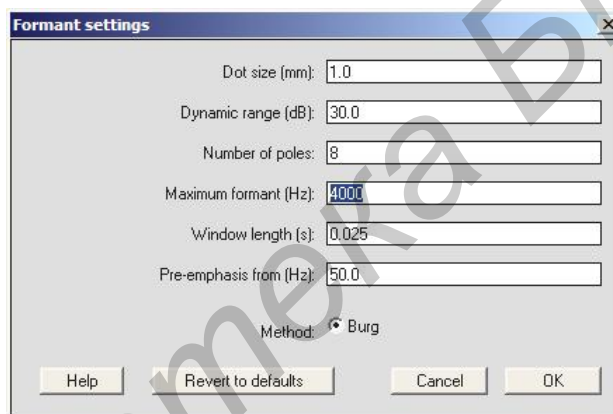


Рисунок 4.4. Параметры расчета формант речевого сигнала

## 4.2. Формантная модель речевого сигнала

Под **формантой** в общем случае понимают полосы передаточной функции речевого тракта, характеризующиеся частотой  $F_i$ , амплитудой  $A_i$  и полосой пропускания  $B_i$ .

На амплитудно-частотном спектре форманты проявляются в виде заметных максимумов (рис. 4.5). Однако не всякий максимум является формантой. Обычно в диапазоне частот первых четырех формант ряд побочных максимумов не является формантами [1] (Лобанов Б.М..2006кн-Речев\_И\_ИС).

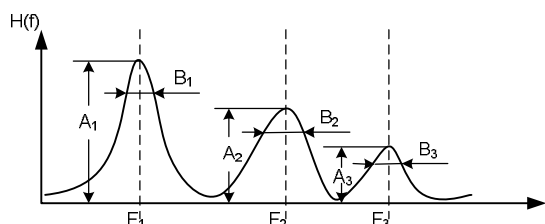


Рисунок 4.5. Выделение формант в спектре речевого сигнала

Анализ вида передаточных функций речевого тракта для различных групп звуков показал, что достаточно полно описать их акустические характеристики можно, используя **формантную модель** [1] (Лобанов Б.М..2006кн-Речев\_И\_ИС).

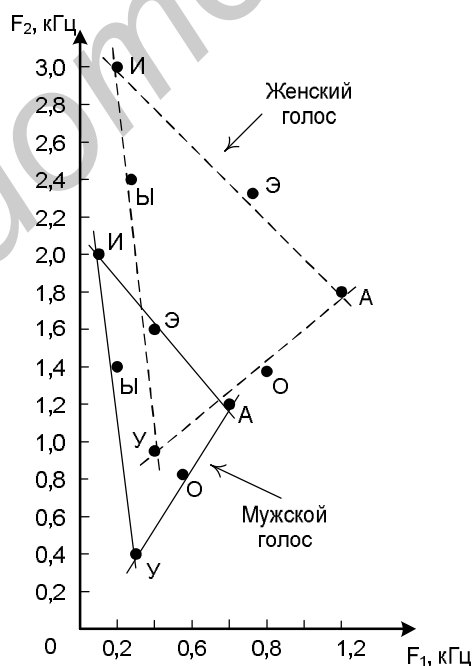
Управляющими параметрами модели являются следующие **10 формантных параметров**:

- $A_p, A_n$  – амплитуды голосового возбуждения ротовых и носовых формант;
- $A_\phi$  – амплитуда шумового возбуждения фрикативных формант;
- $A_a$  – амплитуда аспиративного возбуждения ротовой форманты;
- $F_0$  – частота основного тона;
- $F_1, F_2, F_3$  – частоты первой, второй и третьей формант;
- $F_\phi$  – частота фрикативных формант;
- $V_\phi$  – полоса пропускания фрикативных формант.

Приведенная формантная модель с успехом используется в системах синтеза речи, а ее управляющие параметры являются примером полного набора информативных признаков фонем для систем распознавания речи.

Различия в способе образования фонем определяются амплитудой голосового возбуждения ротовых формант  $A_p$ , носовых формант  $A_n$ , амплитудой шумового возбуждения фрикативных формант  $A_\phi$ , амплитудой аспиративного возбуждения ротовых формант  $A_a$ , а также частотой основного тона  $F_0$ . Причем очень важной является временная организация этих параметров.

