

*Р. К. Потапова, В. В. Потапов*



# РЕЧЕВАЯ КОММУНИКАЦИЯ

От звука к высказыванию



STUDIA PHILOLOGICA

УДК 80/81  
ББК 81  
Р 46

Рецензенты:  
член-корр. РАН В. А. Виноградов,  
д-р филол. наук О. Ф. Кривнова

**Потапова Р. К., Потапов В. В.**

Р 46 Речевая коммуникация: От звука к высказыванию. — М.:  
Языки славянских культур, 2012. — 464 с. — (Studia philologica).

ISSN 1726-135X  
ISBN 978-5-9551-0559-8

В общей семиотической системе, обеспечивающей успешность межличностной коммуникации, *устная речь* занимает особое место в силу своих природных свойств (скорости процессов речепорождения и речевосприятия, органической связи со звуковой материей окружающего мира и т. д.), полифункциональности, связанной с актуализацией в речи различных функциональных установок (например, коммуникативной, эмотивной и др.), и полиинформативности, соотносящейся со всем объемом знаний о говорящем, его собеседнике, ситуации общения и т. д.

Содержательный диапазон монографии охватывает основные аспекты речеведения как раздела общей семиотики, включающего описание акустической природы звуков речи (сегментного строя языков) и просодии (супрасегментных характеристик различных речевых единиц).

Детально представлен материал, позволяющий отразить все многообразие средств микро- и макросегментации речевого потока. Антропоцентрический подход к изучению звуковой материи обусловлен всесторонним анализом проблематики, связанной с триадой «*вербалика — паравербалика — коммуникация*». Особое внимание уделено различным аспектам речеведения с позиций прагмалингвистики.

Монография включает результаты целого ряда авторских экспериментальных исследований и оригинальных концептуальных разработок.

**ББК 81**

*В оформлении переплета использован фрагмент картины  
В. Кандинского «Доминирующая кривая» (1936)*

ISBN 978-5-9551-0559-8

© Потапова Р. К., Потапов В. В. 2012

© Языки славянских культур, 2012

Электронная версия данного издания является собственностью издательства, и ее распространение без согласия издательства запрещается.

# ОГЛАВЛЕНИЕ

ВВЕДЕНИЕ .....	13
----------------	----

## I. АКУСТИЧЕСКАЯ ПРИРОДА ЗВУКОВ РЕЧИ

<b>ГЛАВА 1. Экспериментально-фонетическое исследование сегментного строя языков.....</b>	<b>19</b>
1.1. Сущность спектрального анализа речи .....	21
1.2. Основные понятия спектрального анализа .....	28
1.3. Опорные признаки сегментации речи на спектрограммах.....	41
1.3.1. Акустические ключи, используемые для сегментации речевых сигналов и для определения признаков способа образования звуков .....	49
1.3.2. Акустические ключи, используемые для определения признаков твердости—мягкости и места образования согласных и качества гласных .....	56
1.4. Коартикуляция в процессе речеобразования и ее спектральные корреляты .....	58
1.5. Синтез речевых сегментов по данным акустического анализа.....	64
1.6. Использование данных спектрального анализа в речеведческих исследованиях.....	72
1.7. Принципы подбора экспериментального материала .....	78
<b>ГЛАВА 2. Супraseгментные (просодические) характеристики речи.....</b>	<b>90</b>
2.1. Введение .....	90
2.2. Собственные акустические характеристики звуков речи.....	95
2.2.1. Частота основного тона.....	95
2.2.2. Интенсивность.....	98
2.2.3. Длительность .....	105
2.3. Просодические характеристики слога.....	107
2.4. Число компонентов слога и частота встречаемости основных словесных структур .....	108
2.5. Акустические признаки компонентов слога.....	110

2.5.1. Длительность гласного в составе слога .....	110
2.5.2. Частота основного тона компонентов слога .....	112
2.5.3. Характер примыкания компонентов слога.....	113
2.6. Просодические характеристики фонетического слова.....	116
2.7. Изменение акустических характеристик компонентов слога в составе фонетического слова и фразы.....	119
2.8. Модификации границ фонетического слова в потоке речи .....	122
2.9. Подбор экспериментального материала с учетом просодических характеристик слога и фонетического слова.....	124
2.10. Просодия, просодика, просодемика .....	127

