

Рис. 3.2.2

а) Изобразить цифровую структурную схему, соответствующую данной системе в каскадной и прямой формах.

б) Определить, устойчива ли эта система? Привести пример.

2.14. Для системы на рис. 3.2.2:

а) Записать разностное уравнение, описывающее эту систему.

б) Определить передаточную функцию для данной системы.

2.15. Определить  $a_1$ ,  $a_2$  и  $a_3$  через  $b_1$  и  $b_2$  так, чтобы обе системы на рис. 3.2.3 обладали одной и той же передаточной функцией.

2.16. Передаточная функция обычного резонатора имеет вид

$$H(z) = \frac{1 - 2e^{-aT} \cos(bT)z^{-1} + e^{-2aT}}{1 - 2e^{-aT} \cos(bT)z^{-1} + e^{-2aT}z^{-2}}$$

а) Найти полюса и нули  $H(z)$  и изобразить их на  $z$ -плоскости.

б) Найти импульсную характеристику для  $T=10^{-4}$ ;  $b=1000$ ;  $a=200$  и изобразить ее на графике.

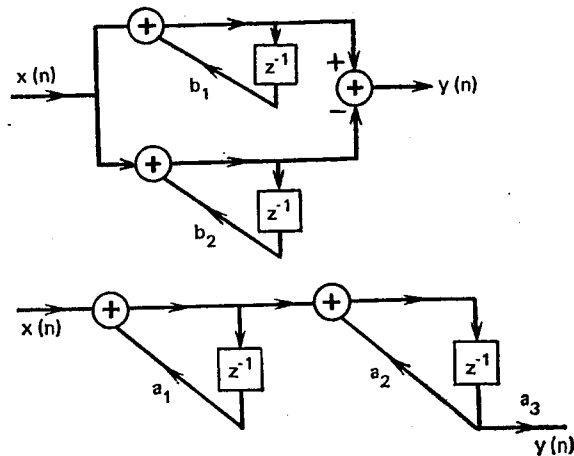


Рис. 3.2.3

2.17. Рассмотрим последовательность конечной длительности  $x(n) = \delta(n) + 0,5\delta(n-5)$ .

а) Найти  $z$ -преобразование и преобразование Фурье для  $x(n)$ .

б) Найти  $N$ -точечное ДПФ для  $x(n)$  при  $N=50$ ;  $10$ ;  $5$ .

в) Сравнить ДПФ при  $N=5$  и  $50$ .

г) Определить взаимосвязь между ДПФ и преобразованием Фурье.

2.18. Речевой сигнал дискретизирован с частотой 20 000 отсч./с. Для вычисления 1024-точечного ДПФ выделен сегмент, равный 1024 отсчетам.

а) Определить длительность сегмента.

б) Какое частотное разрешение достигается в ДПФ?

в) Как изменятся ответы на вопросы а) и б), если сегмент будет содержать 512 отсчетов (для вычисления ДПФ сегмент дополняется 512 нулями)?

## Цифровые модели речевых сигналов

### 3.0. Введение

Для того чтобы научиться применять методы цифровой обработки сигналов в задачах связи, надо хорошо представлять основные положения как теории речеобразования, так и теории цифровой обработки сигналов. В данной главе приведен обзор положений акустической теории речеобразования и показано, как из нее вытекают различные способы представления речи. Особенно большое внимание уделяется моделям в дискретном времени, с помощью которых описывается дискретизированный речевой сигнал. Эти модели служат основой применения методов цифровой обработки.

Назначение данной главы близко к назначению гл. 2, в которой излагаются предварительные сведения об изучаемых вопросах. Более детальное изложение можно найти в [1—5]. Повышенного внимания заслуживают книги Фанта [1] и Фланагана [2]. В [1] подробно рассмотрены положения акустической теории речеобразования и содержится большое количество данных о системах измерения акустических характеристик речи и ее моделях. В книге Фланагана, охватывающей более широкий круг вопросов, дается богатое множество способов физического моделирования процесса речеобразования и указываются пути их применения для представления и обработки речевых сигналов. Эти книги могут быть рекомендованы читателю, желающему глубоко изучить состояние проблемы.

Перед изложением акустической теории речеобразования и ее математических положений полезно познакомиться с различными типами звуков, из которых состоит речь. Поэтому глава начинается с краткого введения в акустическую фонетику, в котором приводятся основные классы фонем английского языка, рассматриваются особенности их произнесения. Далее излагаются основы акустической теории речеобразования. Рассматриваются вопросы распространения звуковых волн в голосовом тракте, электрические аналоги голосового тракта, квазистационарное поведение артикуляторного аппарата при произнесении протяжных звуков речи. Теория позволяет представить речевой сигнал в виде отклика нестационарной линейной системы (голосового тракта), возбуждаемой либо шумом, либо квазипериодической последовательностью импульсов. Такое представление применяется для получения моделей речевого сигнала в дискретном времени. Эти модели, разработанные на основе положений акустической теории, формируются с позиций теории цифровой обработки сигналов и используются далее при изложении основного предмета книги — методов цифровой обработки речевых сигналов.

### 3.1. Процесс образования речи

Речь состоит из последовательности звуков. Звуки и переходы между ними служат символическим представлением информации. Порядок следования звуков (символов) определяется правилами языка. Изучение этих правил и их роли в общении между людьми составляет предмет *лингвистики*, анализ и классификация самих звуков речи — предмет *фонетики*. В подробном изучении фонетики и лингвистики здесь нет необходимости. Однако при обработке речевых сигналов с целью повышения их информативного содержания либо для выделения содержащейся в сигнале информации полезно располагать как можно большим количеством сведений о структуре сигнала, например о способе кодирования информации в сигнале. Таким образом, прежде чем подробно рассматривать математические модели речеобразования, уместно обсудить основные группы звуков речи. Этим упоминанием о фонетике и лингвистике мы и ограничимся. Укажем, однако, что эта краткость не умаляет важности этих наук, особенно в области распознавания и синтеза речи.

#### 3.1.1. Механизм речеобразования

На рентгеновском снимке (рис. 3.1) показаны наиболее важные органы речеобразующей системы человека [6]. *Голосовой тракт*, который на рисунке обведен пунктиром, начинается с прохода между голосовыми связками, называемого *голосовой щелью*, и заканчивается у губ. Голосовой тракт, таким образом, состоит из *гортани* (от пищевода до рта) и рта, или *ротовой полости*. У взрослого мужчины общая длина голосового тракта составляет примерно 17 см. Площадь поперечного сечения голосового тракта, которая определяется положением языка, губ, челюстей и небной занавески, может изменяться от нуля (тракт полностью перекрыт) до примерно 20 см<sup>2</sup>. *Носовая полость* начинается у небной занавески и заканчивается ноздрями. При опущенной небной занавеске носовая полость акустически соединена с голосовым трактом и участвует в образовании носовых звуков речи. При изучении процесса речеобразования полезно изображать основные органы физической системы в таком виде, при котором становится ясной математическая сторона вопроса. На рис. 3.2 показано подробное схематическое изображение речеобразующей системы. Для полноты в диаграмму включены и такие органы, как легкие, бронхи и трахея, расположенные ниже гортани. Совокупность этих органов служит источником энергии для образования речи. Речь представляет собой акустическую волну, которая вначале излучается этой системой при выталкивании воздуха из легких и затем преобразуется в голосовом тракте. В качестве примера на рис. 3.3а показано речевое колебание, соответствующее фразе «Should we cha (se)», произнесенной мужским голосом. Основные особенности колеба-

ния легко объяснить на основе подробного анализа механизма образования речи.

Звуки речи могут быть разделены на три четко выраженные группы по типу возбуждения. *Вокализованные* звуки образуются

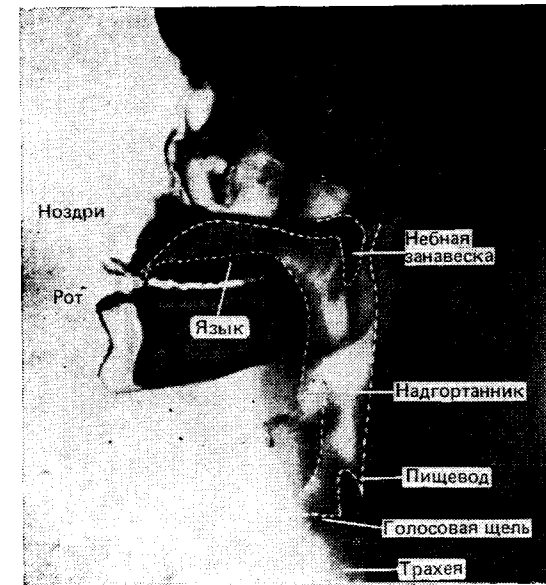


Рис. 3.1. Рентгеновский снимок речеобразующих органов человека [6]

проталкиванием воздуха через голосовую щель, при котором периодически напрягаются и расслабляются голосовые связки и возникает квазипериодическая последовательность импульсов потока воздуха, возбуждающая голосовой тракт. Вокализованные сегменты обозначены на рис. 3.3а знаками  $|U|$ ,  $|d|$ ,  $|\omega|$ ,  $|i|$ ,  $|e|$ . Фрикативные или невокализованные звуки генерируются при сужении

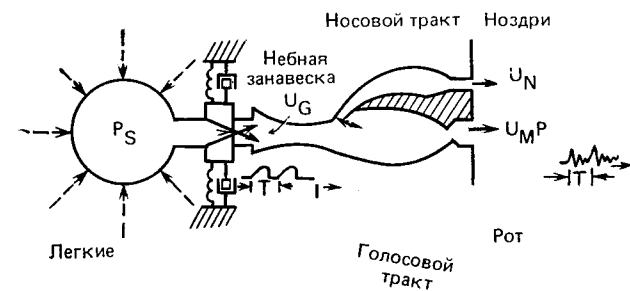


Рис. 3.2. Схематическое изображение речеобразующих органов человека [6]

голосового тракта в каком-либо месте (обычно в конце рта) и проталкивании воздуха через суженное место со скоростью, достаточно высокой для образования турбулентного воздушного потока. Таким образом, формируется источник широкополосного шу-

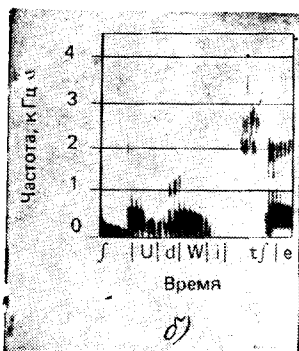
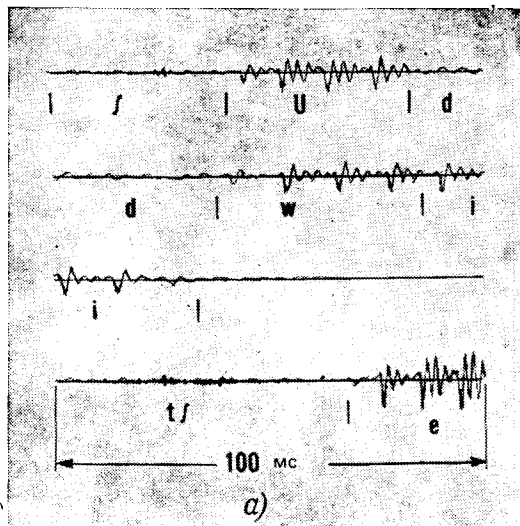


Рис. 3.3. Временная диаграмма (а) и спектрограмма (б) фразы «Should we cha(se)»

ма, возбуждающего голосовой тракт. На рис. 3.3а сегмент, обозначенный знаком  $|\int|$ , соответствует фрикативному звуку «sh». При произнесении *взрывных звуков* голосовой тракт полностью закрывается (обычно в начале голосового тракта). За этой смычкой возникает повышенное сжатие воздуха. Затем воздух внезапно высвобождается. Такое явление имеет место при произнесении звука, обозначенного на рис. 3.3а символом  $|\int|$ . Область малого уровня в конце третьей линии, которая предшествует шумоподобному участку колебания, соответствует периоду полного закрытия голосового тракта. Голосовой тракт и носовая полость показаны на рис. 3.2 в виде труб с переменной по продольной оси площадью поперечного сечения. При прохождении звуковых волн через эти трубы их частотный спектр изменяется в соответствии с частотной избирательностью трубы. Этот эффект похож на резонансные явления, происходящие в трубах органов и духовых музыкальных инструментов. При описании речеобразования резонансные частоты трубы голосового тракта называют *формантными частотами* или просто *формантами*. Формантные частоты зависят от конфигурации и размеров голосового тракта: произвольная форма тракта может быть описана набором формантных частот. Различные звуки образуются путем изменения формы голосового тракта. Таким образом, спектральные свойства речевого сигнала

изменяются во времени в соответствии с изменением формы голосового тракта.

