

МИНИСТЕРСТВО
ВЫСШЕГО И СРЕДНЕГО СПЕЦИАЛЬНОГО ОБРАЗОВАНИЯ СССР

Московский государственный педагогический институт
иностранных языков имени Мориса Тореза

Р. К. ПОТАПОВА

экспериментальное

ЭКСПЕРИМЕНТАЛЬНО-ФОНЕТИЧЕСКОЕ

ИССЛЕДОВАНИЕ

СЕГМЕНТНОГО УРОВНЯ ЯЗЫКОВ

(текст лекций спецкурса)

Москва — 1979

МИНИСТЕРСТВО ВЫСШЕГО И СРЕДНЕГО СПЕЦИАЛЬНОГО ОБРАЗОВАНИЯ
СССР

Московский государственный педагогический институт
иностранных языков имени Мориса Тореза

Р.К. ПОТАПОВА

ЭКСПЕРИМЕНТАЛЬНО-ФОНЕТИЧЕСКОЕ ИССЛЕДОВАНИЕ
СЕМАНТИЧЕСКОГО УРОВНЯ ЯЗЫКОВ

(текст лекций спецкурса)

МОСКВА - 1979

ПРЕДИСЛОВИЕ

Лекции спецкурса "Экспериментально-фонетическое исследование сегментного уровня языков" предназначены для студентов, аспирантов и соискателей, специализирующихся в области экспериментальной фонетики. Лекции содержат последовательное изложение принципов проведения спектрального анализа различных речевых сегментов (звуков, слогов и т.д.), проверки данных спектрального анализа путем синтезирования речевого сигнала, подбора экспериментального материала.

Рассмотрение в отдельном спецкурсе перечисленных выше вопросов представляется целесообразным ввиду возросшего интереса к изучению сегментного строя языков и отсутствия соответствующего учебного материала по данному ряду вопросов в систематизированном и доступном для студентов старших курсов и аспирантов, приступающих к самостоятельным экспериментально-фонетическим исследованиям, виде.

В основу данного пособия положены результаты научных исследований, проводимых в течение многих лет на кафедре фонетики немецкого языка и в Лаборатории экспериментальной фонетики и психологии речи МПНИИ им.М.Тореза.

по договору между МПНИИ

Пособие подготовлено на кафедре фонетики немецкого
языка.

Составитель: к.ф.н., доц. Р.К.Потапова

ВВЕДЕНИЕ

Предметом настоящего пособия является описание акустического анализа, проводимого с помощью спектрального метода. В экспериментальной фонетике прибегают обычно к трем типам анализа речи: слуховому, артикуляторному, акустическому. В связи с этим обычно говорят о "трех фонетиках": фонетике слушающего, говорящего и наблюдателя. Аспект акустического изучения речи относится к фонетике наблюдателя, т.к. акустические методы исследования предполагают фиксацию и анализ особенностей речевого сигнала с позиций наблюдателя-экспериментатора. В связи с вышеуказанным в экспериментально-фонетических исследованиях используют аппаратуру, предназначенную для проведения слухового, артикуляторного и акустического типов анализа.

Слуховой анализ речи проводится на базе использования фонограмм, различного рода сепараторов и сегментаторов, позволяющих экспериментатору вычленять речевые сегменты с определенной протяженностью 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 (статической палатографией, рентгенографией, фотографией) в настоящее время широко используются динамические артикуляторные методы: динамическая палатография, кинорентгенография, электромиография, динамическое тензометрирование, воздушная плетизмография и т.д. (Подробнее об этом см: Р.К. Потапова, "Основные современные способы анализа и синтеза речи", Пособие для студентов отделения приклад-

ной лингвистики, М., 1971)

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

Спектральный анализ речевого сигнала является одним из видов акустического анализа. Понятие "спектра" 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) по длинам 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 колеблется каждая из ее частей: половина, четверть, восьмая 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. Разложение звуковой волны на ее составляющие, различающиеся по своей частоте и форме, аналогично разложению света на составляющие, различающиеся длиной волн. 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. С п е к т р п о к а з ы в а е т , к а к о г о р о д а к о л е б а н и я п р е о б л а д а ю т в д а н н о м п р о ц е с с е , к а к о в а е г о в н у т р е н н я я с т р у к т у р а .

Спектральный анализ, имеющий строгое математическое выражение, является с физической точки зрения 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 с исследованием сегментного строя различных языков.

Лекция I. 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. I).

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 рече-производства играют полости глотки, носоглотки, рта и носа. Изменение размеров резонаторов и шелей, соединяющих их, вызывает изменение резонансных частот.

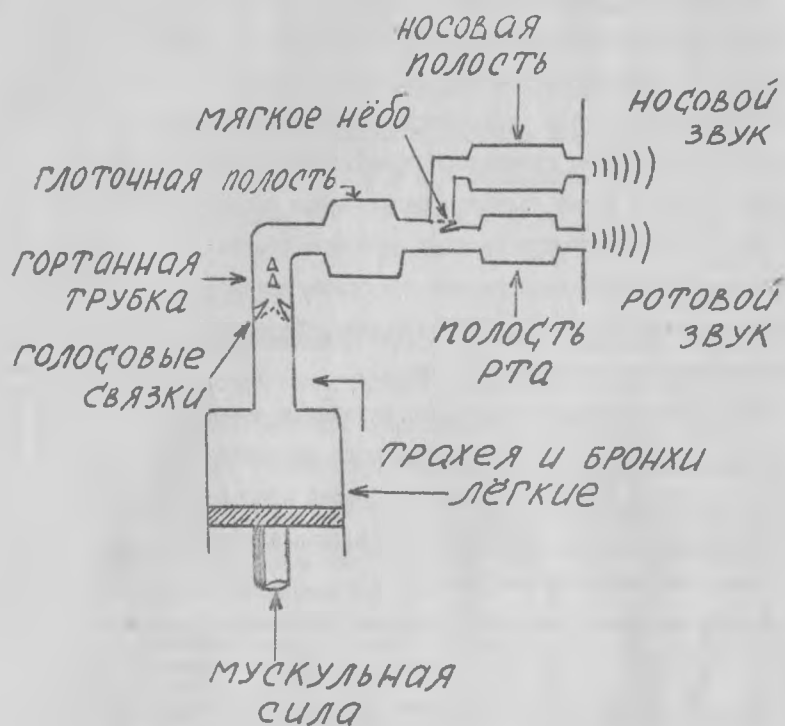


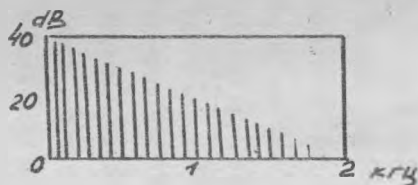
Рис. I. СХЕМА МЕХАНИЗМА ОБРАЗОВАНИЯ РЕЧИ