Группа гласных отличается от согласных тем, что для них  $A_p = 1$ , а  $A_\phi = A_n = A_a = 0$ . Кроме того, внутри группы гласных фонем, т.е. по месту образования, гласные отличаются значениями формантных частот  $F_1, F_2$  и  $F_3$ , причем основную роль играют  $F_1$  и  $F_2$ . На рис. 4.6 приведен пример распределения формантных частот гласных для мужских и женских голосов. Приведенные здесь зависимости значений первой и второй формант для одного голоса образуют так называемые **формантные треугольники**, в рамках которых распределяются формантные частоты гласных большинства дикторов. Данные на рисунке значения являются усредненными и могут варьироваться в рамках допустимых диапазонов частот формант для разных голосов.



**Рисунок 4.6.** Распределение формантных частот гласных для мужских и женских голосов

При внимательном рассмотрении данных формантных треугольников можно сделать определенные выводы о близости (схожести) различных фонем. Например, хорошо видно,

что фонемы [ы] и [и] очень незначительно различаются по значению первой форманты. При этом фонемы [y] и [o] близки по значению второй форманты. Кроме того, диапазон частот формант женского голоса значительно выше и более разбросан по сравнению с данными для мужского голоса. Этот факт, в частности, привносит дополнительные трудности в решение задачи распознавания речи женского голоса.

Наибольшую роль в разделении согласных по месту образования играет формантная частота  $F_2$ . Причем вследствие коартикуляции  $F_2$  зависит от места образования не только согласного, но и окружающих его гласных (рис. 4.7). Как видно из рисунка, наибольшему коартикуляционному воздействию подвергается группа небных твердых согласных [к], [г], [х], а наименьшему – мягкие согласные.

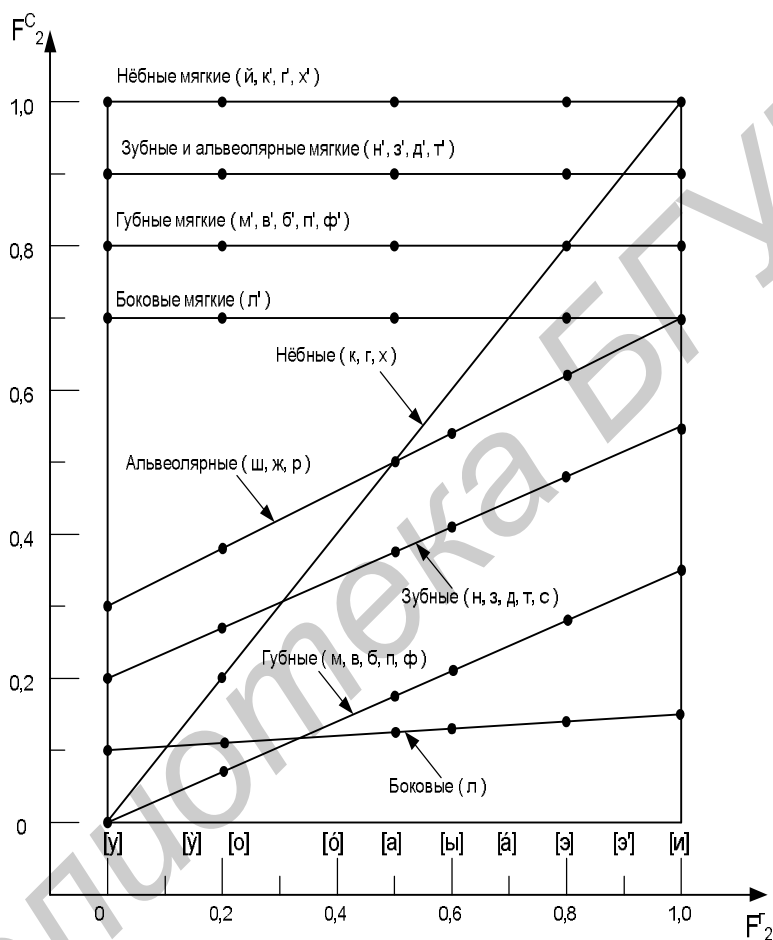


Рисунок 4.7. Зависимость частоты второй форманты согласных от частоты гласных в слоге

### 4.3. Порядок выполнения лабораторной работы №4

#### Лабораторная работа №4. Часть 1

**Задание 4.1.** Научитесь работать с программой *Praat*. Запустите программу *Praat*, изучите содержание меню программы, назначение кнопок на панели инструментов.