## **II. СЕГМЕНТНО-СТРУКТУРНАЯ АРХИТЕКТОНИКА РЕЧЕВОГО ВЫСКАЗЫВАНИЯ**

<b>ГЛАВА 3. Фонетические средства сегментации звучащей речи.....</b>	<b>135</b>
3.1. Микросегментные (звуковые, слоговые) признаки членения речевого потока.....	135
3.1.1. Фонологические критерии микросегментации речи.....	135
3.1.2. Фонетические признаки микросегментации речи на звуковые единицы.....	139
3.1.3. Принципы слоговой сегментации речевого потока.....	149
3.2. Фонетические характеристики членения речевого потока на фонетические слова (ритмические структуры).....	157
3.2.1. Структурно-типологические подходы к исследованию речевого ритма.....	157
3.2.2. Просодические признаки сегментации речевого потока на фонетические слова (ритмические структуры).....	165
3.2.3. Качественные и частотные характеристики фонетических слов (ритмических структур).....	172
3.3. Макросегментация звучащих текстов.....	178
3.3.1. Единицы членения.....	178
3.3.2. Соотношение фонетических и структурных единиц в звучащей речи (условия просодической актуализации потенциальных синтагм).....	184

3.3.3. Обусловленность фонетического членения речи текстовым механизмом.....	191
3.4. Роль просодических средств в перцептивно-слуховой макросегментации звучащей речи .....	194
3.4.1. Физическая пауза .....	195
3.4.1.1. Восприятие длительности временных интервалов.....	195
3.4.1.2. Перерывы фонации, связанные с процессом дыхания .....	199
3.4.1.3. Надежность распознавания пауз в зависимости от длительности временных интервалов и их синтаксического и просодического контекста .....	200
3.4.2. Паузы без перерыва в фонации (нетемпоральная пауза) ....	205
3.4.3. Средства макросегментации спонтанной речи.....	209
<b>ГЛАВА 4. Просодические характеристики макросегментации звучающей речи.....</b>	<b>218</b>
<b>ГЛАВА 5. Теоретические и прикладные аспекты речевой сегментологии .....</b>	<b>238</b>
<b>ГЛАВА 6. Сегментация звучащих текстов с опорой на критерий вербальной многоуровневой комплексности .....</b>	<b>252</b>
6.1. Результаты опытов по перцептивно-слуховой оценке сегментации звучащих текстов .....	255
6.2. Акустические корреляты перцептивно-слуховых признаков сегментации слитной речи .....	268
<b>ГЛАВА 7. Феномен речевой сегментологии с учетом когнитивной рефлексии и коммуникативно- деятельностного подхода.....</b>	<b>281</b>
 <b>III. АНТРОПОЦЕНТРИЧЕСКИЙ ПОДХОД К ИЗУЧЕНИЮ ЗВУКОВОЙ МАТЕРИИ</b> 	
<b>ГЛАВА 8. Сопоставительная паравербалика и ее просодические корреляты .....</b>	<b>293</b>

<b>ГЛАВА 9. Фонетические средства оптимизации речевого воздействия.....</b>	<b>300</b>
<b>ГЛАВА 10. Восприятие эмоционального поведения иноязычных и инокультурных коммуникантов .....</b>	<b>318</b>
<b>ГЛАВА 11. Идентификация говорящего и его эмоционального состояния по голосу и речи .....</b>	<b>334</b>
<b>ГЛАВА 12. Определение «облика говорящего» по данным перцептивно-слухового анализа звучащей речи.....</b>	<b>356</b>
<b>ГЛАВА 13. Восприятие звучащей речи в затрудненных условиях .....</b>	<b>366</b>
<b>ГЛАВА 14. Фоностилистика научной речи с позиций прагмафонетики .....</b>	<b>379</b>
<b>ЗАКЛЮЧЕНИЕ.....</b>	<b>401</b>
 <b>ПРИЛОЖЕНИЯ</b>	
<b>Приложение 1. Иллюстрации акустической обработки речевого сигнала на базе новых информационных технологий.....</b>	<b>405</b>
<b>Приложение 2. Интонационная специфика русской речи: от теории к практике в Интернете.....</b>	<b>407</b>
<b>Приложение 3. <i>Rodmonga K. Potapova, Vsevolod V. Potapov. Kommunikative Sprechfähigkeit. Rußland und Deutschland im Vergleich. Böhlau Verlag. Köln; Weimar; Wien, 2011. 312 S. (на нем. яз.)</i> .....</b>	<b>418</b>
<b>Литература.....</b>	<b>425</b>
<b>Именной указатель .....</b>	<b>456</b>

## ГЛАВА 1. Экспериментально-фонетическое исследование сегментного строя языков

В экспериментальной фонетике обычно обращаются к трем типам анализа речи: слуховому, артикуляторному, акустическому. В связи с этим обычно говорят о «трех фонетиках»: фонетике слушающего, фонетике говорящего и фонетике наблюдателя. Аспект акустического исследования речи относится к фонетике наблюдателя, так как акустические методы исследования предполагают фиксацию и анализ особенностей речевого сигнала с позиций наблюдателя-экспериментатора. В связи с вышесказанным в экспериментально-фонетических исследованиях используют инструментальные (аппаратно-программные) средства, предназначенные для проведения слухового, артикуляторного и акустического видов анализа.