Переменные во времени спектральные характеристики речевого сигнала с помощью звукового спектрографа могут быть высвечены в виде графика. Этот прибор позволяет получить двумерный график, называемый *спектрограммой*, на которой по вертикальной оси отложена частота, а по горизонтальной — время. Плотность зачернения графика пропорциональна энергии сигнала. Таким образом, резонансные частоты голосового тракта имеют вид затемненных областей на спектрограмме. Вокализированным областям сигнала соответствует появление четко выраженной периодичности временной зависимости, в то время как невокализованные интервалы выглядят почти сплошными. Спектрограмма фразы рис. 3.3а показана на рис. 3.3б. На спектрограмме отдельные участки помечены теми же символами, что и на рис. 3.3а, так что особенности сигнала во временной и частотной областях могут быть сопоставлены.

Звуковой спектрограф весьма долго служил основным инструментом исследования речевого сигнала, и хотя в настоящее время с помощью цифровой обработки можно получить более гибкие устройства визуального изображения (см. гл. 6), основные принципы спектрографа используются широко и в настоящее время. Хорошим пособием по спектрографическому представлению речи, не утратившим своего значения и до сих пор, является книга [8]. Хотя книга написана для обучения «чтению» спектрограмм, она является замечательным введением в акустическую фонетику.

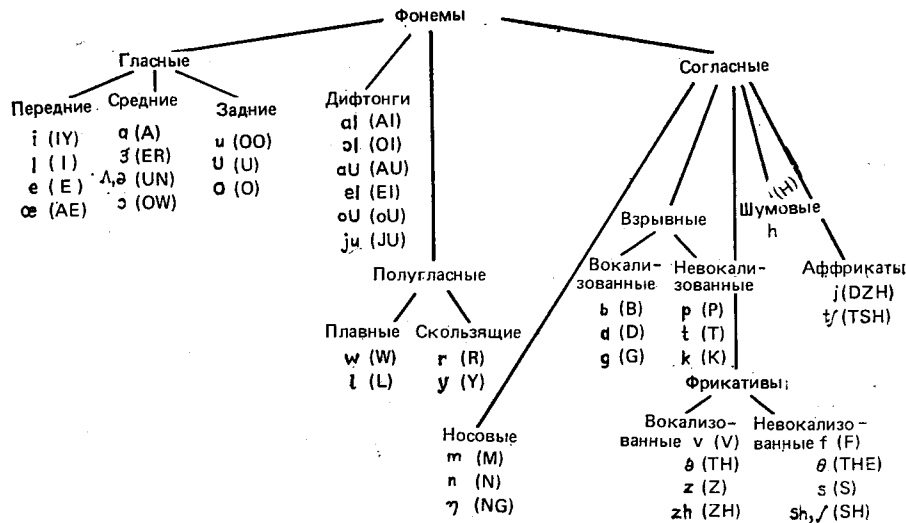
### 3.1.2. Акустическая фонетика

Многие языки, в том числе и английский, можно описать набором отдельных звуков или *фонем*. В частности, в американском произношении английского языка существует 42 фонемы, которые подразделяются на гласные, дифтонги, полугласные и согласные. Изучать фонему можно по-разному. Лингвисты, например, изучают отличительные характеристики фонем [9, 10]. Нам достаточно рассмотреть акустические свойства различных звуков, в том числе место и способ артикуляции, форму акустического колебания, характеристики спектрограмм. В табл. 3.1 приведены различные классы фонем английского языка в его американском произношении<sup>1</sup>. Четыре широких класса звуков образуют гласные, дифтонги, полугласные и согласные. Каждый из классов разбит на подклассы по способу и месту образования звука в голосовом тракте. Каждая фонема табл. 3.1 может быть отнесена к классу протяжных или кратковременных звуков. Протяжные звуки образуются при фиксированной (инвариантной во времени) форме голосового тракта, который возбуждается соответствующим источником. К этому классу относятся гласные, фрикативные (вокализованные и невокализованные) носовые согласные. Остальные звуки (дифтонги, полугласные, аффрикаты и взрывные согласные) произносятся при изменяющейся форме голосового тракта. Они образуют класс кратковременных звуков<sup>2</sup>.

<sup>1</sup> В таблице указаны как фонетическое, так и орфографическое представление фонем, которые используются далее во всей книге. (Прим. ред.)

<sup>2</sup> Имеется в виду, что при их произнесении решающую роль играет кратковременная динамика артикуляционных движений. (Прим. ред.)

Таблица 3.1



**Гласные.** Гласные образуются при квазипериодическом возбуждении голосового тракта неизменной формы импульсами воздуха, возникающими вследствие колебания голосовых связок. Как будет показано ниже, зависимость площади поперечного сечения голосового тракта от координаты (расстояния) вдоль его продольной оси определяет резонансные частоты тракта (форманты) и характер произносимого звука. Эта зависимость называется *функцией площади поперечного сечения*. Функция площади поперечного сечения для каждой гласной зависит в первую очередь от положения языка; вместе с тем на характер звука оказывают влияние положение челюстей, губ и, в меньшей степени, небной занавески. Например, при произнесении звука [a], как в слове «father», голосовой тракт открыт в начале, а в его конце тело языка образует сужение. Наоборот, при произнесении звука [i], как в слове «eve», язык образует сужение в начале голосового тракта и оставляет его открытым в конце. Таким образом, каждому гласному звуку может быть поставлена в соответствие форма голосового тракта (функция площади поперечного сечения), характерная для его произношения. Очевидно, что это соответствие неоднозначное, так как у разных дикторов голосовые тракты различны. Другим представлением гласного звука является его описание с помощью набора резонансных частот голосового тракта. Это описание также зависит от диктора. Петерсон и Барней [11] провели измерения формантных (резонансных) частот с помощью звукового спектрографа для гласных, произнесенных различными дикторами. Эти результаты приведены на рис. 3.4, где показан график зависимости частоты второй форманты от частоты первой форманты для некоторых гласных, произнесенных взрослыми дик-

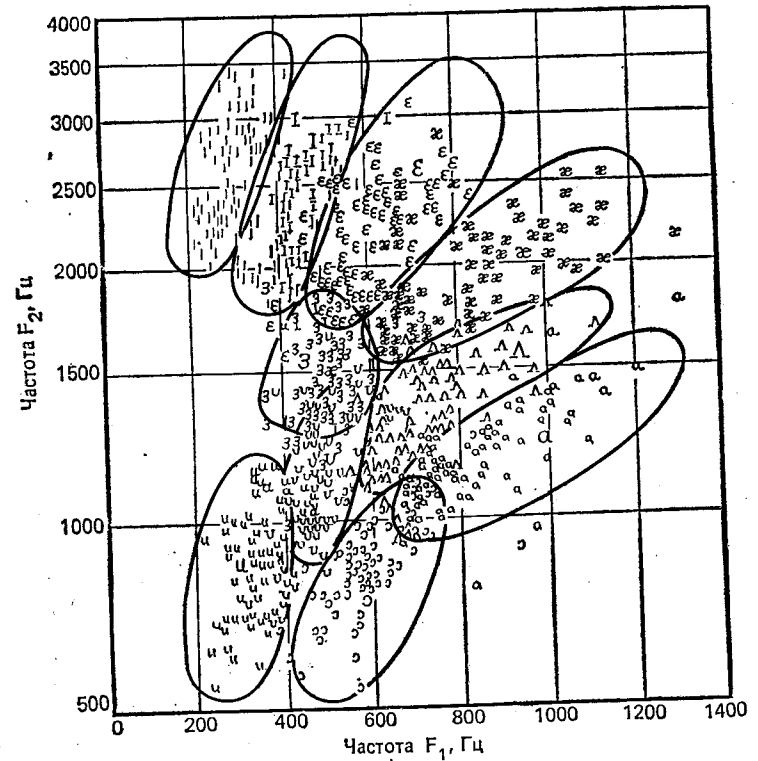


Рис. 3.4. Зависимость частоты второй форманты от частоты первой форманты для гласных, произнесенных разными дикторами [11]

торами и детьми. Эллипсы на рис. 3.4 ограничивают область изменения формантных частот для каждой из гласных. В табл. 3.2 приведены средние значения первых трех формантных частот для гласных, произнесенных мужскими голосами. Хотя существует

Таблица 3.2  
Средние значения формантных частот для гласных [11]

Письменный символ	Транскрипция	Типичное слово	$F_1$	$F_2$	$F_3$
IY	i	beet	270	2290	3010
I	ɪ	bit	390	1990	2550
E	e	bet	530	1840	2480
AE	æ	bat	660	1720	2410
UH	ʌ	but	520	1190	2390
A	a	hot	730	1090	2440
OW	ɔ	bought	570	840	2410
U	ʊ	foot	440	1020	2240
OO	u	boot	300	870	2240

большой разброс формантных частот, данные табл. 3.2 являются полезной характеристикой гласных. На рис. 3.5 приведен график зависимости частоты второй форманты от частоты первой форманты для гласных табл. 3.2. В верхнем левом углу так называемого треугольника гласных расположена гласная  $|i|$  с низкой частотой

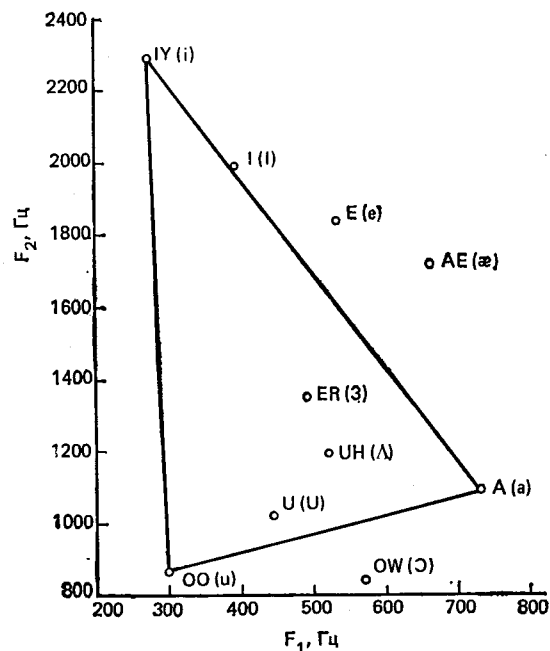


Рис. 3.5. Треугольник гласных

первой форманты и высокой частотой второй форманты. В нижнем левом углу расположена гласная  $|u|$  с низкими частотами первой и второй формант. В третьей вершине треугольника находится гласная  $|a|$  с высокой частотой первой форманты и низкой частотой второй форманты. Далее будет показано, как форма голосового тракта влияет на частоты формант гласных.