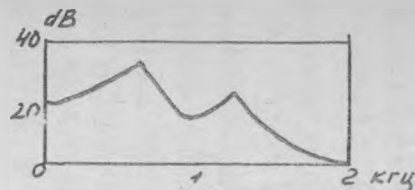
ИМПУЛЬСЫ ГОЛОСОВЫХ СВЯЗОК



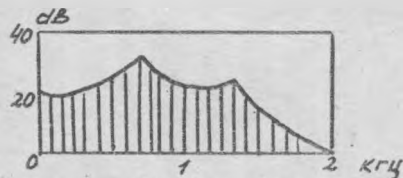
ИЗЛУЧАЕМЫЕ КОЛЕБАНИЯ



СПЕКТР ИСТОЧНИКА



ПЕРЕДАТОЧНАЯ ХАРАКТЕРИСТИКА
РЕЧЕВОГО ТРАКТА



РЕЗУЛЬТИРУЮЩИЙ СПЕКТР ИЗЛУЧЕНИЯ ГЛАСНОЙ

Рис. 2. РЕЧЕОБРАЗОВАНИЕ КАК ПРОЦЕСС ФИЛЬТРАЦИИ КОЛЕБАНИЙ ГОЛОСА

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.

Использование ~~аналогов речевого тракта~~ дает возможность глубже изучить анатомические и физиологические аспекты речи. Первый электрический ~~аналог голосового тракта~~ был продемонстрирован в 1950 году. Он имел три регулировки: одну - для положения наивысшей точки на языке, другую - для высоты этой точки, третью - для степени сужения губ. В 1953 году был создан новый электрический артикуляторный аналог в МТИ (Массачусетском технологическом институте). Речевой тракт человека рассматривается как акустическая труба, в нижнем конце которой расположена головная щель, в противоположном конце - губы, а все органы, расположенные между ними, могут изменяться в разной степени. Эта труба образует различные звуки речи путем изменения расположения источников возбуждения (голосовой щели или точек турбулентности на поверхности труб). Изменение форм и объемов полостей трубы ведет к образованию разных резонансов. 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 идеализации голосового тракта в виде последовательного соединения 35 цилиндров.

В 1960 году аналог был усовершенствован. Если первый аналог МТИ был статическим ~~голосовым~~ трактом, который способен

50
60
был "произносить" только изолированные гласные и некоторые согласные, то второй аналог мог уже производить слоги. Можно сказать, что в 1960 году была получена динамическая модель артикуляторного аналога.

В 1961 году к аналогу голосового тракта был добавлен аналог носовой полости, появилась возможность исследовать образование носовых согласных и назализованных гласных. Позже в 1965г. 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 ЭВМ.

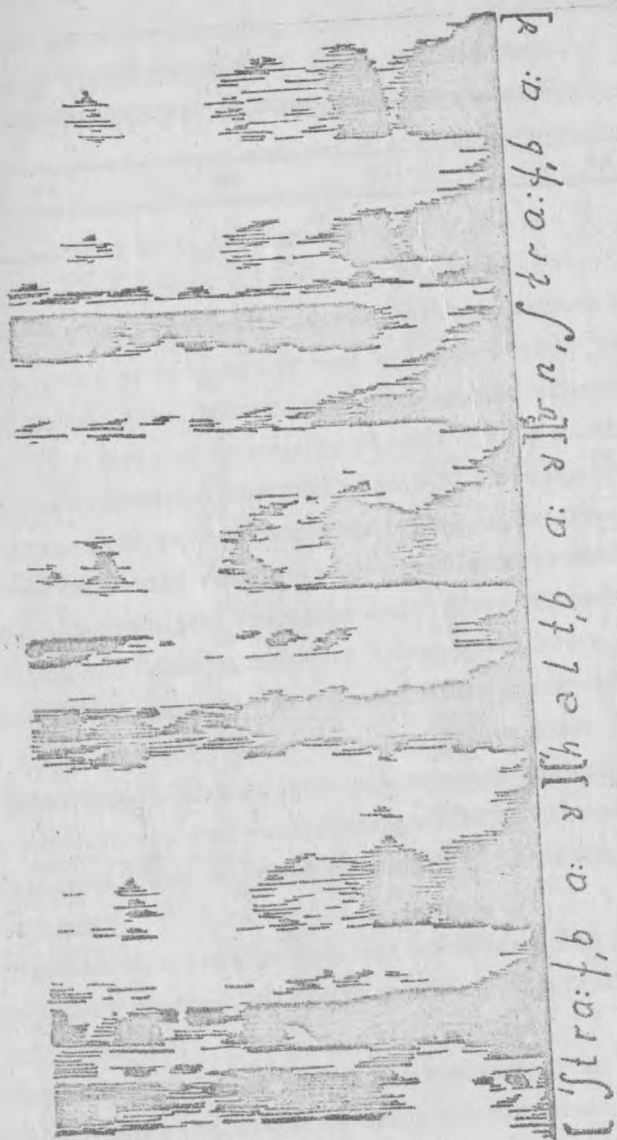
Изучение речи при помощи спектрального анализа было начато в 1941 году, когда была предложена конструкция спектрографа, названная "Видимая речь". Этот аппарат сыграл большую роль в деле развития и дальнейшего совершенствования методики исследования звуков в потоке речи. В настоящее время спектрограф "Видимая речь", претерпевший целый ряд модификаций, является одним из основных приборов, используемых при акустических исследованиях речи.

Последняя модель "Кей-Сонаграфа" существенно отличается от своих предшественниц. После записи сигнал многократно воспроизводится на большой скорости. На специальной бумаге, укрепленной на вращающемся барабане, регистрируется картина в пространстве:

время - интенсивность - частота. Ось времени расположена горизонтально, ось частоты - вертикально, интенсивность передается степенью потемнения (см. иллюстрацию I). Кроме этого возможна регистрация спектральных "срезов", "кадров" в координатах амплитуда - частота для любого момента времени.