**Задание 4.2.** Исследуйте формантные характеристики гласных фонем русской речи: измерьте форманты гласных фонем мужского или женского\* голосов и отобразите их на формантной плоскости. Для этого выполните следующие указанные ниже действия.

**\* Примечание.** В данной части лабораторной работы необходимо проанализировать какой-нибудь из двух голосов: мужской либо женский. При этом можно руководствоваться либо своим интересом, либо логикой. Например, если вы являетесь «обладателем» мужского голоса, то в данной работе выполняйте все задания, связанные с анализом мужского голоса (в том числе, своего собственного). Если вы являетесь представительницей женского пола, то анализируйте сначала женский голос. Голоса «другой» категории вы сможете аналогичным образом проанализировать при выполнении второй части этой лабораторной работы (см. задание 4.12).

**Задание 4.2.1.** Измерьте форманты гласных фонем мужского (женского) голоса.

Для этого необходимо открыть файл *1m.wav* (*1f.wav*) с записью гласных (У-О-А-Э-Ы-И), произнесенных мужским (женским) голосом.

Для этого в окне *Praat Objects* выберите из меню *Read* пункт *Read from file...* или нажмите комбинацию клавиш *Ctrl+O*. После этого в окне объектов и функций появится *Sound 1m* (*Sound 1f*). Для определения формант используется подробный спектр Фурье, получить который можно нажатием кнопки *Edit* справа от окна. Откроется окно *Sound 1m*, в котором в верхней части отображен сам сигнал (его осциллограмма), а в нижней – его разложение в спектр. Необходимо (для удобства работы) в меню *Pitch* отключить пункт *Show Pitch*, в меню *Intensity* -> *Show Intensity* (если они уже не отключены). В пункте *Formant* -> *Formant Settings* установить: *Number of poles* – 8, *Maximum formant* – 4000.

Для анализа каждой отдельной фонемы, ориентируясь с помощью осциллограммы, выделяйте последовательно каждую из них нажатием левой кнопки мыши и ее движением от одной границы фонемы к другой. Вы увидите меняющуюся размеры рамочки розового цвета. Выделив нужную фонему, прослушайте ее, нажав левой кнопкой мыши на длинную полосу в виде кнопки внизу выделенного фрагмента. Убедитесь, что это именно та фонема, которую вы собираетесь анализировать. После этого, не убирая выделение, для удобства работы и большей наглядности вы можете с помощью клавиш *Ctrl+N* увеличить анализируемый фрагмент на все окно. Если вам понадобится вернуться обратно к изображению всего речевого сигнала полностью, используйте нажатие клавиш *Ctrl+O* до тех пор, пока не получите требуемый результат.

Перемещаясь по оси частот спектра, необходимо определить две формантные частоты  $F_1$  и  $F_2$  (в Гц) всех гласных звуков. Для этого необходимо курсор мыши подвести на требуемую форманту (определять формантные частоты лучше всего в самой середине спектра гласного) и нажать левую кнопку мыши. Далее, используя либо пункт меню *Formant* -> *Get first formant*, либо клавишу *F1*, определите значение частоты первой форманты, которое появится в отдельном окне *Info* (для удобства дальнейшей работы разместите все окна так, чтобы они не накладывались друг на друга). Запомните или запишите это значение. Вернувшись в окно *Sound*, нажмите клавишу *F2* (или аналогичным образом используя пункты меню) и получите значение второй форманты в рассматриваемой точке спектра. Таким образом найдите на спектре точку, в которой значения  $F_1$  и  $F_2$  являются максимальными. Это и есть значения формантных частот  $F_1$  и  $F_2$ , которые необходимо занести в приведенную ниже табл. 4.1.

**Таблица 4.1.** Формантные частоты гласных

	У	О	А	Э	Ы	И
$F_1$						
$F_2$						

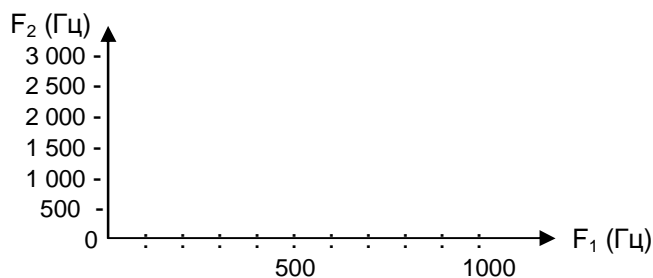
Аналогичным образом проанализируйте спектры всех гласных.