На рис. 3.6 приведены акустические колебания и спектрограммы для всех гласных английского языка. На спектрограммах четко выделяются различные резонансные области, характерные для каждой гласной. Акустические колебания, иллюстрируя периодичность вокализованных звуков, позволяют также путем анализа одного периода выявить грубые спектральные характеристики. Например, акустическое колебание звука  $|i|$  состоит из низкочастотного затухающего колебания, на которое накладывается относительно высокочастотная составляющая. Это соответствует низкой частоте первой форманты и высоким частотам второй и третьей формант (см. табл. 3.2). Два резонанса, расположенных на близких частотах, расширяют спектр колебания. Наоборот, в акусти-

ческом колебании гласной  $|u|$  энергия высокочастотных составляющих относительно мала, что соответствует низким частотам первой и второй формант. Подобный анализ может быть проведен для всех гласных, акустические колебания которых приведены на рис. 3.6.

**Дифтонги.** Дифтонгом называется участок речи, соответствующий одному слогу, который начинается с одной гласной и затем постепенно переходит в другую. На основе этого определения в американском произношении можно выделить шесть дифтонгов:  $|eI|$  (как в слове «bay»),  $|oU|$  (как в слове «boat»),  $|aU|$  (как в слове «how»),  $|oI|$  (как в слове «boy»),  $|aI|$  (как в слове «by») и  $|ju|$  (как в слове «you»).

Дифтонги образуются путем плавного изменения формы голосового тракта. Для иллюстрации этого положения на рис. 3.7 показана временная зависимость частоты второй форманты от частоты первой форманты для дифтонгов [12]. Стрелками показаны направления изменения формантных частот во времени. Пунктирными линиями обведены средние значения формант для гласных. Дифтонги можно описать изменением во времени функции площади поперечного сечения голосового тракта от значения, соответствующего первой гласной, до значения, соответствующего второй гласной дифтонга.

**Полугласные.** Группу звуков, содержащих  $|w|$ ,  $|l|$ ,  $|r|$  и  $|y|$ , описать довольно трудно. Эти звуки называются полугласными, так как по своим свойствам они напоминают гласные звуки. Обычно их характеризуют плавным изменением функции площади поперечного сечения голосового тракта между смежными фонемами. Таким образом, акустические характеристики этих звуков существенно зависят от произносимого текста. Нам удобно рассматривать эти звуки как переходные, сходные с гласными. Их структура близка к структуре гласных и дифтонгов. Пример сигнала, соответствующего полугласному звуку  $|w|$ , показан на рис. 3.3.

**Носовые звуки.** Носовые согласные  $|m|$ ,  $|n|$  и  $|ŋ|$  образуются при голосовом возбуждении. В полости рта при этом возникает полная смычка. Небная занавеска опущена, поэтому поток воздуха проходит через носовую полость и излучается через нос. Полость рта, которая вначале закрыта, акустически соединена с гортанью. Таким образом, рот служит резонансной полостью, в которой задерживается часть энергии при определенных частотах воздушного потока. Эти резонансные частоты соответствуют антирезонансам или нулям передаточной функции тракта речеобразования [2]. Более того, для носовых согласных и гласных (т. е. гласных, расположенных перед носовыми согласными) характерны менее выраженные резонансы, чем для гласных. Расширение резонансных областей происходит из-за того, что внутренняя поверхность носового тракта напрягается и при этом носовая полость имеет большое отношение площади поверхности к площади поперечного сечения. Вследствие этого потери за счет теплопроводности и вязкости оказываются большими, чем обычно.

Рис. 3.6. Акустические колебания (а, б)

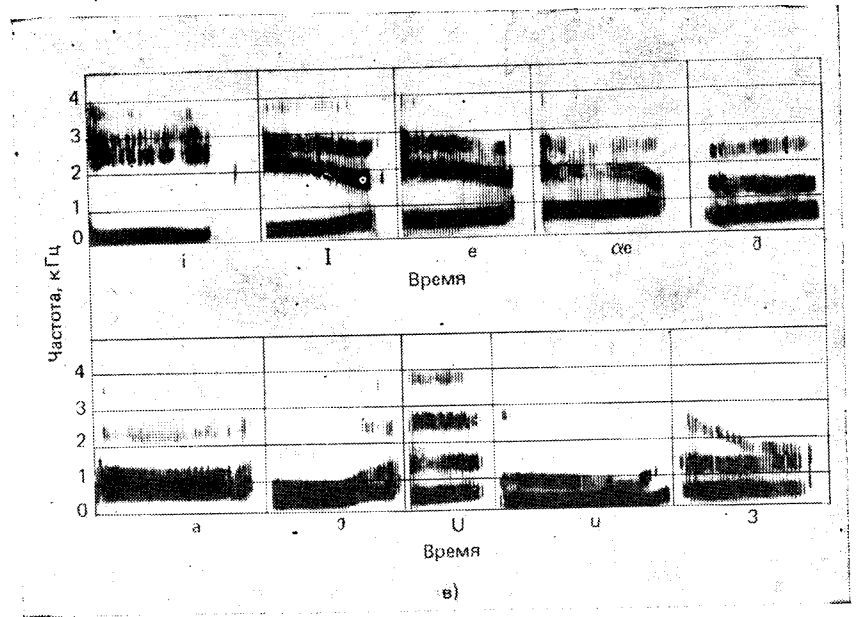
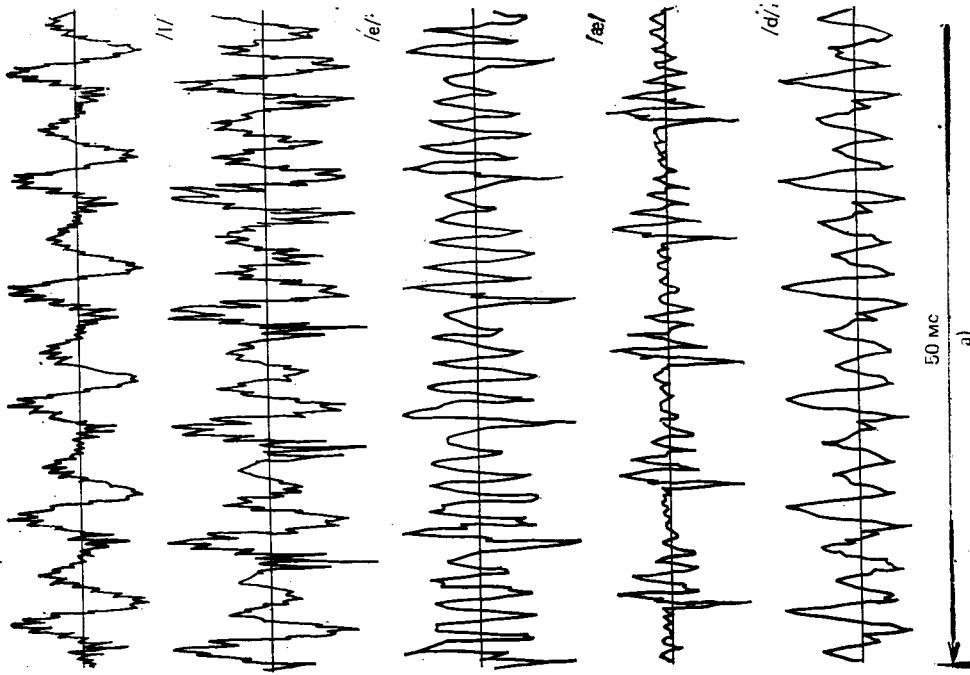
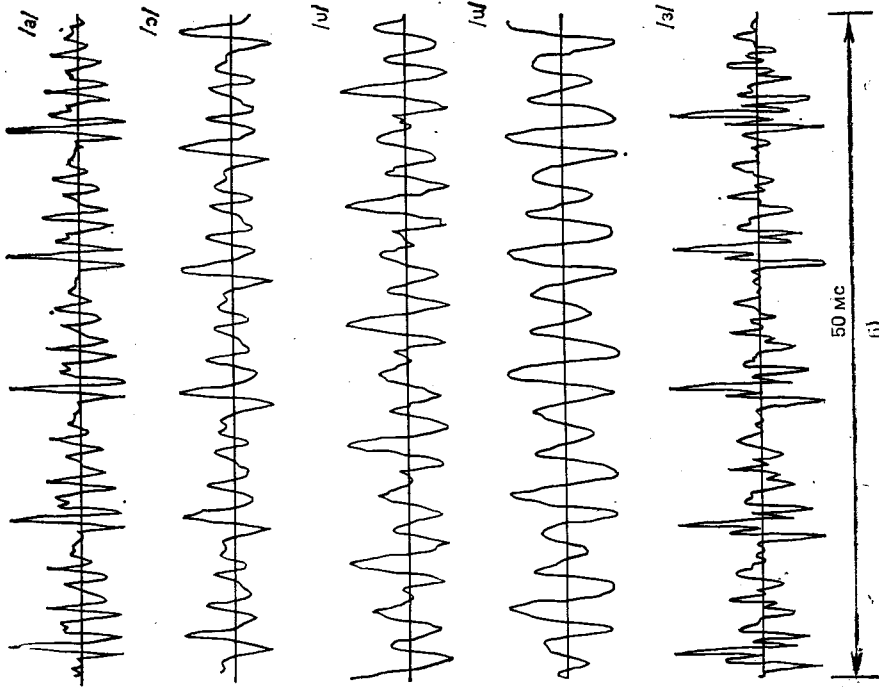


Рис. 3.6.в. Спектрограммы в некоторых гласных английского языка (американское произношение)

Три носовых согласных различаются местом расположения полной смычки. При произнесении звука  $[m]$  смычка образуется между губами,  $[n]$  — у внутренней стороны зубов и  $[ŋ]$  — у небной занавески. На рис. 3.8 приведены типичные колебания и спектрограммы для двух носовых согласных в сочетании гласный — носовой согласный — носовой согласный — гласный. Из рисунка видно, что временные колебания согласных  $[m]$  и  $[n]$  очень похожи. Спектрограммы иллюстрируют подъем спектра на низких частотах и отсутствие четко выраженных резонансов в диапа-

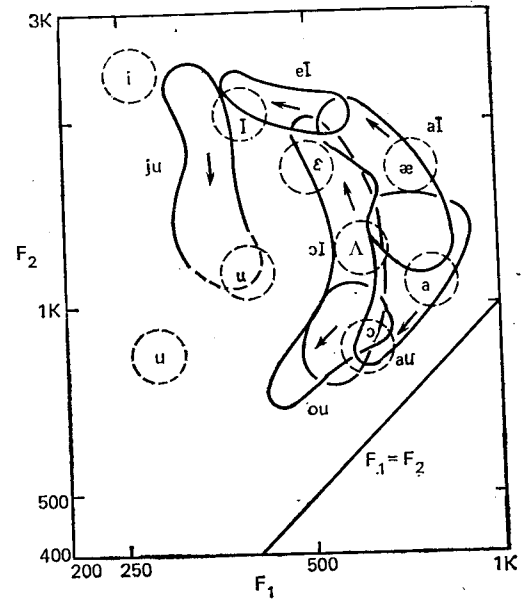


Рис. 3.7. Временные изменения двух первых формант дифтонгов [27]

зоне средних частот. Это происходит вследствие взаимного влияния резонансов и антирезонансов, образующихся за счет взаимодействия полостей носа и рта [13].