В течение последующих лет было сконструировано и применено на практике большое число спектрографов, причем модификации в конструкции были вызваны своеобразием стоящей перед исследователями задачи. В 1945 году в Лаборатории экспериментальной фонетики и психологии речи МПШИЯ им. М. Тореза были начаты исследования по изучению звукового строя языка с помощью 23-канального спектрографа. Спектрограф имел 23 фильтра, каждый из которых был настроен на определенную частоту. Участие тех или иных фильтров в процессе анализа речевой волны было обусловлено наличием в анализируемой речевой волне колебаний соответствующей частоты.

Несколько позже в той же лаборатории был разработан и функционировал 48-канальный спектрограф "С-48". Этот прибор представлял собой спектроанализатор с последовательной показывающей регистрацией результатов анализа. Благодаря наличию 48 фильтров звуковой сигнал разделялся на частотные составляющие в диапазоне от 80 до 12 050 гц. В настоящее время в лаборатории эксплуатируется 50-канальный спектрограф "С-50" с определенной настройкой фильтров по частотам (Таблица I). Динамический диапазон по амплитуде равен 40 дб. Запись речевого сигнала производится с помощью кинорегистрации на 16 мм-киноплёнку. В результате спектрального анализа исследователь получает покадровое изображение



ИЗДАНИЕ I. ВЪВЕДЕНИЕ ИЛИ ПРЕДИСЛОВИЕ

Таблица I.

характеристики полос пропускания фильтров
50-канального динамического спектроанализатора

№ - I	Гц-100	№ - 26	Гц-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

речевого сигнала в его динамике. Длительность каждого кадра равна 14,5 мсек. (См. иллюстрацию 2).

Представляет интерес оптическая аналоговая аппаратура анализа. Оптическая аналоговая аппаратура обладает сравнительно высокими параметрами, характеризующими качество спектрального анализа. Максимальное число независимых точек в спектре 500, динамический диапазон сигнала на выходе 50 дб. Время одного цикла анализа равно 0,5 сек. Трехмерные (время, частота, амплитуда) образы могут служить материалом для классификации речевых сигналов.

В настоящее время исследователи располагают возможностью использовать в целях спектрального анализа ЭВМ. С помощью ЭВМ осуществляется спектральный анализ речевого сигнала, данные в цифровом виде поступают на выход устройства. Создание быстродействующих вычислительных машин позволило использовать сложные программы обработки речи благодаря способности ЭВМ запоминать и обрабатывать с большой скоростью значительный объем данных.

Следует подчеркнуть, что количественные изменения речевого тракта и качественная сторона звукового потока может быть в полной мере изучена только путем комплексного подхода. В связи с этим кинофлуорография, синхронизированная со спектрографией речи, является одним из эффективных средств исследования речевого сигнала.

Подобная синхронизация была осуществлена в Лаборатории экспериментальной фонетики института языка и литературы Академии наук Эстонской ССР.

Точная синхронизация кадров рентгенокинофильма и спектрограмм достигается при помощи специального механизма, соединенного с ведущим валом кинокамеры. Каждый оборот ведущего вала вы-

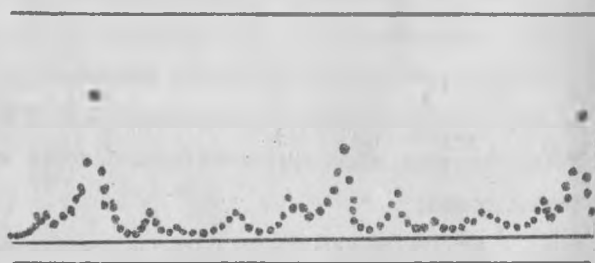
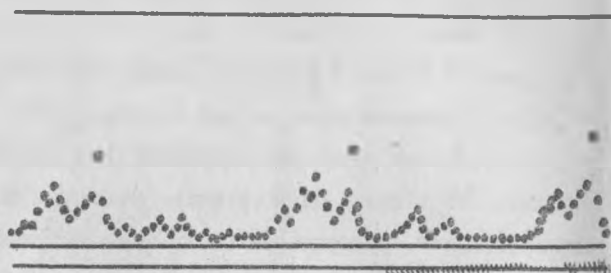


ИЛЛЮСТРАЦИЯ 2. СПЕКТРОГРАММЫ РЕАЛИЗАЦИИ ЗВУКА "Е" В
СЛОВЕ *fraisiers*

21
41
30

зывает импульс, отмечающий единичные кадры, а при помощи зубчатой передачи вызываются усиленные импульсы, отмечающие первый и десятый кадры. Полученные импульсы записываются на одну звуковую дорожку стереофонического магнитофона "Дуза-10". В то же самое время на вторую звуковую дорожку записывается речь, произносимая во время кинофлуорграфирования.

Исследование акустической структуры речевого сигнала производится при помощи быстродействующего 52-канального динамического спектрографа, имеющего следующие характеристики: диапазон анализируемых частот от 40 до 14000 гц, динамический диапазон - 35 дб, предварительное усиление частот выше 1000 гц - 6 дб на октаву.

Результаты анализа регистрируются 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 соответствует отрезку речевого сигнала длительностью около 8 мсек.

Спектральные исследования проводились с целью определения акустических коррелятов звуков речи, отдельных сегментов звуков речи, с целью установления статистического объема данных от различных групп дикторов, для определения вариативности "речевого поведения" человека. Первоначально информация, например, о согласных была далеко не полной, так как верхняя граница частотного диапазона была равна 4000 гц, что соответствовало верхней границе частотного диапазона в практике телеграфирования. Последующие модели звукового спектрографа существенно раздвинули верхние

границы частотного диапазона.

Таким образом, подводя итог всему сказанному, можно утверждать, что на спектрограммах отражена сложная природа речевого сигнала. Речь в ее динамике представляет собой своеобразную комбинацию непрерывных и дискретных характеристик. Движения артикуляционных органов в процессе речепроизводства отражаются 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 неба 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.

Таким образом, на спектрограммах отображены следующие характеристики:

- а) изменения спектральной структуры звуков речи по принципу: частота - время;
- б) изменение энергетической структуры частотных составляющих звуков речи по принципу: интенсивность-время;

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

Г) наличие квазистационарных и переходных участков.