Слуховой анализ речи проводится на базе использования фонограмм, различного рода программ сегментации и т. д., позволяющих экспериментатору вычленять речевые сегменты определенной протяженности во времени и с определенными фонетическими характеристиками.

Все существующие методы артикуляторного анализа речи подразделяются на статические и динамические. Наряду со статическими методами исследования артикуляции звуков речи (например, статической палатографией, рентгенографией, фотографией) в настоящее время широко используются динамические методы: например, динамическая палатография, кинорентгенография, электромиография, динамическое тензометрирование, воздушная плетизмография и т. д. (подробно об этом см. [Потапова 1971а]).

Акустические методы исследования связаны с физической природой речевого сигнала. Речеобразующий аппарат человека рассматривается как сложная система, генерирующая колебания определенной частоты. Колебательные процессы фиксируются и анализируются с помощью специальных компьютерных программ. Наиболее полное представление о характере изучаемого явления дает комплексный подход, то есть обращение ко всем трем типам анализа речевого сигнала: слуховому, артикуляторному, акустическому.

Спектральный анализ речевого сигнала является одним из видов акустического анализа (см., например, [Артемов 1960; Бондарко 1961; 1977; Варшавский, Литвак 1955; Деркач 1983; Зиновьева 1986; Князев, Пожарицкая 2005: 55—100; Кодзасов, Кривнова 2001: 99—195; Кузнецов 2004; Мальмберг 1962; Потапова 1977; Фант 1970; Черри 1972: 153—198; Fant 1973; Jassem 1973; Ladefoged 1996; Lieberman, Blumstein 1988; Olive, Greenwood, Coleman 1993; O'Shaughnessy 1987; Potarov 1993] и др.). Понятие «спектр» встречается, например, в одном из разделов физики — оптике, где спектром в узком смысле слова называют цветные полосы, получающиеся в результате разложения светового луча призмой (или другим прибором) с учетом длины волн. Если обратиться к природе звука, генерируемого каким-либо источником, то можно установить, что всякое звучащее тело совершает не одно, а целый ряд колебательных движений. Например, струна колеблется не только по всей своей длине, но и по длине каждой из ее частей: половины, четверти, восьмой и т. д. В процессе колебания происходит слияние звуковых волн разной длины. Колебание звучащего тела по всей его длине дает основной тон звучания, остальные колебания — колебания частей этого тела — дают обертоны, которые в речевых исследованиях принято называть *гармониками*. Каждое из колебаний, входящее в состав сложного колебательного целого, имеет свою форму. Таким образом, применительно к речевому сигналу принято говорить об акустическом *спектре*.

Понятие спектра встречается не только в физике и технике. Оно широко применяется в математике. Если какой-либо колебательный процесс представляется в виде суммы гармонических колебаний различных частот (гармоник), то спектром колебательного процесса называется функция, описывающая распределение амплитуд по различным частотам. Спектр показывает, какого рода колебания преобладают в данном процессе, какова его внутренняя структура<sup>1</sup>. Установлено, что механизм речеобразования является квазистационарным источником звуков с изменяющимися во времени характером возбуждения и частотной характеристикой. Следовательно, спектральная характеристика речевого сигнала может быть описана в терминах математической связи между аperiодической функцией времени  $f(t)$  и ее комплексным спектром плотности амплитуд  $F(\omega)$ , что выражается парой преобразований Фурье [Фланаган 1968: 161].

---

<sup>1</sup> Более подробно об этом см. [Сапожков 1963; Фант 1964; Фланаган 1968].

Спектральный анализ, имеющий строгое математическое выражение, является с позиций физики некоторой аппроксимацией. Однако, несмотря на известную относительность, понятие спектрального анализа прочно укоренилось в экспериментально-фонетических исследованиях. Известно, что спектральный анализ широко применяется при решении целого ряда задач, связанных прежде всего с исследованием звукового (сегментного) строя различных языков.

### 1.1. Сущность спектрального анализа речи

Акустический анализ речевого сигнала базируется на целом ряде определенных принципов, основным из которых является положение о том, что всякий речевой сигнал представляет собой реакцию резонансной системы голосового тракта на возбуждение его одним или несколькими генераторами звуковых колебаний. Иными словами, речевая волна является результатом воздействия источника звука на фильтрующую систему речевого тракта. Речевая волна определяется характеристиками источника звука и характеристиками фильтрующей системы речевого тракта. Речеобразование может быть представлено блок-схемой (рис. 1).