**Глухие фрикативные звуки.** Глухие фрикативные звуки  $|f|$ ,  $|\theta|$ ,  $|s|$  и  $|sh|$  образуются путем возбуждения голосового тракта турбулентным воздушным потоком, возникающим в области смычки голосового тракта. Расположение смычки характеризует тип фрикативного звука. При произнесении звука  $|f|$  смычка возникает около губ,  $|\theta|$  — около зубов,  $|s|$  — в середине полости рта и  $|sh|$  — в конце полости рта. Таким образом, система образования глухих фрикативных звуков содержит источник шума, расположенный в области смычки, которая разделяет голосовой тракт на две полости. Звуковая волна излучается через губы, т. е. через переднюю полость. Другая полость служит, как и в случае произнесения носовых звуков, для задерживания акустического потока, и таким образом в речеобразующем тракте возникают антирезонансы [2, 14]. На рис. 3.9 приведены колебания и спектрограммы фрикативных звуков  $|f|$ ,  $|s|$  и  $|sh|$ . Непериодическая структура возбуждения отчетливо видна на временных диаграммах колеба-

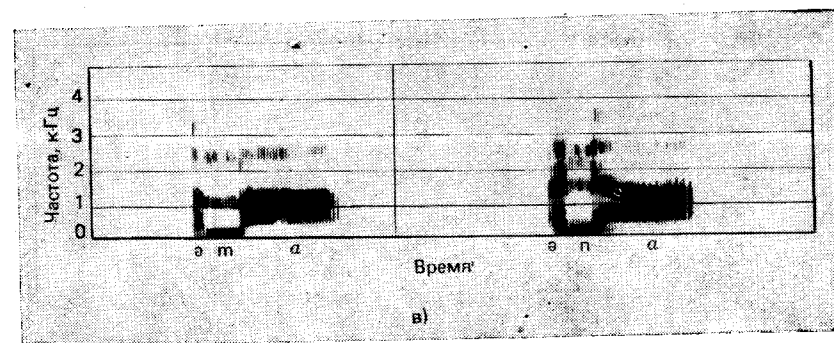
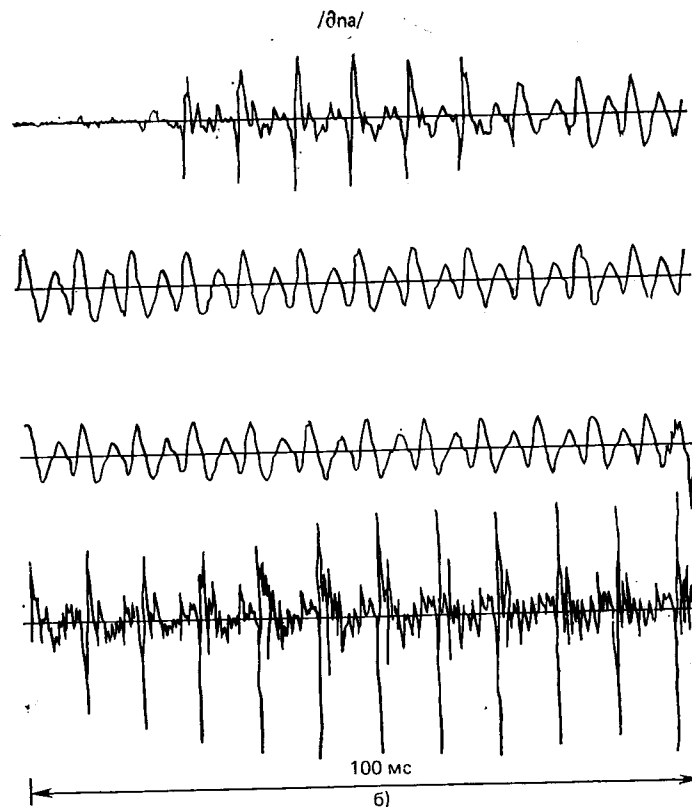
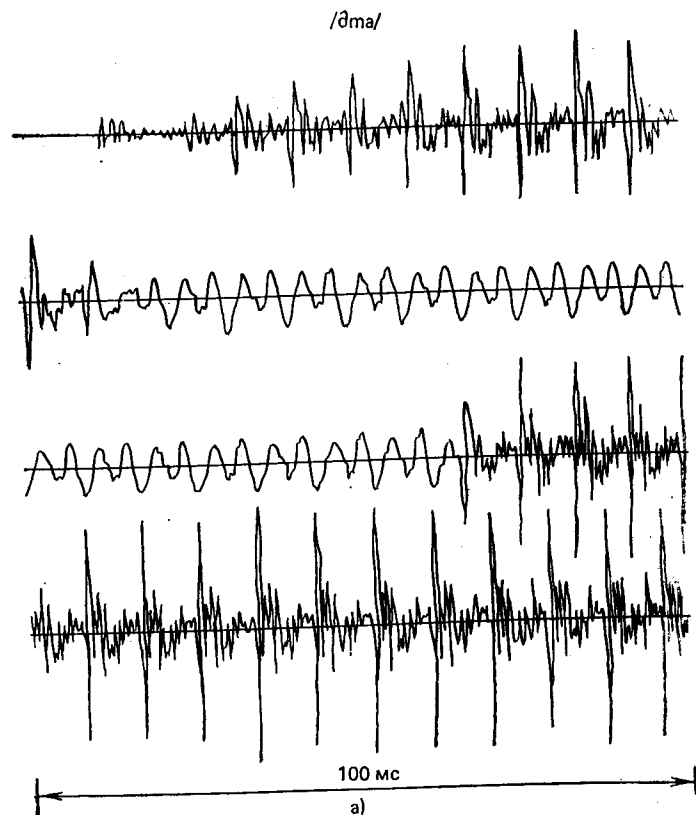


Рис. 3.8. Акустические колебания (а, б) и спектрограммы (в) для сочетаний  $|UH-M-A|$  и  $|UH-N-A|$

ний. Спектральные отличия фрикативных звуков легко определяются по спектрограммам.

**Звонкие фрикативные звуки.** Звонкие фрикативные звуки  $|v|$ ,  $|th|$ ,  $|z|$  и  $|zh|$  являются прототипами глухих звуков  $|f|$ ,  $|\theta|$ ,

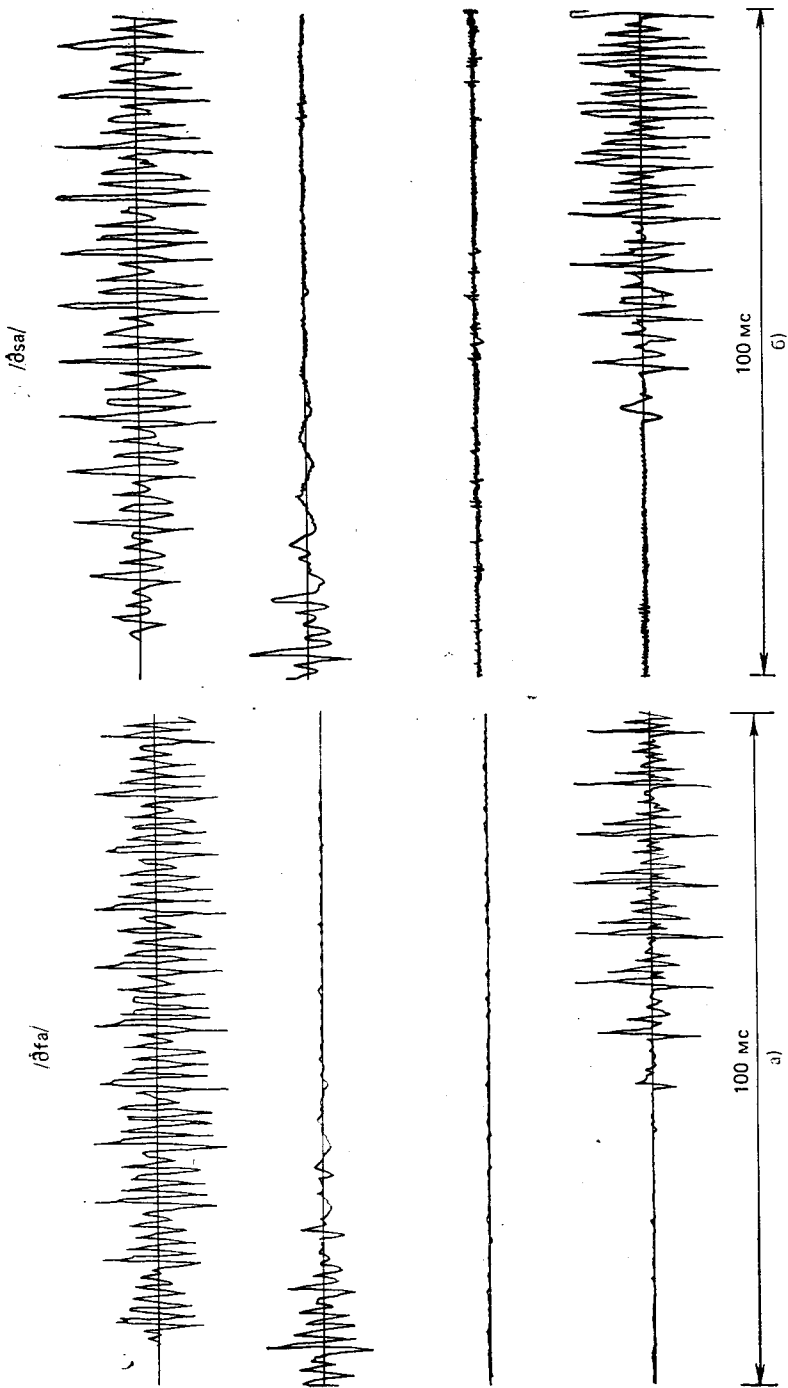


Рис. 3.9. Акустические колебания (а, б)

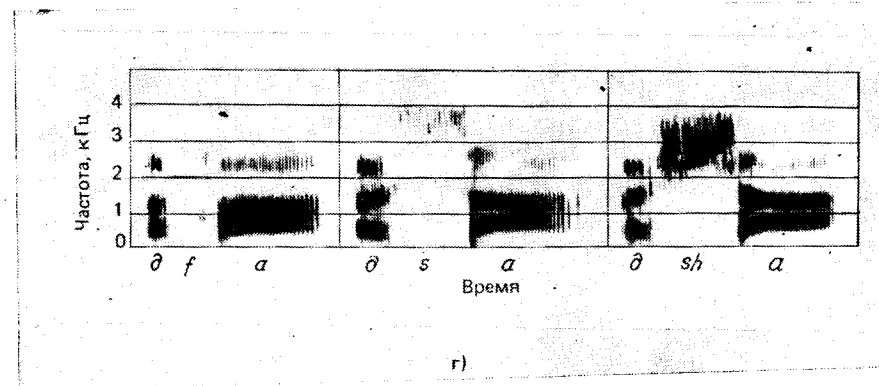
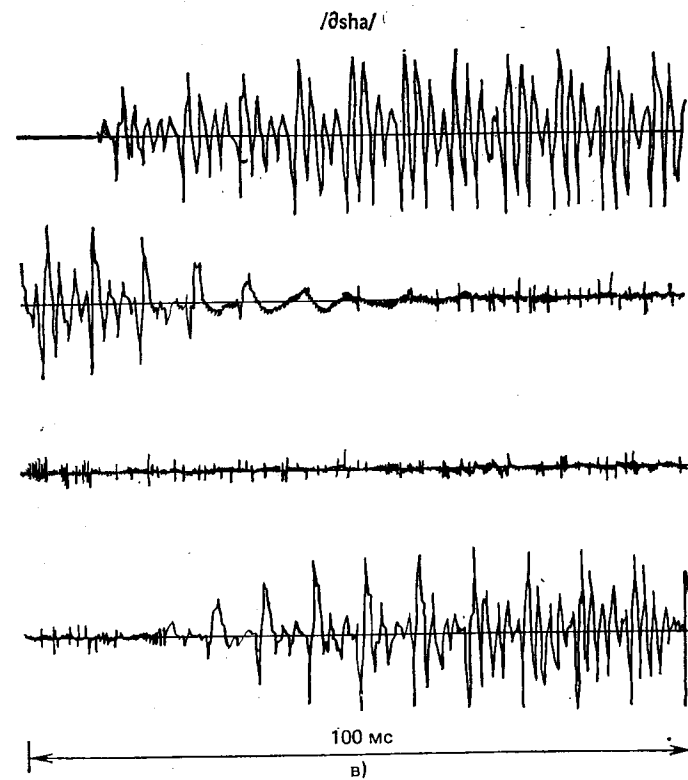


Рис. 3.9. Акустические колебания (б) и спектрограммы (в) для |UH-F-A| и |UH-SH-A|

|s| и |sh| соответственно. Место расположения смычки для этих пар звуков совпадает. Однако звонкие фрикативные отличаются от своих глухих аналогов тем, что при их образовании участвуют



два источника возбуждения. При образовании звонких звуков голосовые связки колеблются и, таким образом, один источник возбуждения находится в гортани. Однако, так как в голосовом тракте образуется смычка, поток воздуха в этой области становится турбулентным. Можно ожидать, что в спектре звонких фрикативных звуков будут две различные составляющие. Эти особенности возбуждения отчетливо видны на рис. 3.10, на котором приведены типичные колебания и спектры для нескольких звонких фрикативных звуков. Сходство структуры звонкого  $[v]$  и глухого  $[f]$  также легко установить путем сравнения соответствующих спектрограмм (рис. 3.9 и 3.10). Аналогично можно сравнить и спектрограммы звуков  $[sh]$  и  $[zh]$ .