**Задание 4.2.2.** Измерьте форманты гласных фонем своего голоса.

Для этого запишите с помощью микрофона произнесенные гласные: У-О-А-Э-Ы-И – примерно с тем же темпом, как и в рассмотренном выше примере. Форманты гласных определите так же, как в задании 4.2.1 и поместите их значения в таблицу, аналогичную табл. 4.1.

**Задание 4.2.3.** Отобразите гласные звуки на формантной плоскости.

Для этого постройте два формантных треугольника гласных (для данного в файле и своего собственного голоса в отдельности) в следующих координатах на графике, общий вид которого представлен на рис. 4.8.



**Рисунок 4.8.** Координатная плоскость для графика, содержащего формантные треугольники гласных

Сравните полученные результаты анализа двух мужских (женских) голосов и сделайте выводы.

**Задание 4.3.** Исследуйте формантные характеристики коартикуляции «согласный – гласный».

**Задание 4.3.1.** Определите и постройте график зависимости частоты форманты **твердого губного согласного** от частоты гласного в слоге, рассчитайте коэффициент коартикуляции «губной твердый согласный – гласный».

Для этого в программе *Praat* откройте файл *2m.wav*, записанный мужским голосом (или *2f.wav* – женским голосом), с записью слогов «бу – бо – ба – бэ – бы». По спектру Фурье определите величину **второй форманты** согласного ( $F_2^c$ ) и гласного звука ( $F_2^r$ ). Величина  $F_2^c$  определяется в самом начале области гласного (эту область легко определить по спектрограмме Фурье), а величина  $F_2^r$  – в середине области гласного. Полученные данные поместите в табл. 4.2.

**Таблица 4.2.** Формантные частоты твердых губных согласных в зависимости от гласных

	БУ	БО	БА	БЭ	БЫ
$F_2^r$					
$F_2^c$					

Постройте график зависимости частоты губных согласных от гласного контекста в виде графика, общий вид которого представлен на рис. 4.9.



**Рисунок 4.9.** Координатная плоскость для графика, содержащего зависимость частоты губных согласных от гласного контекста

В результате получится кривая зависимости частоты второй форманты губных согласных от частоты гласных в слоге. Для определения коэффициента коартикуляции проведите на графике аппроксимирующую прямую для этой кривой и измерьте угол  $\alpha$  между этой прямой и осью ординат. Посчитайте коэффициент коартикуляции «губной согласный – гласный»  $K_{гс}$  по формуле:  $K_{гс} = \operatorname{tg} \alpha$ .

**Задание 4.3.2.** Определите и постройте график зависимости частоты форманты **твердого зубного согласного** от частоты гласного в слоге, рассчитайте коэффициент коартикуляции «зубной твердый согласный – гласный».

Для этого в программе *Praat* откройте файл *3m.wav*, записанный мужским голосом (или *3f.wav* – женским голосом), с записью слогов «ду – до – да – дэ – ды». По спектру Фурье определите величину второй форманты согласного ( $F_2^c$ ) и гласного звука ( $F_2^r$ ). Полученные данные поместите в табл. 4.3.



**Таблица 4.3.** Формантные частоты твердых зубных согласных в зависимости от гласных

	ДУ	ДО	ДА	ДЭ	ДЫ
$F_2^r$					
$F_2^c$					

Используя данные табл. 4.3, постройте кривую зависимости частоты **второй форманты** зубных согласных от частоты гласного на графике, аналогичном графику на рис. 4.9. Для определения коэффициента коартикуляции проведите на графике аппроксимирующую прямую для этой кривой и измерьте угол  $\alpha$  между этой прямой и осью ординат. Посчитайте коэффициент коартикуляции «зубной согласный – гласный»  $K_{зс}$  по формуле  $K_{зс} = tg \alpha$ .

**Задание 4.3.3.** Определите и постройте график зависимости частоты форманты **твердого нёбного согласного** от частоты гласного в слоге, рассчитайте коэффициент коартикуляции «нёбный твердый согласный – гласный».

Для этого в программе *Praat* откройте файл *4m.wav*, записанный мужским голосом (или *4f.wav* – женским голосом), с записью слогов «гу – го – га – гэ – гы». По спектру Фурье определите величину **второй форманты** согласного ( $F_2^c$ ) и гласного звука ( $F_2^r$ ). Полученные данные поместите в табл. 4.4.