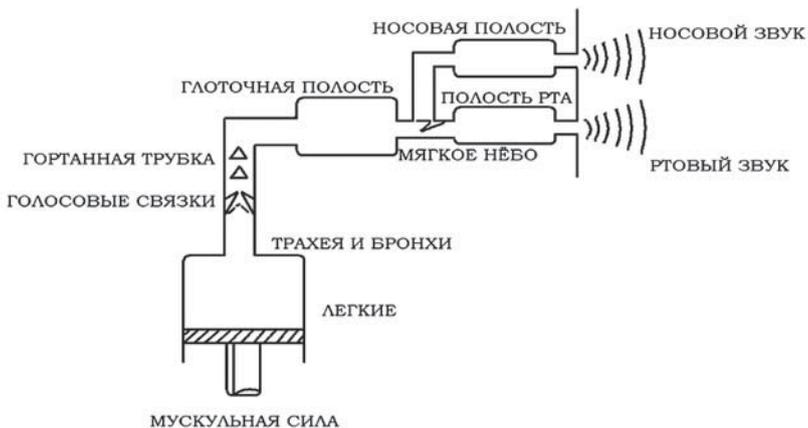


Рис. 1. Схема механизма образования речи  
(по Дж. Флангану [Фланган 1968: 37])

Основным свойством голосового источника является периодичность создаваемого звукового эффекта. Структура речевого сигнала сложна по своей природе. На форму звуковой волны, исходящей от источника, накладываются частотные характеристики резонирующих полостей.

Поскольку такой генератор звуковых колебаний, как гортань, дает монотонный по форме спектр, то изменение результирующей огибающей спектра может определяться только изменением конфигурации речевого тракта (рис. 2). Из акустики известно, что при резонансе энергия от одного звучащего тела возбуждает энергию другого тела. Первое тело обычно называют вибратором, второе — резонатором. Передача энергии происходит через волновые колебания воздушной среды. Роль резонаторов в процессе речеобразования играют полости глотки, носоглотки, рта и носа. Изменение размеров резонаторов и щелей, соединяющих их, вызывает изменение резонансных частот.

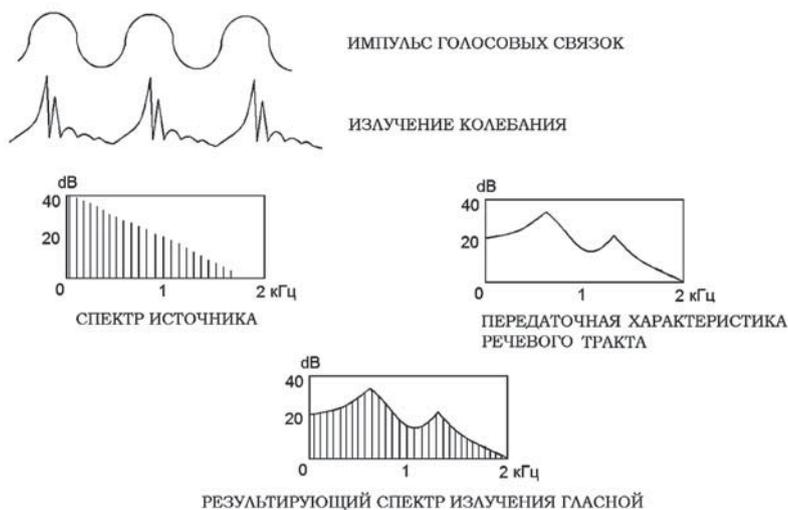


Рис. 2. Речеобразование как процесс фильтрации колебаний гортани  
(по М. А. Сапожкову [Сапожков 1963: 26])

Основным допущением теории речеобразования является допущение о независимости характеристик резонансной системы речевого тракта от характеристик источника голоса. Это означает, что

связь резонансных частот и частот гармоник основного тона имеет случайный характер.

«...Согласно известному разложению Фурье любую сложную периодическую кривую можно представить в виде суммы простых гармонических колебаний с определенными амплитудами и начальными фазами, причем, если частоту сложного колебания принять за единицу, то частоты его составляющих (гармоник) должны находиться в отношении 1 : 2 : 3 : 4 и т. д.» [Поспелов 2002: 22]. Таким образом, простую гармоническую функцию можно представить с помощью трех детерминантов: частоты, амплитуды и начальной фазы [Там же: 23].