Лекция 2. ОСНОВНЫЕ ПОНЯТИЯ СПЕКТРАЛЬНОГО АНАЛИЗА

При проведении спектрального анализа имеют дело с двумя типами частотных параметров: первый – это частота основного тона, второй – это результат наложения резонансных частот речевого тракта на периодическую картину вибрации голосовых связок. 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.

В ходе спектрального анализа исследователь сталкивается с рядом специальных терминов. К числу последних следует отнести такие основные термины спектральных описаний речи, 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, форманте обычно визуируют максимум энергии в спектре звука. Частотная локализация форманта 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.

Среднее расстояние между формантами вычисляется по формуле $C/2l_4$, где $C = 35300$ см/сек (скорость звука), а l_4 - средняя длина речевой полости (ширина речевого тракта не оказывает влияния на расстояние между формантами). Обычно $l_4 = 17,5$ см, соответственно $C/2l_4 = 1000$ 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 обусловливается особенностями голоса диктора.

Формантные полосы более тесно соотносятся с отдельными моделями формант, т.е. базируются на артикуляции. Ширина формантных полос (B_n) определяется разницей 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 приблизительно следующим величинам:

$$F_1 = 150 - 850 \text{ гц}$$

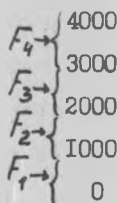
$$F_2 = 500 - 2500 \text{ гц}$$

$$F_3 = 1500 - 3500 \text{ гц}$$

$$F_4 = 2500 - 4500 \text{ гц}$$

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:



ФОРМАНТНЫЕ МОДЕЛИ ГЛАСНЫХ

1. Подъем языка вверх способствует:

- а) повышению F_2
- б) сближению F_2 с F_3

2. Подъем языка и его движение к передней палатальной позиции способствует:

- а) повышению F_3
- б) сближению F_2 и F_3

3. Подъем языка и движение языка назад способствует:

- а) повышению F_1
- б) сближению F_1 и F_2

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

- а) понижению F_1
- б) сближению F_2 и F_1

5. Движение языка назад и дополнительная увулярная артикуляция способствует:
 - а) умеренному возрастанию F_3
 - б) промежуточной позиции F_2
6. Движение языка вперед с соответствующей ретрофлексной артикуляцией способствует:
 - а) понижению F_3
 - б) сближению F_2 и F_3
7. Полное препятствие в любой точке речевого тракта способствует:
 - а) совпадению формант F_2 и F_1 или F_2 и F_3
8. Полное препятствие при подъеме языка вверх способствует:
 - а) совпадению F_2 и F_3
 - б) приближению F_1 к 0
9. Полное препятствие при движении языка назад способствует:
 - а) совпадению F_2 и F_1
 - б) приближению F_1 к 0
10. Эффект округления или выдвижения губ всегда понижает частоту всех формант безотносительно к объему речевой полости.

Соответственно артикуляторная оппозиция нелабиализованный /лабиализованный акустически может быть выражена соотношением:

$$\begin{aligned}
 &\text{нелабиализованный} - \text{лабиализованный} \\
 &(F_1 + F_2) > (F_1 + F_2) \\
 &(F_1 + F_2 + F_3) > (F_1 + F_2 + F_3)
 \end{aligned}$$

II. Фонетическая значимость передних гласных недостаточно четко выражается отношением F_1/F_2 .

Для большей точности используются параметры:

$$F_2 - \frac{1}{2} \left(\frac{F_2 - F_1}{F_3 - F_1} \right) \times (F_3 - F_2)$$

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

Уровень интенсивности, выраженный в дБ , пропорционален высоте амплитуды кривой спектра.

Звуковые колебания $/W/$ рассматриваются как производные линейной системы (т.е. 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 по разному активны при образовании различных звуков.

Иногда в литературе резонансы называют полосами. Когда на 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, т.е. 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.

Полюса (резонансы) зависят от конфигурации речевого тракта; нулевые точки — от взаимодействия двух резонирующих систем, которые регулируются посредством: а) открытия дополнительного прохода, б) неконечной локализацией источника.

Если локализация нулевой точки близка к локализации полюса, нулевая точка подавляет эффект полюса. По мере того, как между ними увеличивается расстояние, подавление уменьшается.

Гласные фонемы 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:

- 1) По источнику;
- 2) по особенностям динамики;
- 3) по резонансным особенностям.

Каждый из этих признаков имеет определенную спектральную характеристику.

I. По источнику различают два основных признака звуков речи:

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

Звуки, обладающие первым признаком, характеризуются совершенно определенной структурой. Портрет гласных состоит из ряда F_n (формант), местоположение которых строго определено: оно обусловлено характеристикой данного гласного. Каждая из этих F_n соответствует различным формантам гласных. При определении спектральных характеристик гласных оперируют обычно тремя форман-

тами (F_1, F_2, F_3). Их распределение по частотной шкале выражается относительными числами (R) - рацию.

Например, R_1 показывает, находятся ли две первые форманты близко друг от друга или удалены на значительное расстояние. Другими словами, R_1 является показателем того, массирована ли основная энергия спектра в высоких или в низких частотах: R_1 передних гласных значительно больше, чем R_1 задних гласных.

В целях осмысления спектральных данных используют следующие противопоставления:

I. Низкий - высокий

Акустически этот признак выражается в концентрации энергии в спектре: в том случае, если энергия сосредоточена в низких частотах спектра, то звук характеризуется как низкий; если соответственно энергия концентрируется в высоких частотах спектра - как высокий.

Артикуляционно этот признак выражается в "вариациях полости". Низкие гласные или согласные звуки обусловлены большим объемом и меньшей расчлененностью ротовой полости, тогда как высокие - меньшим объемом и большей расчлененностью.

Соответственно к низким относятся лабиальные согласные, которые противоплагаются согласным дентальным, и велярные, противоплагаемые палатальным. Задние гласные, артикулируемые оттягиванием языка, являются низкими; передние гласные, образованные движением языка вперед - высокими. Таким образом, соотношение "низкий-высокий" соответствует горизонтальному разделению резонатора.