**Звонкие взрывные согласные.** Звонкие взрывные согласные  $[b]$ ,  $[d]$  и  $[g]$  являются переходными непротяжными звуками. При их образовании голосовой тракт смыкается в какой-нибудь области полости рта. За смычкой воздух сжимается и затем внезапно высвобождается. При произнесении звука  $[b]$  смычка образуется между губами,  $[d]$  — с внутренней стороны зубов,  $[g]$  — вблизи небной занавески. В течение периода, когда голосовой тракт полностью закрыт, звуковые волны практически не излучаются через губы. Однако слабые низкочастотные колебания излу-

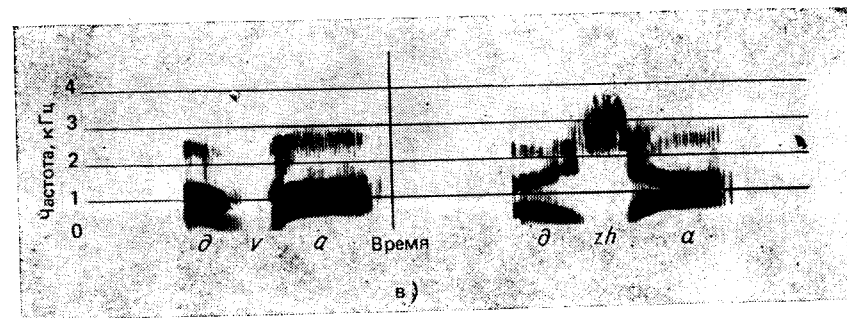
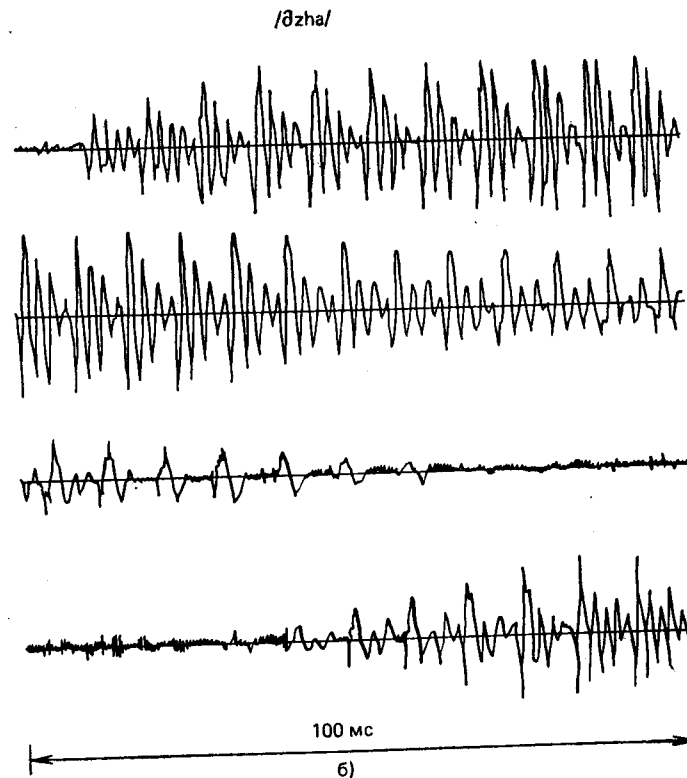
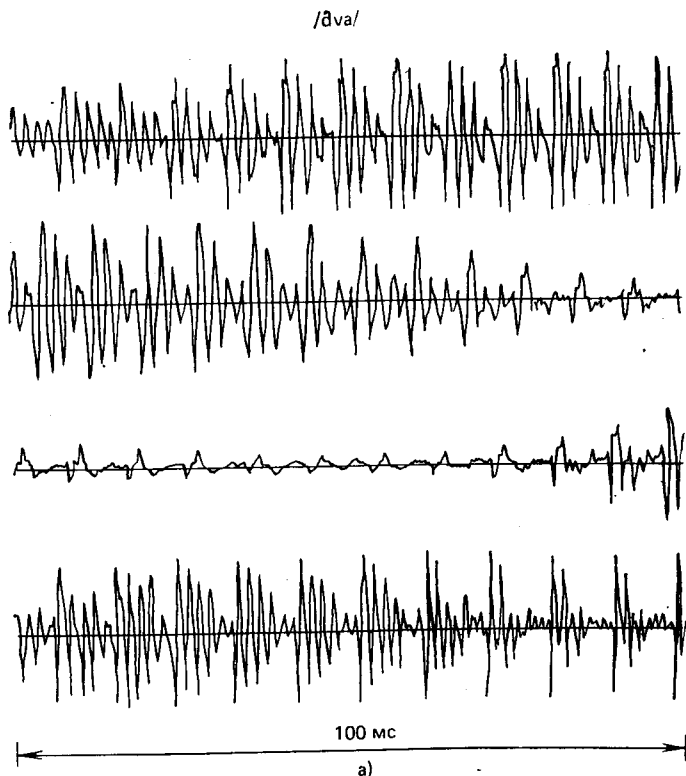


Рис. 3.10. Акустические колебания (а, б) и спектрограммы (в) для  $[UH-V-A]$  и  $[UH-ZH-A]$

чаются стенками горла (эту область иногда называют голосовым затвором — «voice bar»). Колебания возникают из-за того, что голосовые связки могут вибрировать даже тогда, когда голосовой тракт перекрыт.

Так как структура взрывных звуков изменчива, их свойства существенно зависят от последующего гласного [15]. В этой связи

характер временных колебаний несет мало сведений о свойствах этих согласных. На рис. 3.11 показаны временная диаграмма колебания и спектрограмма сочетания  $[UH-B-A]$ . Из временной диаграммы, соответствующей звуку  $[b]$ , видно лишь, что при его произнесении имеет место голосовое возбуждение и в сигнале отсутствуют высокочастотные составляющие.

**Глухие взрывные согласные.** Глухие взрывные согласные  $[p]$ ,  $[t]$  и  $[k]$  подобны своим звонким прототипам  $[b]$ ,  $[d]$  и  $[g]$ , но

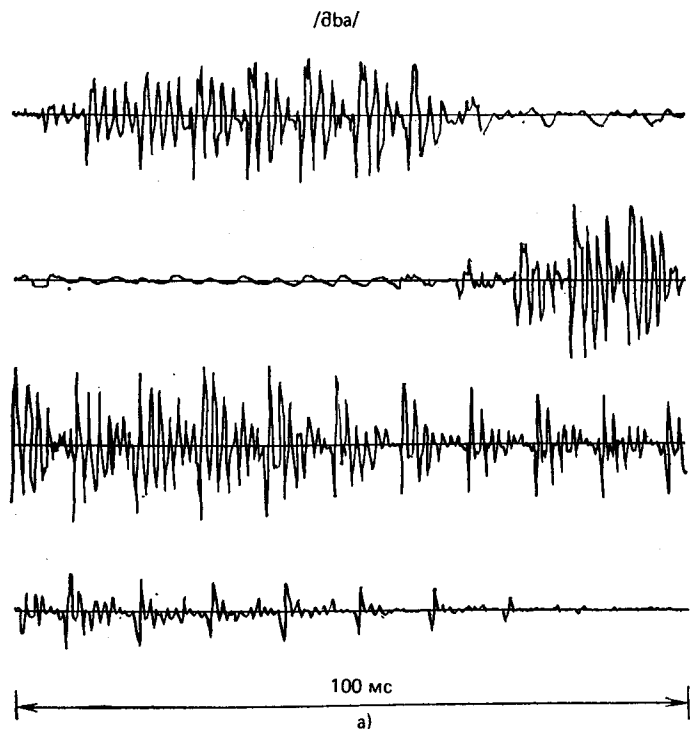


Рис. 3.11. Акустическое колебание (а) и спектрограмма (б) для  $[UH-B-A]$

имеют одно важное отличие. В течение периода полного смыкания голосового тракта голосовые связки не колеблются. После этого периода, когда воздух за смычкой высвобождается, в течение короткого промежутка времени потери на трение возрастают из-за внезапной турбулентности потока воздуха. Далее следует период придыхания (шумовой воздушный поток из голосовой щели возбуждает голосовой тракт). После этого возникает голосовое возбуждение.

На рис. 3.12 приведены временные колебания и спектрограммы согласных  $[p]$  и  $[t]$ . На рисунке отчетливо виден период смычки (интервал, в течение которого сжимается воздух за смычкой). Видно также, что длительность и частотный состав шума в периоды повышенного трения и придыхания существенно зависят от типа взрывной согласной.

**Аффрикаты и звук  $[h]$ .** Остальными согласными американского произношения являются аффрикаты  $[tʃ]$  и  $[dʒ]$  и фонема  $[h]$ . Глухая аффриката  $[tʃ]$  является динамичным звуком, который можно представить как сочетание взрывного  $[t]$  и фриктивного согласного  $[ʃ]$  (см. рис. 3.3а). Звонкий звук  $[dʒ]$  можно представить как сочетание взрывного  $[d]$  и фриктивного звука  $[ʒh]$ . Наконец, фонема  $[h]$  образуется путем возбуждения голосового тракта турбулентным воздушным потоком, т. е. без участия голосовых связок, но при возникновении шумового потока в голосовой щели<sup>1</sup>. Структура звука  $[h]$  не зависит от следующей за ним гласной. Поэтому голосовой тракт может перестраиваться для произнесения следующей гласной в процессе произнесения звука  $[h]$ .

## 3.2. Акустическая теория речеобразования

В предыдущем параграфе дано качественное описание звуков речи и способов их образования. В настоящем параграфе изучим математическое описание речеобразования, которое служит основой анализа и синтеза речи.

### 3.2.1. Распространение звуков

Понятие звука почти совпадает с понятием колебаний. Звуковые волны возникают за счет колебаний. Они распространяются в воздухе или другой среде с помощью колебаний частиц этой среды. Следовательно, образование и распространение звуков в голосовом тракте подчиняется законам физики. В частности, основные законы сохранения массы, сохранения энергии, сохранения количества движения вместе с законами термодинамики и механики жидкостей применимы к сжимаемому воздушному потоку с низкой вязкостью, который является средой распространения звуков речи. Используя эти основные физические законы, можно соста-

<sup>1</sup> Этот способ возбуждения характерен и для шепота.