**Таблица 4.4.** Формантные частоты твердых нёбных согласных в зависимости от гласных

	ГУ	ГО	ГА	ГЭ	ГЫ
$F_2^r$					
$F_2^c$					

Используя данные табл. 4.4, постройте кривую зависимости частоты второй форманты твердых нёбных согласных от частоты гласного на графике, аналогичном графику на рис. 4.9. Для определения коэффициента коартикуляции проведите на графике аппроксимирующую прямую для этой кривой и измерьте угол  $\alpha$  между этой прямой и осью ординат. Посчитайте коэффициент коартикуляции «нёбный согласный – гласный»  $K_{ас}$  по формуле:  $K_{ас} = tg \alpha$ .

**Задание 4.3.4.** Определите и постройте график зависимости частоты форманты **мягкого согласного** от частоты гласного в слоге, рассчитайте коэффициент коартикуляции «мягкий согласный – гласный».

Для этого в программе *Praat* откройте файл *5m.wav*, записанный мужским голосом (или *5f.wav* – женским голосом), с записью слогов «дю – дё – дя – де – ди». По спектру Фурье определите величину **второй форманты** согласного ( $F_2^c$ ) и гласного звука ( $F_2^r$ ). Полученные данные поместите в табл. 4.5.

**Таблица 4.5.** Формантные частоты мягких согласных в зависимости от гласных

	ДЮ	ДЁ	ДЯ	ДЕ	ДИ
$F_2^r$					
$F_2^c$					

Используя данные табл. 4.5, постройте кривую зависимости частоты второй форманты мягких согласных от частоты гласного на графике, аналогичном графику на рис. 4.9. Для определения коэффициента коартикуляции проведите на графике аппроксимирующую прямую для этой кривой и измерьте угол  $\alpha$  между этой прямой и осью абсцисс ( $F_2^r$ ). Посчитайте коэффициент коартикуляции «мягкий согласный – гласный»  $K_{мс}$  по формуле:  $K_{мс} = tg \alpha$ .

**Задание 4.4.** Проведите аналогичные исследования для своего собственного голоса и сделайте выводы.

**Задание 4.5.** В качестве отчета по лабораторной работе представить: заполненные таблицы; построенные графики; сравнительный анализ результатов для двух мужских (женских) голосов.

**Задание 4.6.** Если у вас осталось свободное время, исследуйте другие функциональные возможности программы *Praat* самостоятельно.

## Лабораторная работа №4. Часть 2

**Задание 4.7.** Запустите программу *Praat*, вспомните ее функциональные возможности и содержание меню программы, основные команды, назначение кнопок на панели инструментов.

**Задание 4.8.** Постройте формантные траектории и траекторию частоты основного тона для одной фразы, произнесенной мужским голосом.

**Задание 4.8.1.** Откройте файл *7m.wav* с записью фразы: «А я уеду в Ялту!».

**Задание 4.8.2.** Для построения формантных траекторий по спектру Фурье определите значение частоты первой ( $F_1$ ) и второй форманты ( $F_2$ ) для каждого звука фразы. Величины формантных частот отмечайте точками на графике (рис. 4.10), затем соедините точки и получите две формантные траектории.



**Рисунок 4.10.** Координатная плоскость для значений частот первой ( $F_1$ ) и второй форманты ( $F_2$ ) для фразы «А я уеду в Ялту!»

**Задание 4.8.3.** Сравните формантные траектории, построенные на графике на рис. 4.10, с приближенными формантными траекториями, которые показаны на спектрограмме Фурье в программе *Praat* (для этого в меню *Formant* должен быть выбран пункт *Show formants*) с помощью красных точек. Сделайте выводы.

**Задание 4.8.4.** Постройте траекторию частоты основного тона по двумерному представлению автокорреляционной функции, отмечая точками на графике, аналогичном графику на рис. 4.10, величину частоты основного тона  $F_0$  в каждый отсчет времени. Для этого установите вычисление интенсивности  $A$  и частоты основного тона  $F_0$  в программе *Praat*: в меню *Intensity* выберите пункт *Show intensity*, а в меню *Pitch* – пункт *Show pitch* (при этом желательно, чтобы в меню *Formant* пункт *Show formants* не был выбран). Пропорционально перенесите значения интенсивности  $A$  и частоты основного тона  $F_0$  для каждого звука фразы на графике.