Обычно вспомогательным фактором в образовании низких звуков является сжатие заднего отверстия ротового резонатора посредством сужения фаринкса, тогда как высокий ^{звук} образуется при расширенной глотке. Так, например, при произнесении некоторых звуков ширина глоточного отверстия, равная при отсутствии артикуляции звуков 13,3 мм, варьирует следующим образом:

<u>низкие</u>	<u>высокие</u>
/u/- 3,8	/i/- 15,2
/o/ - 5,5	/ɛ/- 12,7
	/n/- 8,9
/k/ - 3,8	

Спектральное противопоставление "низкий - высокий" выражается изменением второй форманты, ее соотношением с другими формантами спектра. Показателем этого противоположения является отношение: $R_1 = F_2 : F_1$

Иногда используется более точный вариант:

$$R_1 : R_2 = \frac{F_2 : F_1}{F_3 : F_2}$$

Сближение расстояния $F_2 - F_1$ у гласных заднего ряда дает возможность синтезировать эти звуки при наличии только одной форманты. Для гласных /o/, /ɔ/ эта единственная форманта должна занимать среднее положение между F_1 и F_2 , а для гласных /o/, /u/ находиться ближе к F_2 .

Для синтеза передних гласных необходимы всегда две форманты, причем вторая форманта должна быть ближе к F_2 .

2. Диезный — простой

Акустически — это противоположение спектра звука, форманты которого сдвинуты в область высоких частот, и спектра звуков, не обладающих этим признаком. Артикуляционно при этом признаке ротовая полость сокращается путем подъема средней спинки языка к верхнему нёбу. Обычно при этом у фонем, относящихся к категории диезных, сильно расширен фарингальный проход.

В палатализованных звуках, которые произносятся с участием губ, F_2 сдвигается в направлении к F_3 .

$$\begin{array}{ccc} \text{непалатализованные} & & \text{палатализованные} \\ R_2 = F_3 : F_2 & > & R_2 = F_3 : F_2 \end{array}$$

3. Компактный — диффузный

Компактные звуки характеризуются преобладанием центрально локализованного формантного района. У диффузных звуков энергия рассеяна по большому отрезку частотной шкалы.

Основное артикуляционное различие между компактными и диффузными звуками заключается в отношении между объемом резонирующих полостей перед сужением и позади него.

У гласных компактность возрастает с увеличением поперечного сечения прохода. Широкие гласные более компактны, чем узкие гласные. Аналогичное отношение резонирующих полостей достигается при изменении объема глотки. Объем глотки у диффузных звуков всегда больше, чем у соответствующих компактных. Измерением этих отношений у гласных звуков служат отношения $R_3 = F_3 : F_4$.

Другими словами, отношение R_3 соответствует вертикальному сужению резонатора.

В произнесении русского диктора форманты гласных (*i, u, a*) характеризуются следующими величинами:

$$/i/ F_1 = 250 \text{ в}$$

$$F_2 = 2750 \text{ в}$$

$$F_3 = 3400 \text{ в}$$

$$/u/ F_1 = 350 \text{ в}$$

$$F_2 = 600 \text{ в}$$

$$F_3 = 2600 \text{ в}$$

$$/a/ F_1 = 650 \text{ в}$$

$$F_2 = 1250 \text{ в}$$

$$F_3 = 2200 \text{ в}$$

$$R_3 \text{ для } /i/ = 13,6$$

$$/u/ = 7,4$$

$$/a/ = 3,4$$

Таким образом, R_3 показывает, что у гласных $/i/$, $/u/$ энергия рассеяна по большому отрезку частотной шкалы спектра, у гласной же $/a/$ основная энергия сосредоточена на узком участке спектра.

У согласных компактность представлена преобладанием формантного района, расположенного в центре. Подобные согласные четко отличаются от звуков, где ярко выражены нецентральные районы.

Соответственно к группе компактных относятся согласные $k, g,$

q, j , а к группе диффузных — согласные t, d, n, p, b, f, m .

4. Напряженный — ненапряженный

Акустически эта оппозиция выражается высокой общей энергией спектра и ее широкой разверткой во времени сравнительно с низкой общей энергией и ее узкой разверткой во времени. Артикуляционно напряженные фонемы произносятся с большей отчетливостью и большим давлением, чем соответствующие ненапряженные фонемы. Большая мускульная напряженность языка, голосовых связок и всего речевого тракта связаны с известной деформацией стенок речевого тракта в сравнении с его нейтральной позицией.

Напряженные звуки характеризуются более длительными интервалами звучания и большей интенсивностью, чем звуки ненапряженные.

У напряженных гласных сумма значений ^{разностей} формант больше, чем у соответствующих ненапряженных. Это признак и является решающим при отнесении звуков к категории напряженных и ненапряженных.

5. Носовой – неносовой.

Признак назализации может быть представлен как у гласных, так и у согласных.

Спектр ^{носовых} фонем в отличие от неносовых характеризуется наличием большего числа формантных областей. У носовых гласных появляются дополнительные форманты, расположенные между первой и второй формантой. Одновременно наблюдается спад интенсивности F_1 и F_2 . У гласных с высокой F_1 , например, гласной /a/, дополнительная носовая форманта появляется обычно ниже F_1 и реже – выше F_1 .

Дополнительные полюса и нулевые точки, связанные с назализацией согласных, представляют локальные искажения в спектре, не оказывающие влияния на другие резонансные признаки.

6. Огубленный – неогубленный (Беззубный – простой)

Акустически это противоположение спектра звуков, у которых форманты сдвинуты в область более низких частот, спектру звуков, не обладающих этим признаком.

Противоположение "огубленный – неогубленный" имеет своим физическим коррелятом изменения частотных характеристик в сторону понижения:

неогубленные	$F_1 + F_2 + F_3$	$>$	$F_1 + F_2 + F_3$	огубленные
гласные				гласные

Эффект лабиализации проявляется в наибольшей степени у той форманты, формирование которой обуславливается речевой полостью. Резкое смещение формантной характеристики вверх при губной артикуляции прослеживается, в основном, у гласного /u/ в F_1 , у глас-

ного ($\mathfrak{z}$) в F_3 и у гласного (i') в F_3 , т.е. в тех формантах, которые являются решающими в формировании частотной характеристики данного гласного.

В качестве примеров следует привести некоторые результаты формантного анализа гласных, полученные на материале русского, английского и немецкого языков.