При произнесении звуков речи через речевой тракт проходит или тональный импульсный сигнал, или шумовой, или тот и другой вместе, что приводит к выводу: речевой сигнал можно рассматривать как своего рода модулированную несущую [Сапожков 1989: 46]. При этом гортань рассматривается как источник первичного речевого сигнала, являющегося результатом чисто механических колебаний тканей гортани [Там же: 47].

Колебательными движениями (или колебаниями) в физике и технике принято называть такие виды движений (или изменения состояния), которые обладают какой-либо степенью повторяемости. При этом периодические колебания характеризуются тем, что каждое значение изменяющейся величины повторяется неограниченное число раз через одинаковые промежутки времени.

Использование аналогов речевого тракта дает возможность глубже изучить анатомические и физиологические особенности речи. Так, первый электрический аналог голосового тракта<sup>2</sup> был продемонстрирован в 1950 г. Он имел три регулировки: одну — для места подъема на теле языка, вторую — для высоты этого подъема, третью — для изменения степени сужения губ. В 1953 г. был создан новый электрический артикуляторный аналог в МТИ (Массачусетском технологическом институте — США). Речевой тракт человека рассматривался в виде акустической трубы, на нижнем конце которой располагалась голосовая щель, на противоположном конце — губы, а все другие артикуляторные органы, локализующиеся между ними, имели возможность изменяться с различной степенью точно-

<sup>2</sup> В акустике наибольшее распространение получил резонатор Гельмгольца. Подробнее об этом см. [Поспелов 2002: 29].

сти. Данная модель аналога голосового тракта была способна образовывать различные звуки речи путем изменения расположения источников возбуждения (голосовой щели или точек турбулентности на поверхности трубы). Изменение форм и объемов полостей трубы вело к образованию разных типов резонансов. Следует подчеркнуть, что теоретически число резонансов бесконечно, но в реальной речи энергия в высоких частотах быстро затухает и в практических исследованиях не учитывается. Таким образом был сконструирован электрический аналог артикуляторной трубы, основанный на идеальном представлении голосового тракта в виде последовательного соединения 35 цилиндров.

В 1960 г. аналог был усовершенствован. Если первый аналог МТИ был статическим голосовым трактом, который был способен «произносить» только лишь изолированные гласные и некоторые согласные, то второй аналог мог уже производить целые слоги, что свидетельствует о том, что в 1960 г. была получена динамическая модель артикуляторного аналога.

В 1961 г. к аналогу голосового тракта был добавлен аналог носовой полости, появилась возможность исследовать образование носовых согласных и назализованных гласных. Позже в 1965 г. решалась задача придания артикуляторной модели еще большего числа характеристик, свойственных речи человека, путем непосредственного наблюдения за артикуляторным процессом с помощью фотографической и кинорентгенографической техники. По диаграммам движений различных участков голосового тракта человека, полученным путем тщательного анализа кадра за кадром, был подготовлен ряд программ для управления секциями артикуляторного аналога с помощью компьютера.

Изучение речи при помощи спектрального анализа было начато в 40-х гг. XX в., когда была предложена конструкция спектрографа, названная «Видимая речь» («Visible speech») [Potter, Kopp, Green 1947]. Этот аппарат сыграл большую роль в деле развития и дальнейшего совершенствования методики исследования звуков в потоке речи. В настоящее время спектрограф «Видимая речь», претерпевший целый ряд модификаций, лег в основу специальных программных средств, используемых для акустического анализа речи.

В ходе анализа регистрируется картина в пространстве: время — интенсивность — частота. Ось времени расположена горизонтально, ось частоты — вертикально, интенсивность передается степенью

потемнения. Кроме этого возможна регистрация спектральных «срезов», «кадров» в координатах «амплитуда — частота» для любого момента времени.

В качестве исторического экскурса следует остановиться на развитии спектрографической техники в отечественной науке. В 1945 г. в Лаборатории экспериментальной фонетики и психологии речи МГПИИЯ им. М. Тореза, возглавляемой выдающимся ученым и организатором научных изысканий в области экспериментальной фонетики и психологии речи доктором психологических наук, профессором Владимиром Алексеевичем Артемовым, были начаты исследования по изучению звукового строя различных языков мира с помощью 23-канального спектрографа. Спектрограф имел 23 фильтра, каждый из которых был настроен на определенную резонансную частоту. Участие тех или иных фильтров в процессе анализа речевой волны было обусловлено наличием в анализируемой речевой волне колебаний соответствующей резонансной частоты.