вить систему дифференциальных уравнений в частных производных, описывающую движение воздуха в речеобразующей системе [16—20]. Составление и решение этих уравнений весьма затруднительны даже для простых предположений относительно формы голосового тракта и потерь энергии в речеобразующей системе. Полная акустическая теория должна учитывать следующие факторы:

- изменение во времени формы голосового тракта;
- потери энергии на стенках голосового тракта за счет вязкого трения и теплопроводности;
- мягкость стенок голосового тракта;
- излучение звуковых волн через губы;
- влияние носовой полости;
- возбуждение голосового тракта.

Построение акустической теории, охватывающей все эти факторы, выходит за рамки этой главы и, кроме того, создание такой теории пока еще невозможно. Дадим обзор этих явлений и ссылки на соответствующую литературу. Если подходящей литературы по какому-либо вопросу не имеется, придется ограничиться качественным описанием.

Простейшая физическая интерпретация системы речеобразования показана на рис. 3.13а. Голосовой тракт здесь представлен в виде неоднородной трубы с переменной во времени площадью по-

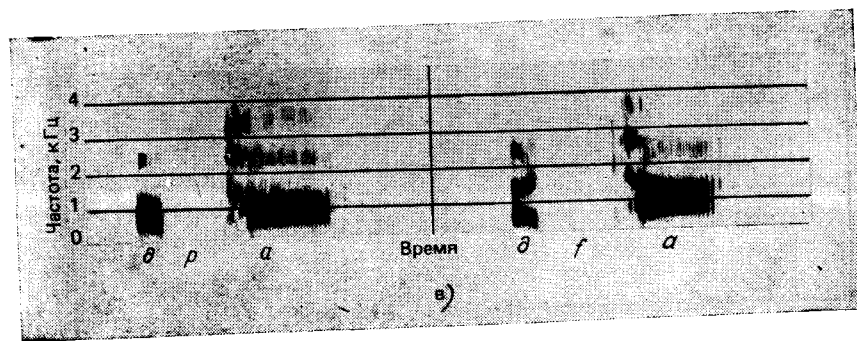
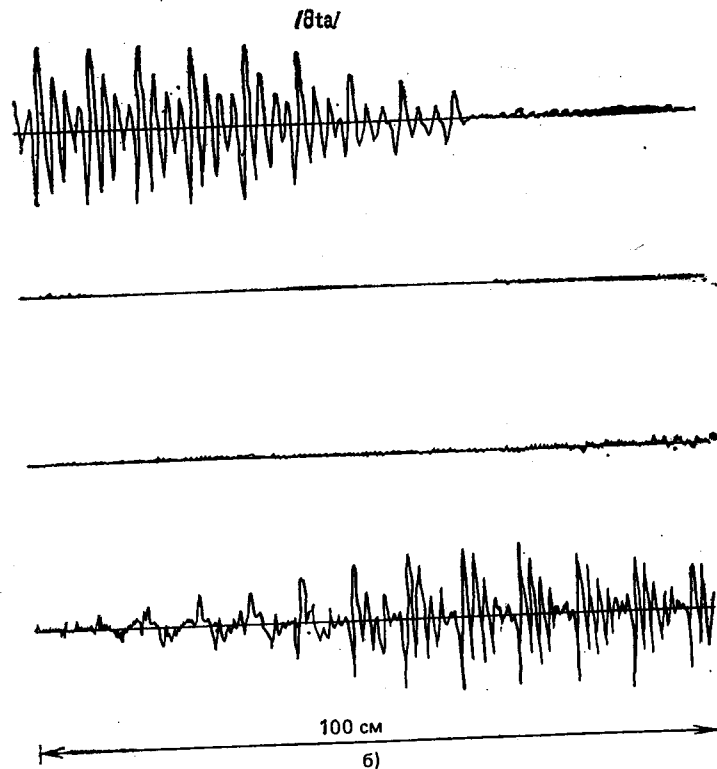
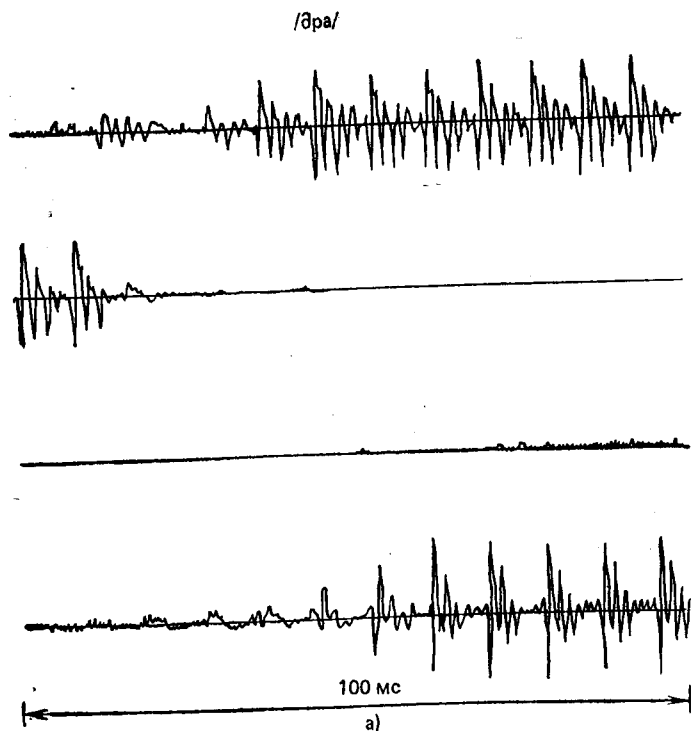


Рис. 3.12. Акустические колебания (а, б) и спектрограммы (в) для  $[UH-P-A]$  и  $[UH-T-A]$  поперечного сечения. Для колебаний, длина волны которых превышает размеры голосового тракта (это обычно имеет место на частотах ниже 4000 Гц), можно допустить, что вдоль продольной оси трубы распространяется плоская волна. Дальнейшее упрощение состоит в предположении отсутствия потерь на вязкость и тепло-

проводность как внутри воздушного потока, так и на стенках трубы. На основе законов сохранения массы, количества движения и энергии с учетом перечисленных допущений Портнов [18] показал, что звуковые волны в трубе удовлетворяют следующим уравнениям:

$$-\frac{\partial p}{\partial x} = \rho \frac{\partial (u/A)}{\partial t}; \quad (3.1a)$$

$$-\frac{\partial u}{\partial x} = \frac{1}{\rho c^2} \frac{\partial (pA)}{\partial t} + \frac{\partial A}{\partial t}, \quad (3.1б)$$

где  $p=p(x, t)$  — звуковое давление как функция  $x$  и  $t$ ;  $u=u(x, t)$  — скорость воздушного потока (volume velocity) как функция  $x$  и  $t$ ;  $\rho$  — плотность воздуха в трубе;  $c$  — скорость распространения

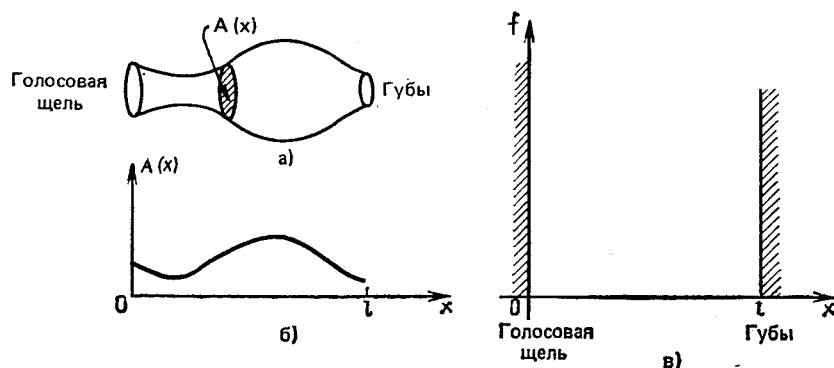


Рис. 3.13. Схематическое изображение голосового тракта (а), функция площади его поперечного сечения (б) и плоскость  $x-t$  для решения волнового уравнения (в)

звуча;  $A=A(x, t)$  — «функция площади», т. е. площадь поперечного сечения в направлении, перпендикулярном продольной оси трубы, как функция расстояния вдоль этой оси и времени. Сходная система уравнений была получена Сондхи [20].

Замкнутое решение уравнений (3.1) получить невозможно даже для простых форм трубы. Однако могут быть получены численные решения. Полное решение дифференциальных уравнений предполагает заданными давлением и скоростью потока для значений  $x$  и  $t$  в области голосовой щели и около губ, т. е. для получения решения должны быть заданы граничные условия у обоих концов трубы. Со стороны губ граничные условия должны отображать эффект излучения, а со стороны голосовой щели — характер возбуждения.

Кроме граничных условий необходимо задать функцию площади  $A(x, t)$ . На рис. 3.13б показана функция площади для трубы рис. 3.13а в некоторый момент времени. Для протяжных звуков можно предположить, что  $A(x, t)$  не изменяется во времени.

Однако это предположение неверно для непротяжных звуков. Подробные измерения  $A(x, t)$  весьма затруднительны и могут быть выполнены только для протяжных звуков. Одним из методов проведения таких измерений является рентгеновская киносъемка. Фант [1] и Перкелл [21] провели несколько таких экспериментов, однако подобные измерения могут быть выполнены лишь в ограниченном объеме. Другим методом является вычисление формы голосового тракта по акустическим измерениям. В [22] описан подобный метод, предполагающий возбуждение голосового тракта внешним источником. Оба метода являются полезными для получения сведений о динамике речеобразования. Тем не менее они не могут быть применены для получения описания речевых сигналов, например, в задачах связи. В работе Атала [23] описаны результаты прямого измерения  $A(x, t)$  по сигналу речи, произнесенной в нормальных условиях.

Точное решение уравнений (3.1) является весьма сложным, даже если значение  $A(x, t)$  точно известно. Вместе с тем для изучения структуры речевого сигнала нет необходимости в точном и общем решении этих уравнений. Достаточно рассмотреть приближенные решения для простых ситуаций.

### 3.2.2. Однородная труба без потерь (пример)

Для изучения структуры речевого сигнала весьма полезно рассмотреть простую модель речеобразования, в которой функция площади постоянна по обоим аргументам (однородная труба инвариантная к временному сдвигу). Такая форма голосового тракта является прототипом для нейтрального гласного [UH]. Рассмотрим эту модель, оставив для более позднего анализа исследование сложных ситуаций. На рис. 3.14а приведена схема однородной трубы, воз-

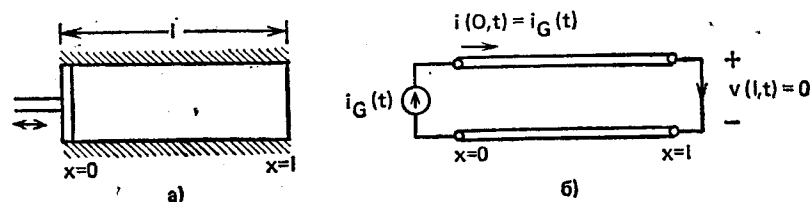


Рис. 3.14. Однородная труба без потерь с идеальной нагрузкой (а) и эквивалентная электрическая линия (б)

буждаемая идеальным источником. Источник представляет собой поршень, который может передвигаться независимо от давления в трубе. Дальнейшее упрощение состоит в предположении, что конец трубы открыт. Звуковое давление, таким образом, постоянно. Изменяется только скорость воздушного потока. Очевидно, что эти существенные упрощения не имеют места в действительности. Однако на этом примере мы убедимся в том, что основной подход к анализу и особенности полученного решения имеют много общего и с более сложной ситуацией. Покажем также, что более общие модели могут быть получены путем соединения однородных труб.