Таблица 2. Русские гласные (вне фонетического контекста)⁺⁾

Гласный	F_1	F_2	F_3
у	300	625	2500
о	400	750	2400
а	$R_1 = 410$ 700 $R_2 = 1100$ 1250 $R_3 = 3100$		
э	440	1800	2550
и	240	2250	3200
ы	300	1480	2230

^{+) Приведены данные Г.Фанта из работы "Акустическая теория речеобразования", М., 1964. Для гласных /о, а/ внесены уточнения по данным М.Халле (The Sounds Patterns of Russian, 1959). Следует указать на то, что по данным некоторых исследователей F_3 практически постоянна для всех гласных и равна 2500 гц, исключением является F_3 для гласной /и/, равная 3000 гц.}

Таблица 3. Английские гласные (вне фонетического контекста)⁺⁺⁾

Гласный	F_1	F_2	F_3
i	200	2600	3500
ɪ	300	2200	3000
e	300	2200	3600
ɛ	550	1950	2600
æ	600	1800	2400
ʌ	450	1200	1650
ʌ	600	1150	2600
a	680	1100	2500
ɔ	600	950	2500
o	460	840	2500
ʊ	460	910	2400
u	390	950	2300
oɪ	600	910	2500
aɪ	750	1300	2600
au	750	1500	2600

⁺⁺⁾ Данные взяты из работы Х.Траби. H.M.Truby. Acoustico-cine-radiographic analysis considerations with especial reference to certain consonantal complexes. Acta Radiologica.Suppl., Stockholm, 1959.

Таблица 4. Средние значения формант гласных в односложных словах на материале немецкого языка +++)

Гласный	F_1	F_2	F_3
i:	300	1920	3000
ɪ	340	1620	2580
e:	360	1920	2800
ɛ	470	1800	2800
ə	600	1400	2800
ɑ:	500	1200	2800
ɔ	550	1120	2500
o:	400	980	2800
ʊ	470	1120	2800
u:	400	980	2580

В спектральной характеристике согласных звуков содержится информация о способе, месте образования звука, точно так же, как и общей конфигурации речевого тракта в момент произнесения звуков. Но определить и выявить формантную модель такого согласного, которая являлась бы закрытым физическим коррелятом определенной конфигурации речевого тракта на материале изучения этого звука не всегда возможно.

+++) Приведены данные Л.И.Прокоповой.См.: Структура слога в немецком языке. Киев, 1973

Поэтому для спектральной характеристики согласных звуков особенно важным является определение прежде всего модификаций, происходящих в предшествующих или последующих гласных фонемах, т.е. характеристика согласных складывается из каких-то постоянных и переменных величин, не являющихся самостоятельными единицами в сообщении. При описании звуков они соединяются воедино, но при спектральной характеристике согласных они измеряются и рассматриваются изолированно.

Соответственно при исследовании согласных основным моментом является проблема сегментации или проблема выяснения константных и варьирующих звуковых признаков. Так, например, произнесение глухого смычного звука обычно складывается из 4 сегментов: 1.смычка + 2.раскрытие смычки + 3. фрикция + 4.аспирация.

Сегмент взрыва и фрикции отличаются друг от друга насыщенностью энергии. Различие между взрывом и фрикцией зависит от источника: взрыв, будучи образован посредством резкого возбуждения полостей, связан с использованием давления, тогда как фрикция образуется турбулентной струей воздуха, стремящейся в щель. Для спектрального портрета согласных решающим является ширина этого прохода, степень заполнения которого и рассматривается. При этом важным артикуляционным условием для выявления концентрации энергии в спектре служит наличие полости определенного объема перед источником, т.е. в палатальной или велярной областях. В том случае, если подобная полость отсутствует, энергия звука распространяется по всему спектру. Особенно широкая развертка спектра губных согласных связана именно с отсутствием

такой полости¹⁾.

В отличие от относительно константного характера формант гласных звуков форманты согласных отличаются неустойчивым лабильным характером. "Локус" согласных, т.е. определенное место в частотной шкале спектра, является лишь исходной точкой, от которой начинается частотный сдвиг к устойчивому частотному уровню формант гласных.

Локус первой форманты согласных связан со способом их образования. Локус второй форманты связан с относительно фиксированным местом артикуляции согласных. Эта связь подтверждается тем, что группы согласных (b, p, m) или (d, t, n) или (g, k, ŋ) характеризуются одними и теми же сдвигами второй форманты. У смычно-взрывных согласных первая часть интервала между локусом этих согласных и формантами соседних гласных представлена на спектре отсутствием всяких значений.

Частотный сдвиг локуса второй форманты имеет различный характер (восходящий или нисходящий) в зависимости от спектральной характеристики следующего гласного.

Неустойчивость локуса формант согласных ясно обнаруживается при анализе взрывного согласного "g". В спектре этого согласного не удалось установить постоянного локуса для второй форманты. Он резко смещается по частотной шкале в зависимости от того, приносится за ним гласный переднего или заднего ряда.

-
1. Поэтому для исследования губных согласных верхний частотный лимит может превышать 8000 гц - частота, достаточная для рассмотрения остальных согласных.

Как показал целый ряд исследований, наиболее значимыми являются две первые форманты. Однако известно также, что всякая часть голосового тракта влияет в той или иной степени на настройку всех формант. Следует также отметить, что ротовой резонатор в большей степени влияет на F_1 , чем на F_2 . Уровень интенсивности формант определяется как источником звуковых колебаний, так и резонансной системой речевого тракта. Длительность звука определяется длительностью работы голосовых связок и артикуляционного аппарата.

Итак, существует два основных вида источников, источник периодических колебаний и источник шума. В зависимости от этого трактуются спектральная картина. На спектрограммах периодические колебания представлены характерной гармонической структурой, а шум представлен иррегулярным распределением энергии во времени.

Качество гласной определяется, главным образом, распределением энергии в спектре, что в значительной степени зависит от положения формант F_1, F_2, F_3 и т.д. Как показала практика спектрального анализа, конкретные формантные картинки одной и той же гласной зависят от многих факторов, включая индивидуальные особенности произнесения диктора, его пол, возраст, состояние здоровья, эмоциональное состояние, темп произнесения, ударения, позиционные факторы и т.д.

Однако в то же время существуют относительные структурные формантные признаки, которые отличают принадлежность конкретной формантной картинки одного звука от формантной картинки другого звука.