**Задание 4.8.5.** Сравните значения и динамику интенсивности  $A$  и частоты основного тона  $F_0$  для ударных и безударных гласных звуков заданной фразы. Сделайте выводы.

**Задание 4.8.6.** Вычислите средние значения интенсивности  $A$  и частоты основного тона  $F_0$  для всей фразы, произнесенной мужским голосом:  $A_{\text{ср}}^{\text{м}}$  и  $F_{\text{ср}}^{\text{м}}$ .

**Задание 4.9.** Постройте формантные траектории и траекторию частоты основного тона для одной фразы, произнесенной женским голосом.

**Задание 4.9.1.** Для этого откройте файл *7f.wav* и выполните для нее задания, аналогичные заданиям 4.8.2 – 4.8.5. В результате должны быть построены соответствующие графики.

**Задание 4.9.2.** Вычислите средние значения интенсивности  $A$  и частоты основного тона  $F_0$  для всей фразы, произнесенной женским голосом:  $A_{\text{ср}}^{\text{ж}}$  и  $F_{\text{ср}}^{\text{ж}}$ .

**Задание 4.10.** Постройте формантные траектории и траекторию частоты основного тона для одной фразы, произнесенной своим голосом.

**Задание 4.10.1.** Для этого запишите с помощью микрофона в новый файл произнесенную своим голосом фразу «А я еду в Ялту!» и выполните для нее задания, аналогичные заданиям 4.8.2 – 4.8.5. В результате должны быть построены соответствующие графики.

**Задание 4.10.2.** Вычислите средние значения интенсивности  $A$  и частоты основного тона  $F_0$  для всей фразы, произнесенной женским голосом:  $A_{cp}^я$  и  $F_{0cp}^я$ .

**Задание 4.11.** Сравните результаты заданий 4.8 – 4.10 и сделайте выводы.

**Задание 4.12\*.** Исследуйте формантные характеристики гласных фонем русской речи: измерьте форманты гласных фонем мужского и женского голосов и отобразите их на формантной плоскости.

**Примечание.** В данной части лабораторной работы необходимо проанализировать второй из двух голосов: мужской либо женский – в зависимости от того, какой голос вы анализировали при выполнении первой части. При этом необходимо осуществить сравнительный анализ мужского и женского голосов.

**Задание 4.12.1\*.** Для этого выполните действия, описанные в задании 4.2 (части 1) данной лабораторной работы. При этом запись своего голоса производить не требуется, а осуществлять сравнительный анализ получаемых при выполнении данного задания результатов с результатами 1-й части лабораторной работы. Для ускорения процесса определения значений формант попробуйте использовать пункт меню *Formant* -> *Formant report*.

**Задание 4.12.2\*.** Поместите значения формант гласных анализируемого голоса в таблицу, аналогичную табл. 4.1.

**Задание 4.12.3\*.** На формантной плоскости отобразите данные о значениях формант мужского и женского голосов, обозначив их линиями разного типа (например, сплошная линия – мужской голос, пунктирная линия – женский). При этом используйте результаты 1-й части данной лабораторной работы.

**Задание 4.12.4\*.** Сравните данные табл. 4.1, 4.2 и таблицы, созданной при выполнении задания 4.12.2, а также соответствующие графики. Сделайте выводы.

**Задание 4.13\*.** Исследуйте формантные характеристики коартикуляции «согласный – гласный» для второго голоса.

**Задание 4.13.1\*.** Для этого выполните действия, описанные в задании 4.3 (части 1) данной лабораторной работы.

**Задание 4.13.2\*.** Результаты работы поместите в таблицы и на графики, аналогичные табл. 4.1–4.4 и графикам части 1 лабораторной работы.

**Задание 4.13.3\*.** Сравните данные всех таблиц и графиков. Сделайте выводы.

**Задание 4.13.4\*.** Сравните значения коэффициентов коартикуляции различных типов согласных для различных голосов (мужской, женский, свой). Сделайте выводы.

**Задание 4.14.** В качестве отчета по части 2 лабораторной работы следует представить: построенные графики; заполненные таблицы; сравнительный анализ результатов для различных голосов (мужской, женский, свой).