Если  $A(x, t) = A$ , то уравнения в частных производных (3.1) запишутся в виде

$$-\frac{\partial p}{\partial x} = \frac{\rho}{A} \frac{\partial u}{\partial t}; \quad -\frac{\partial u}{\partial x} = \frac{A}{\rho c^2} \frac{\partial p}{\partial t}. \quad (3.2a; б)$$

Можно показать (см. задачу 3.3), что решением уравнений (3.2) являются соотношения

$$u(x, t) = [u^+(t-x/c) - u^-(t+x/c)]; \quad (3.3a)$$

$$p(x, t) = (\rho c/A) [u^+(t-x/c) + u^-(t+x/c)]. \quad (3.3б)$$

В (3.3) функции  $u^+(t-x/c)$  и  $u^-(t+x/c)$  могут рассматриваться как волны, распространяющиеся в положительном (прямом) и отрицательном (обратном) направлениях. Взаимосвязь этих волн определяется граничными условиями.

Читатели, знакомые с теорией электрических длинных линий, знают, что для однородной линии без потерь напряжение  $v(x, t)$  и ток  $i(x, t)$  удовлетворяют уравнениям

$$-\frac{\partial v}{\partial x} = L \frac{\partial i}{\partial t}; \quad -\frac{\partial i}{\partial x} = C \frac{\partial v}{\partial t}, \quad (3.4a; б)$$

где  $L$  и  $C$  — индуктивность и емкость на единицу длины соответственно. Таким образом, теория однородных электрических длинных линий без потерь [24, 25] применима непосредственно для анализа однородной акустической трубы (табл. 3.3). Из таблицы следует, что однородная акустическая труба эквивалентна однородной линии без потерь, коротко замкнутой на одном конце ( $v(l, t) = 0$ ) и возбуждаемой источником тока ( $i(0, t) = i_G(t)$ ) на другом. Это показано на рис. 3.14б.

Таблица 3.3

Акустическая переменная	Электрическая переменная (аналог)
$p$ — давление	$U$ — напряжение
$u$ — скорость потока	$i$ — ток
$\rho/A$ — акустическая индуктивность	$L$ — индуктивность
$A/(\rho c^2)$ — акустическая емкость	$C$ — емкость

Весьма полезным является частотное описание таких линейных систем как длинные линии и линейные цепи. По аналогии можно получить такое описание и для однородной трубы без потерь. Для описания такой трубы в частотной области положим, что граничные условия в точке  $x=0$ :

$$u(0, t) = u_G(t) = u_G(\Omega) e^{i\Omega t}. \quad (3.5)$$

Таким образом, труба возбуждается комплексным экспоненциальным потоком с круговой частотой  $\Omega$  и комплексной амплитудой  $U_G(\Omega)$ . Так как уравнения (3.2) линейные, решения  $u^+(t-x/c)$  и  $u^-(t+x/c)$  будут иметь вид

$$u^+(t-x/c) = K^+ e^{i\Omega(t-x/c)}; \quad (3.6a)$$

$$u^-(t+x/c) = K^- e^{i\Omega(t+x/c)}. \quad (3.6б)$$

Подставляя эти соотношения в (3.3) и используя граничные условия

$$p(l, t) = 0 \quad (3.7)$$

на конце трубы со стороны губ и (3.5) — со стороны голосовой щели, можно

решить (3.2) относительно неизвестных коэффициентов  $K^+$  и  $K^-$ . В результате для  $p(x, t)$  и  $u(x, t)$  получаем

$$p(x, t) = i Z_0 \frac{\sin[\Omega(t-x)/c]}{\cos[\Omega l/c]} U_G(\Omega) e^{i\Omega t}; \quad (3.8a)$$

$$u(x, t) = \frac{\cos[\Omega(x)/c]}{\cos[\Omega l/c]} U_G(\Omega) e^{i\Omega t}; \quad (3.8б)$$

по аналогии с длинными линиями называется *характеристическим сопротивлением акустической трубы*.

В дальнейшем будем избегать решения уравнений относительно прямой и отраженной волн, принимая в качестве функций  $p(x, t)$  и  $u(x, t)$  при комплексном возбуждении выражения<sup>1</sup>

$$p(x, t) = P(x, \Omega) e^{i\Omega t}; \quad (3.10a)$$

$$u(x, t) = U(x, \Omega) e^{i\Omega t}. \quad (3.10б)$$

Подставляя эти соотношения в (3.1), получаем уравнения в полных производных относительно комплексных амплитуд

$$-\frac{dP}{dx} = ZU; \quad -\frac{dU}{dx} = YP, \quad (3.11a; б)$$

где

$$Z = i\Omega\rho/A \quad (3.12)$$

называется *акустическим сопротивлением на единицу длины*;

$$Y = i\Omega A/\rho c^2 \quad (3.13)$$

называется *акустической проводимостью на единицу длины*. Дифференциальные уравнения (3.11) имеют следующее решение:

$$P(x, \Omega) = A e^{\gamma x} + B e^{-\gamma x}; \quad (3.14a)$$

$$U(x, \Omega) = C e^{\gamma x} + D e^{-\gamma x}, \quad (3.14б)$$

где

$$\gamma = \sqrt{ZY} = i\Omega/c. \quad (3.14в)$$

Здесь неизвестные коэффициенты могут быть найдены из граничных условий

$$P(l, \Omega) = 0; \quad U(0, \Omega) = U_G(\Omega). \quad (3.15a; б)$$

Очевидно, что полученное решение полностью совпадает с (3.8). Уравнения (3.8) описывают взаимосвязь между синусоидальным потоком источника, звуковым давлением и скоростью потока в каждой точке трубы. В частности, если нас интересует взаимосвязь скорости потока около губ со скоростью потока источника, из (3.8б) получаем

$$u(l, t) = U(l, \Omega) e^{i\Omega t} = \frac{1}{\cos(\Omega l/c)} U_G(\Omega) e^{i\Omega t}. \quad (3.16)$$

Отношение

$$\frac{U(l, \Omega)}{U_G(\Omega)} = V_a(i\Omega) = \frac{1}{\cos(\Omega l/c)}. \quad (3.17)$$

является частотной характеристикой, связывающей скорости потоков на входе и выходе. Эта функция изображена на рис. 3.15а для  $l=17,5$  см и  $c=35\,000$  см/с.

<sup>1</sup> Далее переменные во временной области будут обозначаться малыми буквами, например  $u(x, t)$ , а соответствующие им переменные в частотной области — большими, например  $U(x, \Omega)$ .

Заменяя  $\Omega$  на  $s/i$ , получаем передаточную функцию в виде преобразования Лапласа (системную функцию)

$$V_a(s) = 2e^{-s/c} / (1 + e^{-2s/c}). \quad (3.18)$$

Заметим, что  $V_a(s)$  имеет бесконечное число полюсов, равномерно распределенных на оси  $i\Omega$  так, что (рис. 3.15б)

$$s_n = \pm i [(2n + 1) \pi c / 2l], \quad n = 0, \pm 1, \pm 2, \dots \quad (3.19)$$

Полюса системной функции линейной системы, инвариантной к сдвигу, определяют собственные частоты системы. Кроме того, они указывают на расположение резонансных частот системы. Эти резонансные частоты называют формантными, когда говорят о процессе речеобразования. Как мы увидим далее, аналогичные резонансные явления будут проявляться при разных конфигурациях голосового тракта.

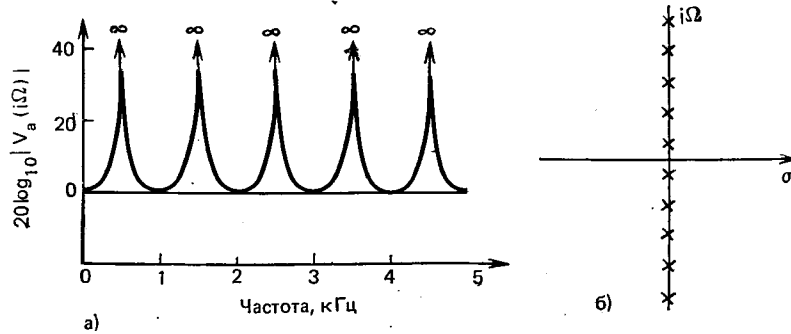


Рис. 3.15. Частотная характеристика (а) и расположение полюсов (б) для однородной трубы без потерь

Следует подчеркнуть, что частотная характеристика позволяет определить отклик системы не только для синусоидальных воздействий, но и для произвольных сигналов с помощью преобразования Фурье. Действительно, (3.17) можно рассматривать как отношение  $V_a(i\Omega)$  преобразования Фурье скорости потока у губ (на выходе) к преобразованию Фурье скорости потока у голосовой щели (на входе). Таким образом, частотная характеристика является удобным средством описания голосового тракта. После того как мы выяснили способ определения частотной характеристики простейшей акустической модели речеобразования, перейдем к анализу более точных моделей.

### 3.2.3. Потери в голосовом тракте

Уравнения, определяющие распространение звуковых волн в голосовом тракте, были получены в предположении отсутствия потерь энергии в трубе. В действительности потери энергии будут возникать за счет вязкого трения между потоком воздуха и стенками трубы, теплопроводности и колебаний стенок голосового тракта. Для анализа этих эффектов необходимо использовать основные законы физики и вывести новую систему уравнений распространения звуковых волн в голосовом тракте. Получение таких уравнений особенно затрудняется тем, что потери зависят от частоты потока. Поэтому наиболее удобно формировать и решать эти уравнения в частотной области [2, 18].

Рассмотрим вначале эффект колебаний стенок голосового тракта. При изменении звукового давления внутри голосового тракта на его стенки будет действовать сила переменной величины. Если стенки тракта мягкие (не жесткие), то площадь поперечного сечения трубы будет изменяться в зависимости от давления. Предполагая, что стенки реагируют на давление «локально» [17, 18], будем считать, что площадь поперечного сечения  $A(x, t)$  является функцией  $p(x, t)$ . Так как изменения звукового давления очень малы, отклонения площади поперечного сечения от «номинального» значения будут небольшими и можно записать

$$A(x, t) = A_0(x, t) + \delta A(x, t), \quad (3.20)$$

где  $A_0(x, t)$  — номинальное значение площади и  $\delta A(x, t)$  — малые отклонения. Это показано на рис. 3.16. С учетом влияния массы и мягкости стенок взаимосвязь между  $\delta A(x, t)$  и  $p(x, t)$  можно записать в виде следующего дифференциального уравнения:

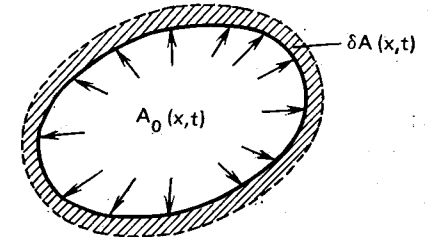


Рис. 3.16. Иллюстрация вибрации стенок голосового тракта

$$m_w \frac{d^2(\delta A)}{dt^2} + b_w \frac{d(\delta A)}{dt} + k_w(\delta A) = p(x, t), \quad (3.21)$$

где  $m_w(x)$  — масса стенок на единицу длины;  $b_w(x)$  — затухание на единицу длины стенок;  $k_w(x)$  — жесткость на единицу длины стенок голосового тракта.

Пренебрегая членами второго порядка относительно величин  $u/A$  и  $pA$ , перепишем (3.1) в виде

$$-\frac{\partial p}{\partial x} = \rho \frac{\partial(u/A_0)}{\partial t}; \quad (3.22a)$$

$$-\frac{\partial u}{\partial x} = \frac{1}{\rho c^2} \frac{\partial(p A_0)}{\partial t} + \frac{\partial A_0}{\partial t} + \frac{\partial(\delta A)}{\partial t}. \quad (3.22b)$$

Итак, распространение звуковых волн в трубе с мягкими стенками, такой как голосовой тракт, описывается системой уравнений (3.20) — (3.22).