Несмотря на те сомнения, которые возникают при интерпретации спектральных данных с точки зрения их соотнесения с реальной речью, можно все же убедиться в пользе спектрографического представления с точки зрения возможности установить непосредственную связь между конкретным видом спектрограммы, особенностями воспринимаемого звука и фонетическими единицами.

Таким образом, формантный анализ речи рассматривается обычно как особый вид спектрального анализа. Задача, стоящая перед исследователем, заключается в определении частот формант речеобразующего механизма. Как указывалось выше, эти изменения обусловлены перестройкой уклада артикуляционных органов в процессе речепроизводства. Один из возможных подходов при этом заключается в ориентации на резонансные максимумы спектра, которые не являются жестко закрепленными и могут изменяться и перемещаться. Определенным артикуляционным признакам могут соответствовать определенные акустические признаки:

Например:	Открытость	компактность
	закрытость	диффузность
	переднерядность	высокая тональность
	заднерядность	низкая тональность
	огубленность	бемольность
	неогубленность	простая тональность
	палатальность	дизезность
	непалатальность	простая тональность

Акустич. признаки: *интенсивность*

847

В ходе спектрального анализа исследователь исходит из того, что "акустические представления неразрывно связаны с представлениями двигательных органов речи, необходимых для осуществления соответствующего звука". +)

+) Л.В. Щерба. Избранные работы по языковедению и фонетике. т. I, Ленинград, 1958, стр. 134.

Лекция 3. ОПОРНЫЕ ПРИЗНАКИ СЕГМЕНТАЦИИ РЕЧИ НА СПЕКТРОГРАММАХ

Как показала практика экспериментально-фонетических исследований, в процессе членения, иными словами, сегментации речи целесообразно идти от описания границ более крупных речевых сегментов (сверхфразового единства, фразы, синтагмы, ритмической структуры) к описанию признаков границ таких сегментов, как слог, звук, субзвуковой сегмент. Все эти сегменты можно рассматривать, как разномасштабные сегменты речевого текста. Вычленение их носит условный характер. В процессе реализации речи природа соотношения между выделенными сегментами представляется намного сложнее.

Сегментация речи является одной из кардинальных проблем теоретической и практической фонетики. Проблема сегментации заключается в поиске относительного соответствия между дискретными единицами языка и участками непрерывной речевой волны. Исследователь располагает информацией о временной последовательности единиц языка и физически реальным временем речевой волны. При этом при сегментации исходят обычно из интерпретации характеристик речевой волны, относительно быстро меняющихся на определенных участках. Участки, на которых происходят наибольшие изменения, принято условно рассматривать в качестве границ; участки, находящиеся между этими границами, принято рассматривать как некоторые соответствия единицам языка. Сказанное выше означает, что сегментация наблюдаемой физи-

ческой субстанции исходит из положения о наличии временной выраженности единиц языка.

Как указывалось выше, в речи существует три типа источников возбуждения:

1. Первый служит для генерирования сонорных звуков и представляет собой пульсирующий воздушный поток, проходящий через вибрирующие голосовые связки.

2. Второй генерирует шум, возникающий в результате задержаний потока воздуха в суженных проходах. Результат — артикуляция фрикативных и аспирированных звуков.

Известно, что максимальная эффективность шумового источника достигается в том случае, когда воздух с большой скоростью ударяется об острый угол или же в шероховатую поверхность.

3. Третий генерирует эффект, наблюдающийся при резком спаде давления после раскрытия закрытого до этого прохода в речевом тракте, что имеет место при артикуляции смычных.

Передаточная функция речевого тракта или функция фильтрации $T(f)$ включает три основных компонента $T(f) = T(\text{плюсов}) + T(\text{нулей}) + T(\text{излучения})$. Функция излучения присутствует всегда, она прямо пропорциональна частоте и обратно пропорциональна расстоянию от рта говорящего до микрофона. Зависимость от расстояния учитывается понижением звукового давления на 6 дБ с удвоением расстояния до микрофона. Функция излучения характеризует интенсивность речевого сигнала.

Передаточная функция включает далее функцию полюсов $T(\text{полюсов})$. $T(\text{полюсов})$ является обязательной компонентой для всех типов звука. Выражается эта функция суммой резонансных

кривых, данных в двух измерениях: частота (Гц) (или технические единицы) – интенсивность (дБ).

Функция нулей T (нулей) описывается в виде набора антирезонансных кривых. Антирезонанс (или в математическом понимании нуль) нейтрализует действие резонанса на этой же частоте и ослабляет близлежащие составляющие. Антирезонансы появляются:

а) при наличии разветвления воздушного потока;

б) когда источник возбуждения находится не в самом конце речевого тракта. Таким образом, антирезонансы присутствуют при образовании назализованных гласных, сонорных согласных и всех звуков, получающихся с помощью надглоточного источника: фрикативных, аффрикат и взрывных.

Неназализованные гласные звуки описываются характеристикой излучения и характеристикой резонансов, т.е. положением F_1 , F_2 , F_3 , F_4 и т.д. (описание с помощью F-картинки). (Для видимой речи еще и полосами на уровне 3 дБ от центрального (пикового) значения. Нули (или антирезонансы) характеризуют неназализованные гласные, сонанты и согласные: смычные, фрикативные, аффрикаты.

Нули больше проявляют себя, изменяя относительные уровни формант, нежели давая явные спектральные минимумы.

Если антирезонанс локализуется вблизи резонанса, то его влияние на изменение спектра звука невелико. Такой нуль считается связанным с полюсом. В отличие от связанного нуля существуют свободные нули. Свободный нуль располагается не так близко от полюса и заметен в качестве четкого минимума. Свободный нуль влияет на спектр звука значительно сильнее, чем связанный.

По аналогии с F-картинкой резонансов можно определить Г-картинку антирезонансов. Например, в спектре (*е*), произносимого при перекрытии языком средней части тракта (проходы остаются по бокам), возникает антирезонанс в диапазоне $F_2 - F_4$, который вместе с ближайшим полюсом в районе $F_4 - F_5$ придает звуку характерную окраску, отличающую его от гласного. Однако нуль, возникающий вследствие наличия воздушного пространства за кромкой языка, по данным эксперимента с синтезом, не имеет большого практического значения.