Несколько позже в той же лаборатории был разработан и функционировал 48-канальный спектрограф «С-48». Этот прибор представлял собой спектроанализатор с последовательной покадровой регистрацией результатов анализа. Благодаря наличию 48 фильтров звуковой сигнал разделялся на частотные составляющие в диапазоне от 80 до 12 050 Гц. Позднее в лаборатории эксплуатировался 50-канальный спектрограф «С-50» с определенной настройкой фильтров по частотам (таблица 1). Динамический диапазон по амплитуде был равен 40 дБ. Запись речевого сигнала производилась с помощью кинорегистрации на 16-миллиметровую киноплёнку. В результате спектрального анализа исследователь получал покадровое изображение речевого сигнала в его динамике. Длительность каждого кадра была равна 14,5 мс. В обсуждении технических характеристик упомянутого спектрографа принимал участие Г. Фант, неоднократно посещавший данную лабораторию.

Представляет интерес оптическая аналоговая аппаратура анализа, которая обладала сравнительно высокими характеристиками спектрального анализа речи. К числу параметров относились: максимальное число независимых точек в спектре — 500, динамический диапазон сигнала на выходе — 50 дБ. Время одного цикла анализа равно 0,5 с. Трёхмерные (время, частота, амплитуда) образы служили материалом для классификации речевых сигналов.

Характеристики полос пропускания фильтров  
50-канального динамического спектроанализатора

№№	Гц	№№	Гц
1	100	26	3090
2	200	27	3260
3	300	28	3430
4	400	29	3560
5	500	30	3700
6	600	31	3900
7	700	32	4100
8	800	33	4300
9	900	34	4520
10	1020	35	4700
11	1140	36	4900
12	1270	37	5150
13	1380	38	5400
14	1500	39	5700
15	1640	40	6000
16	1750	41	6350
17	1875	42	6700
18	1980	43	7000
19	2090	44	7400
20	2220	45	7850
21	2350	46	8260
22	2500	47	8930
23	2650	48	9590
24	2800	49	10480
25	2960	50	11220

В настоящее время исследователи располагают возможностью использовать в целях спектрального анализа компьютерные программы (например, Praat, Multi-Speech, Sound Fodge, CSL, Speech Analyzer и др.). С их помощью осуществляется спектральный анализ оцифрованного речевого сигнала. Создание быстродействующих вычислительных машин позволило использовать сложные программы обработки речи благодаря способности компьютера

запоминать и обрабатывать с большой скоростью значительный объем данных.

Следует подчеркнуть, что количественные изменения речевого тракта и качественная сторона звукового потока могут быть в полной мере изучены только путем комплексного подхода. В связи с этим кинорентгенография, синхронизированная со спектрографией речи, может быть примером одного из эффективных способов исследования речевого сигнала.

Точная синхронизация кадров кинорентгенограмм и спектрограмм достигнута при помощи специального механизма, соединенного с ведущим валом кинокамеры. Каждый оборот ведущего вала вызывал импульс, отмечающий единичные кадры, а при помощи специальной передачи вызывались импульсы, отмечающие кадры. Полученные импульсы записывались на звуковую дорожку. В то же самое время на другую звуковую дорожку записывалась речь, произносимая во время кинорентгенографирования.

Исследование акустической структуры речевого сигнала производилось при помощи быстродействующего 52-канального динамического спектрографа, имеющего следующие характеристики: диапазон анализируемых частот от 40 до 14 000 Гц, динамический диапазон — 35 дБ, предварительное усиление частот выше 1000 Гц — 6 дБ на октаву.

Результаты анализа регистрировались одновременно двумя кинокамерами на две киноплёнки: в виде динамической спектрограммы, непрерывно меняющейся во времени, с координатами «время — частота — интенсивность», и в виде спектральных срезов (кадров) с координатами «частота — интенсивность». Каждый кадр соответствовал отрезку речевого сигнала длительностью около 8 мс.

Спектральные исследования проводились, в частности, с целью определения акустических коррелятов звуков речи для установления статистического объема данных от различных групп дикторов, что давало возможность определения вариативности «речевого поведения» человека. Первоначально информация, например, о согласных была далеко не полной, так как верхняя граница частотного диапазона была равна 3400 Гц. Последующие модели звукового спектрографа существенно раздвинули верхние границы частотного диапазона.

Таким образом, на спектрограммах отражена сложная природа речевого сигнала. Речь в ее динамике представляет собой своеобразную комбинацию непрерывных и дискретных характеристик. Движения артикуляционных органов в процессе речевого производства отра-

жаются в непрерывных изменениях характеристик спектрограммы. Вместе с тем дискретные явления также накладывают определенный отпечаток на спектральные изображения речи. Так, например, дискретными являются спектральные картинки таких признаков, как включение или выключение голосового источника, признаки полного или частичного закрытия рта и др. Признаки смыкания и размыкания мягкого нёба выражены менее ярко, но и они относятся к типу признаков, определяющих дискретность спектрального изображения.