Для подробного изучения этого эффекта перейдем в частотную область и, как и раньше, рассмотрим трубу, инвариантную к сдвигу, возбуждаемую потоком с комплексной скоростью. Граничные условия со стороны голосовой щели равны

$$u(0, t) = U_G(\Omega) e^{i\Omega t}. \quad (3.23)$$

Так как уравнения (3.21) и (3.22) линейные и инвариантны к сдвигу, скорость потока и давление могут быть представлены в виде

$$p(x, t) = P(x, \Omega) e^{i\Omega t}; \quad (3.24a)$$

$$u(x, t) = U(x, \Omega) e^{i\Omega t}. \quad (3.246)$$

Подставляя (3.24) в (3.21) и (3.22), получаем

$$-\frac{\partial P}{\partial x} = ZU; \quad -\frac{\partial U}{\partial x} = YP + Y_w P, \quad (3.25a); \quad (3.256)$$

где

$$Z(x, \Omega) = i\Omega \frac{\rho}{A_0(x)}; \quad (3.26a)$$

$$Y(x, \Omega) = i\Omega \frac{A_0(x)}{\rho c^2}; \quad (3.266)$$

$$Y_w(x, \Omega) = \frac{1}{i\Omega m_w(x) + b_w(x) + [k_w(x)/i\Omega]}. \quad (3.26в)$$

Заметим, что (3.25) почти совпадает с (3.11). Различие состоит в том, что появился член  $Y_w$  (за счет учета проводимости стенок) и акустические сопротивление и проводимость оказываются функциями  $x$ . Если труба однородна и  $A_0(x)$  постоянна, то (3.12), (3.13) совпадают с (3.26a, 3.266).

По измерениям, проведенным на теле языка [2], в [18, 19] вычислены параметры (3.26в), после чего дифференциальные уравнения (3.25) были решены для граничных условий  $p(l, t) = 0$  около губ. Отношение

$$V_a(i\Omega) = \frac{U(l, \Omega)}{U_G(\Omega)} \quad (3.27)$$

изображено на рис. 3.17 как функция  $\Omega$  для однородной трубы длиной 17,5 см [18]. Результат близок к рис. 3.15, но имеет существенное отличие. Видно, что резонансы в этом случае расположены не на оси  $i\Omega$  в  $s$ -плоскости. Это означает, что частотная характеристика уже не равна бесконечности на частотах 500, 1500, 2500 Гц и т. д., хотя она и имеет максимумы на этих частотах. Центральные частоты и ширина этих резонансных областей<sup>1</sup> приведены в таблице на рис. 3.17. Отметим несколько существенных особенностей этого примера. Во-первых, центральные частоты расположены выше, чем в случае трубы без потерь. Во-вторых, ширина резонансных областей отлична от нуля, так как на резонансной частоте частотная характеристика имеет конечное значение. Влияние мягкости стенок наиболее существенно проявляется на низких частотах. Этого можно было ожидать, так как массивные стенки незначительно отклоняются на высоких частотах.

Результаты этого примера отражают типичные явления, происходящие при вибрации стенок голосового тракта; центральные частоты несколько смещаются в область верхних частот, резонансные области низких частот оказываются более широкими, чем в других диапазонах. Эффекты вязкого трения и теплопроводности оказывают значительно меньшее влияние, чем вибрация стенок. В

<sup>1</sup> Ширина резонансных областей определяется по уровню 0,707.

[2] Фланаган подробно исследовал эти явления и установил, что влияние вязкого трения можно учесть в частотной области путем добавления в (3.25) действительного зависящего от частоты члена в выражение для акустического сопротивления  $Z$ :

$$Z(x, \Omega) = \frac{S(x)}{[A_0(x)]^2} \sqrt{\Omega \rho \mu / 2} + i\Omega \frac{\rho}{A_0(x)}, \quad (3.28a)$$

где  $S(x)$  — длина окружности сечения трубы;  $\mu$  — коэффициент трения;  $\rho$  — плотность воздуха в трубе.

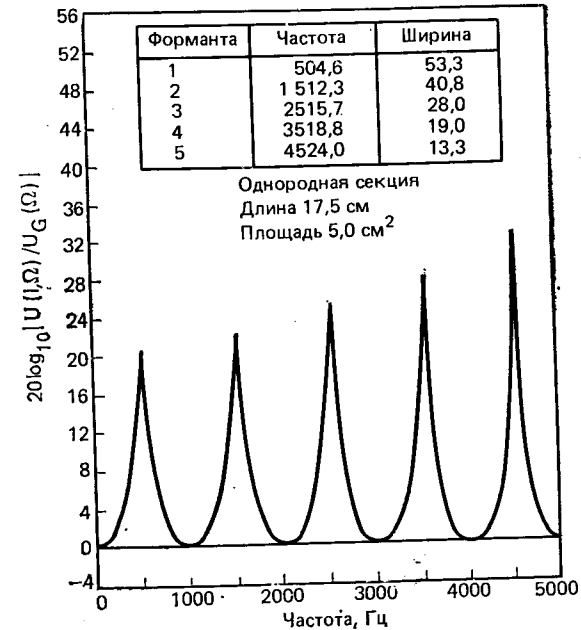


Рис. 3.17. Частотная характеристика однородной трубы с мягкими стенками. Нагрузка — короткое замыкание ( $p(l, t) = 0$ ) [18]

Эффект теплопроводности стенок можно учесть аналогичным образом, добавляя действительный зависящий от частоты член в выражение для акустической проводимости:

$$Y(x, \Omega) = \frac{S(x)(\eta - 1)}{\rho c^2} \sqrt{\frac{\lambda \Omega}{2c_p \rho}} + i\Omega \frac{A_0(x)}{\rho c^2}, \quad (3.286)$$

где  $c_p$  — удельная теплоемкость при постоянном давлении;  $\eta$  — отношение удельной теплоемкости при постоянном давлении к этой же величине при постоянном объеме (адиабатическая постоянная);  $\lambda$  — коэффициент теплопроводности [2]. Значения этих постоянных в (3.28) получены Фланаганом [2]. Здесь достаточно отметить, что потери на трение пропорциональны действительной части  $Z(x, \Omega)$ , т. е.  $\Omega^{1/2}$ . Аналогично потери на теплопро-

водность пропорциональны действительной части  $Y(x, \Omega)$ , т. е. тоже  $\Omega^{1/2}$ . Используя (3.28) для  $Z(x, \Omega)$  и  $Y(x, \Omega)$  и выражение для  $Y_w(x, \Omega)$  (3.26в), уравнения (3.25) можно решить численно [18]. Частотная характеристика, полученная для граничных условий  $p(l, t) = 0$ , показана на рис. 3.18.

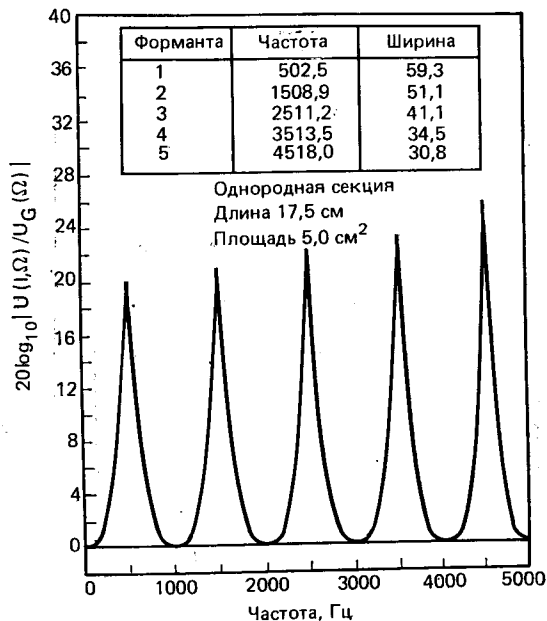


Рис. 3.18. Частотная характеристика однородной трубы с мягкими стенками, потерями на трение и теплопроводность. Нагрузка — короткое замыкание ( $p(l, t) = 0$ ) [18]

Как и раньше, значения резонансных частот и ширины резонансных областей приведены в таблице. Сравнивая рис. 3.18 и 3.17, видим, что центральные частоты при потерях на трение и теплопроводность уменьшаются, а ширина резонансных областей увеличивается. Так как потери на трение и теплопроводность возрастают пропорционально  $\Omega^{1/2}$ , резонансы на высоких частотах расширяются больше, чем в низкочастотной области. Примеры рис. 3.17 и 3.18 отражают типичное влияние потерь в голосовом тракте. Итак, потери на вязкое трение и теплопроводность увеличиваются с возрастанием частоты и оказывают наибольшее влияние на резонансы в высокочастотной области. Потери за счет вибрации стенок в то же время оказывают наибольшее влияние на характеристики в области низких частот. Эффект вибрации стенок приводит к возрастанию резонансных частот, в то время как потери на трение и теплопроводность оказывают обратное влияние. В результате всех этих эффектов резонансы в низкочастотной области несколько смещаются по сравнению с резонансами в трубе

без потерь и с жесткими стенками. Влияние потерь на трение и теплопроводность мало по сравнению с эффектом вибрации стенок на частотах ниже 3—4 кГц. Таким образом, уравнения (3.21), (3.22), в которых не учтены эти потери, хорошо описывают распространение звуковых волн в голосовом тракте. Как мы увидим далее, эффект излучения через губы является источником существенных потерь на высоких частотах. Это еще раз подтверждает правомерность пренебрежения потерями на трение и теплопроводность в моделях речеобразования.

### 3.2.4. Излучение через губы

Мы изучили, каким образом внутренние потери влияют на распространение звуковых волн в голосовом тракте. В рассмотренных примерах граничные условия со стороны губ были всегда равны  $p(l, t) = 0$ . Это соответствует короткому замыканию эквивалентной длинной линии. Получить акустический аналог короткого замыкания очень трудно, так как это предполагает такую конфигурацию трубы, при которой скорость потока может изменяться на конце трубы при неизменном давлении. Однако в реальных условиях труба голосового тракта заканчивается отверстием между губами (или в ноздрах в случае произнесения носовых звуков). Возможная модель такой конфигурации показана на рис. 3.19а, где отверстие между губами показано как отверстие в сфере. В этой модели на нижних частотах отверстие может рассматриваться как излучающая поверхность. Звуковые волны отражаются сферой, которой является голова человека.

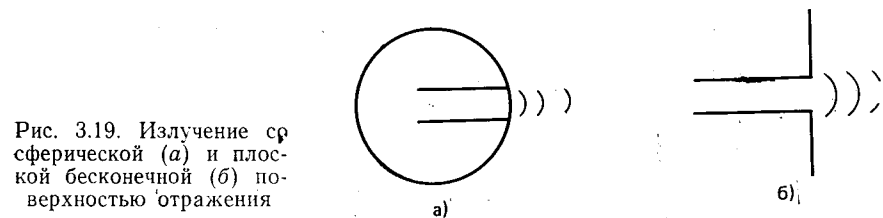


Рис. 3.19. Излучение сферической (а) и плоской бесконечной (б) поверхностью отражения

Эффект отражения очень сложен и труден для описания. Однако для того, чтобы определить граничные условия около губ, достаточно установить взаимосвязь между давлением и скоростью потока у излучающей поверхности. Это тоже оказывается весьма сложным для конструкции, изображенной на рис. 3.19а. Если излучающая поверхность отверстия между губами мала по сравнению с размерами сферы, можно считать отражающую поверхность плоской и бесконечно протяженной (рис. 3.19б). В этом случае можно показать, что связь между комплексными амплитудами давления и скорости потока около губ имеет вид

$$P(l, \Omega) = Z_L(\Omega) U(l, \Omega), \quad (3.29a)$$