**Примечание.** Задания, помеченные звездочкой «\*», не являются основными.

#### 4.4. Вопросы для самопроверки

1. Сформулируйте свое определение понятия «речь».
2. Объясните понятие «звук» с акустической точки зрения.
3. Перечислите и поясните основные способы описания речевого сигнала.
4. Дайте определение форманты и поясните на примере.
5. Дайте определение формантной модели акустики речевого тракта и представьте ее в виде схемы.
6. Перечислите и поясните формантные параметры.
7. Какими значениями формантных параметров различаются гласные и согласные фонемы?
8. В чем заключается различие в распределении формантных частот гласных для мужских и женских голосов?
9. Какой формантный параметр играет наибольшую роль в разделении согласных по месту образования?
10. Каким образом рассчитывается коэффициент коартикуляции на основе анализа формантных характеристик речевого сигнала?
11. Какие согласные фонемы русской речи наибольшим образом подвержены коартикуляционным изменениям?
12. Какие согласные фонемы русской речи наименее подвержены коартикуляционным изменениям?

#### Литература

1. *Лобанов Б.М. 2006кн-Речев\_И\_ИС*  
Лобанов, Б. М. Речевой интерфейс интеллектуальных систем : учеб. пособие / Б. М. Лобанов, О. Е. Елисеева ; под науч. ред. В. В. Голенкова . – Минск : БГУИР, 2006. – 152 с.
2. *Рылов А.С. 2003кн-Анализ\_Р*  
Рылов, А. С. Анализ речи в распознающих системах / А. С. Рылов. – Минск : Бестпринт, 2003. – 264 с.
3. *Рабинер Л.Р. 1981кн-Цифро\_О\_Р\_С*  
Рабинер, Л. Р. Цифровая обработка речевых сигналов / Л. Р. Рабинер, Р. В. Шафер ; пер. с англ. – М. : Радио и связь, 1981. – 495 с.
4. *Фант Г. 1964кн-Акуст\_Т\_Р*  
Фант, Г. Акустическая теория речеобразования / Г. Фант ; пер. с англ.; под ред. В. С. Григорьева. – М. : Наука, 1964. – 284 с.
5. *Радзишевский А. 2006кн-Основ\_А\_и\_Ц\_3*  
Радзишевский, А. Основы аналогового и цифрового звука / А. Радзишевский. – М. : Вильямс, 2006. – 288 с.
6. *Лобанов Б.М. 2002ст-Комп\_К\_П\_Г*  
Лобанов, Б. М. Компьютерное «клонирование» персонального голоса и речи / Б. М. Лобанов // Новости искусственного интеллекта. – 2002. – №5. – С. 35–39.
7. *Цирульник Л.И. 2006ст-Автом\_С\_К*  
Цирульник, Л. И. Автоматизированная система клонирования фонетико-акустических характеристик речи / Л.И. Цирульник // Информатика. – 2006. – № 2(10). – С. 46–55.
8. *Печенева Т.А. 2004сост-Виды\_Я\_Р*  
Виды языкового разбора : краткий комплексный словарь-справочник / Авт.-сост. Т. А. Печенева. – Минск : Изд. центр БГУ, 2004. – 99 с.
9. *Зиндер Л.Р. 1979кн-Общая\_Ф*  
Зиндер, Л. Р. Общая фонетика / Л. Р. Зиндер. – М. : Высш. шк., 1979. – 312 с.

Учебное издание

**Елисеева Ольга Евгеньевна**

**РЕЧЕВОЙ ИНТЕРФЕЙС. ЛАБОРАТОРНЫЙ ПРАКТИКУМ**

Учебно-методическое пособие

В 2-х частях

Часть 1

Редактор *Т. Н. Крюкова*  
Корректор *М. В. Тезина*  
Дизайн обложки *А. А. Макаров*

Подписано в печать 31.01.2008. Формат 60x84 1/8. Бумага офсетная. Гарнитура «Arial».  
Печать ризографическая. Усл. печ. л. 5,58. Уч.-изд. л. 5,0. Тираж 150 экз. Заказ 233.

Издатель и полиграфическое исполнение: Учреждение образования  
«Белорусский государственный университет информатики и радиоэлектроники»  
Лицензия на осуществление издательской деятельности № 02330/0056964 от 1.04.2004.  
Лицензия на осуществление полиграфической деятельности № 02330/0131666 от 30.04.2004.  
220013, Минск, П. Бровки, 6.