Относительная дискретность речевого потока делает более реальной сегментацию спектральных изображений на отрезки, соответствующие разным типам звуков. Резкие изменения спектральной картины во времени, связанные с изменениями конфигурации источников и речевого тракта в целом могут быть использованы в качестве маркеров при членении речи на последовательность сегментов, что позволяет отобразить на спектрограммах следующие характеристики:

- изменения спектральной структуры звуков речи по признаку «частота — время»;

- изменение энергетической структуры частотных составляющих звуков речи по признаку «интенсивность — время»;

- последовательность звуков с гармонической, шумовой, гармонически-шумовой структурой;

- наличие квазистационарных и переходных участков (см. иллюстрации в Приложении I).

## 1.2. Основные понятия спектрального анализа

При проведении спектрального анализа речи имеют дело с двумя типами частотных параметров: первый — это частота основного тона  $F_0$ , второй — это результат наложения резонансных частот речевого тракта на периодическую картину вибрации голосовых связок  $F_n$ . В результате этого наложения возникают новые максимальные значения в огибающей спектра. Фонетические характеристики звука определяются распределением энергии в спектре звука, что в свою очередь обусловлено распределением резонансных частот речевого тракта в процессе генерации звука [Сапожков 1963; Фант 1964; 1970; Златоустова, Потапова, Трунин-Донской 1986; Поспелов 2002].

В ходе спектрального анализа исследователь сталкивается с рядом специальных терминов. К числу последних следует отнести, напри-

мер, такие термины, как *форманта*, *формантный уровень*, *формантная полоса*, *формантная модель*.

Основным недостатком формантной гипотезы является неразработанность формальных критериев определения форманты. О форманте как максимуме на огибающей спектра можно говорить лишь в том смысле, что уровень данной составляющей превышает среднее значение уровней ближайшей левой и правой составляющих. Таким образом, формантой обычно называют максимум энергии (наиболее усиленные резонансные частоты) в спектре звука, а антиформантой — минимум спектральной энергии (область наиболее подавленных резонансных частот). Частотная локализация формант обуславливается формой и объемом резонирующих полостей.

Гармоники, являющиеся характеристикой источника, расположенного в гортани, и форманты могут не совпадать. Другими словами, частотная характеристика форманты может оказаться между двумя гармониками.

Частотная характеристика форманты может быть измерена двояким способом:

1. По центру видимой формантной полосы.
2. Местоположением пика в огибающей кривой.

*Форманта* обозначается через  $F_n = F_1, F_2, F_3, \dots$ , где «n» означает порядковый номер форманты. При этом под частотой форманты обычно понимают либо частоту локального максимума передаточной функции (полюс), либо средневзвешенную частоту (центроид) данной области.

Среднее расстояние между формантами вычисляется по формуле  $C/2l_1$ , где  $C = 35\ 300$  см/с (скорость звука), а  $l_1$  — средняя длина речевой полости (ширина речевого тракта не оказывает влияния на расстояние между формантами). Обычно  $l_1 = 17,5$  см, соответственно  $C/2l_1 = 1000$  Гц. Расстояние между формантами в 1000 Гц рассматривается как среднее исходное идеальное.

*Формантный уровень*, обозначаемый как  $L_n$ , идентичен уровню пик в огибающей спектра. Формантный уровень обуславливается особенностями голоса диктора.

*Формантные полосы* более тесно соотносятся с отдельными моделями формант, то есть базируются на особенностях артикуляции. Ширина формантных полос ( $B_n$ ) определяется различием между точками огибающей спектра. Ширина формантных полос равна приблизительно 40—250 Гц. Среднее расстояние  $B_1 - B_2$ , равно 75 Гц.

Формантная модель является относительным физическим коррелятом какой-либо конфигурации речевого тракта, включающим языковую, глоточную и губную артикуляции. Формантные частоты для мужских голосов равняются в среднем приблизительно следующим величинам:

$$F_1 = 150 — 850 \text{ Гц}$$

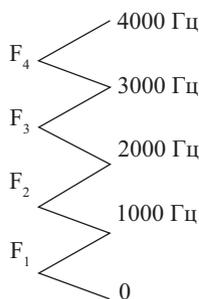
$$F_2 = 500 — 2500 \text{ Гц}$$

$$F_3 = 1500 — 3500 \text{ Гц}$$

$$F_4 = 2500 — 4500 \text{ Гц}$$

Формантные частоты для женских голосов приблизительно на 17% выше, чем для мужских.

Как уже указывалось выше, среднее расстояние между формантами равно 1000 Гц. Поэтому исходная идеальная формантная модель гласных имеет следующий вид:



Распределение формант зависит от особенностей уклада органов артикуляции при произнесении звука [Якобсон, Фант, Халле 1962]. Формантные модели гласных можно представить следующим образом:

1. Подъем тела языка способствует:
  - повышению  $F_2$ ,
  - сближению  $F_2$  с  $F_3$ .
2. Подъем тела языка и его одновременное движение к более передней палатальной позиции способствует:
  - повышению  $F_3$ ,
  - сближению  $F_2$  и  $F_3$ .
3. Подъем тела языка и его одновременное движение назад способствует:
  - повышению  $F_1$ ,
  - сближению  $F_1$  и  $F_2$ .

4. Движение тела языка назад (задняя спинка языка сильно сближена с глоткой) способствует:

- понижению  $F_1$ ,
- сближению  $F_2$  и  $F_1$ .

5. Движение тела языка назад и дополнительная увулярная артикуляция способствует:

- умеренному увеличению  $F_3$ ,
- промежуточной позиции  $F_2$ .

6. Движение тела языка вперед с соответствующей ретрофлексной артикуляцией способствует:

- понижению  $F_3$ ,
- сближению  $F_2$  и  $F_3$ .

7. Наличие полного препятствия (преграды) в любой точке речевого тракта способствует:

- совпадению формант  $F_2$  и  $F_1$  или  $F_2$  и  $F_3$ .

8. Наличие полного препятствия (преграды) при подъеме языка вверх способствует:

- совпадению  $F_2$  и  $F_3$ ,
- приближению  $F_1$  и 0.

9. Наличие полного препятствия (преграды) при движении языка назад способствует:

- совпадению  $F_2$  и  $F_1$ ,
- приближению  $F_1$  к 0.

10. Эффект округления или выдвигания губ понижает частоту всех формант безотносительно к объему речевой полости.

11. Фонетическая значимость передних гласных недостаточно четко выражается отношением  $F_1 / F_2$ . Для большей точности используются параметры:

$$F_2 - 1/2 ((F_2 - F_1) / (F_3 - F_1)) \times (F_3 - F_2)$$

Под *интенсивностью* звука понимается мощность звуковой волны за единицу времени, измеряемая в микроваттах.

*Уровень интенсивности*, выраженный в dB, пропорционален высоте амплитуды кривой спектра.

Звуковые колебания  $|W|$  рассматриваются как производные линейной системы (то есть речевого тракта, соединенного с одним или двумя источниками), где  $W = Ts$  ( $T$  — переменная функция системы,  $s$  — источник).

Существует два основных вида генератора, излучающих периодические колебания и шум. В спектрограммах периодические колебания представлены характерной гармонической структурой, а шум — иррегулярным распределением энергии во времени. Эти два источника при образовании различных звуков активны в разной степени.

Иногда в литературе резонансы называют полюсами. Когда по линии передачи, не имеющей потерь, источник не расположен на той же линии что и выход, наблюдаются различные отклонения. Подобные отклонения ведут к подавлению энергии данного частотного района, то есть действуют подобно обратному резонансу — антирезонансу (нулю). Если система допускает небольшие потери, тогда полюса и нулевые точки определяются, с одной стороны, частотной локализацией резонанса и, с другой стороны, объемом заглушения. Полюсы (резонансы) зависят от конфигурации речевого тракта; нулевые точки — от взаимодействия двух резонирующих систем, которые регулируются посредством открытия дополнительного прохода, неконечной локализацией источника.

Если локализация нулевой точки близка к локализации полюса, нулевая точка подавляется эффектом полюса. По мере того, как между ними увеличивается расстояние, подавление уменьшается. Гласные звуки могут быть определены полностью полюсами без учета нулевых точек, тогда как все остальные звуки речи требуют экспериментального определения минимального объема заглушения (нулевых точек) в каждом отдельном случае.

При рассмотрении звуков речи решающим оказывается характеристика по трем признакам:

- по источнику,
- по особенностям динамики,
- по резонансным особенностям.

Каждый из этих признаков имеет определенную спектральную характеристику.

1. По источнику различают два основных признака звуков речи:

- вокалические (невокалические),
- консонантические (неконсонантические).

Звуки, обладающие первым признаком, характеризуются определенной структурой. «Портрет» гласных состоит из ряда  $F_n$  (формант), местоположение которых строго определено, так как обусловлено артикуляционной характеристикой данного гласного. Каждая из  $F_n$  соответствует различным формантам гласных. При определе-