F_n - сонантов

Носовые характеризуются F_n : 250, 1000, 2000, 3000, 4000. Самая низкая - наиболее интенсивная. Форманта на 1000 гц не всегда может быть обнаружена на спектрограммах ввиду малой степени интенсивности. Более высокие форманты могут быть сдвинуты (2000 и т.д.), что имеет место, например в сочетании с задними гласными.

Форманты носовых обозначаются через N_1, N_2, N_3 и т.д.

Таким образом:

N_1 соответствует F_1

N_2 — F_2

N_3 — F_3

Для / *п* / антирезонанс в районе 1500 гц, для / *н* / - в районе 1000 гц.

Плавные

$F_1 = 210, F_2 = 1700; F_3 = 2500, F_4 = 3050$ (расчетные данные по Фанту)

$F_1 = 230, F_2 = 1600, F_3 = 2300, F_4 = 3100$

(измеренные данные /по Фанту/)

/r/ $F_1 = 500$

$F_2 = 1700$

$F_3 = 2400$

Целевые и аффрикаты

	F_1	F_2	F_3	F_4
x	500-600	1500	3000	7000
s	600-700	1000	2000	8000 10000
f	300-400	1000	3000	10000
ts	500	2000	6000	8000-10000
ty	300	1000	3000-4000	8000
j	500-600	2000-3000	5000	7000-8000

Взрывные (глухие)

k	150-200	1500	4000	7000-8000
p	200-300	700-800	2000	8000
t	300	2000	5000	8000-10000

Взрывные (звонкие)

g	- 200	2000	3000	10000
b	- 500	1000	3000-4000	8-10000
d	- 200	2000	4000-5000	8-10000

Некоторые согласные, например /r/, v, w, ~ часто произносятся как озвученные континанты. Почти совсем без шума, либо с незначительным его уровнем. Поэтому их можно рассматривать как гласные, т.е. их спектр достаточно точно описывается F - картиной.

/v/ $F_2 - 800-900$
 $F_3 - 2000-2500$
 $F_4 - 3500-4000$

/r/

F₂ - 1000-1500

F₃ - 2500-3000

F₄ - 4000

Речевой сигнал не соответствует нашему представлению о речи, как о последовательности дискретных постоянных элементов. То, что отражено на спектрограммах, является комбинацией непрерывных и дискретных процессов. Непрерывность артикуляции в процессе речепроизводства находит отражение в непрерывности F-картины.

Дискретность артикуляции (включение и выключение голосовых связок, полное или частичное закрытие рта, смыкание мягкого нёба) также отражается на спектрограммах. Сегментация не является однозначным соответствием между числом фонем в транскрипции и числом сегментов акустической картинки. Обычно число сегментов превышает число фонем транскрипции; но оно может оказаться и меньше числа фонем. Причиной такого несоответствия между лингвистическими единицами и элементами речевой волны является эффект коартикуляции. Кроме того фонетическая реализация сегмента (звука) зависит от непосредственного фонетического контекста.

Например, один взрывной смычный звук передается несколькими сегментами: начальный переходной сегмент, сегмент смычки, сегмент эксплозии, сегмент фрикции, конечный переходный сегмент. Аспираторный интервал рассматривается, как правило, в качестве составной части эксплозии (взрыва). В специальной литературе термин "взрыв" употребляется неоднозначно. Одни исследователи включают в сегмент взрыва раствор и фрикцию, другие - весь шумо-

вой сегмент смычного взрывного согласного. Аспираторный интервал, включенный в сегмент взрыва, дает возможность дифференцировать согласные по признаку; напряженные – ненапряженные. У напряженных взрывных сегмент эксплозии длительнее, если включить в его состав аспирацию..

Связная речь может быть разделена на последовательность сегментов, отражающих признаки способа и места образования звуков.

I. Характеристики сегментов, отражающих
признаки способа образования

№ пп	Название сегмента	Характеристики артикуляции (образования)	Характеристики акустической картины (речевой волны)
I	2	3	4
I	Звонкий	Вибрации голосовых связок, вызывающие модуляции воздушного потока	Квазипериодическая, четкая по частоте и времени структура
2	Носовой	Турбулентный воздушный поток через узкие артикуляторные проходы	Относительно четкая по частоте и времени структура
3 °	Отсутствие звука (пауза)	Отсутствие активного состояния в голосовом тракте	Отсутствие спектральной энергии

1	2	3	4
4	Полугласный	Свободное прохождение воздуха через гортань и ротовую полость	Наличие F-картины (F_1 , F_2 , F_3 и т.д.)
5	Латеральный (боковой)	Центральное сжатие с латеральной щелью в ротовой полости	Спектральная картина, похожая на гласную, но с антирезонансом в области $F_2 - F_4$ и дополнительным резонансом в F_4 - области
6	Окклюдивный (смычный)	Полное сжатие в ротовой полости или гортани	В спектре звонкой смычки доминирует F_1 очень низкой частоты. Более высокие форманты обычно не просматриваются.
7	Назальный	Движение мягкого нёба, которое связывает в дан-	Преобладание формантного

ном случае назальную систему с остальной частью голосового тракта

максимума в районе 250 гц.

В районе 2200 гц слабая форманта; F_2 - очень слабая или вообще отсутствует

8 Фрикативный

Поток воздуха проходит через узкий канал, образованный в ротовой части или гортани

Наличие заполненной по всему спектру F - картины (вплоть до F_4 , но с более низкой частотой F_1 , чем у полугласных

По сравнению с глухим полугласным похожей артикуляции спектр фрикативного смещен в область более высоких частот

9 Промежуточный

Скользкая артикуляция "глайд" определяется артикуляторами, движущимися в пределах сегмента

Спектральная картина в сегменте меняется с высокой скоростью

II. Применение метода образования

Литература	Артикуляция	Речевая волна (акустический коррелят)
п/п		
I	2	3
1	Вытянутый вперед язык	Большое значение $F_2 - F_1$
2	Оттянутый назад язык	Малое значение $F_2 - F_1$
3	Узко открытый рот	Низкая F_1
4	Губы почти закрыты и вытянуты вперед (образуя маленькое ротовое отверстие)	$F_1 + F_2 + F_3$ меньше, чем при широком раскрытии губ и такой же артикуляции языка
5	Палатальная артикуляция	F_3 низкая и F_3 близка к F_2
6	Альвеолярная артикуляция	F_4 низкая и близка к F_3
7	Билабиальная артикуляция (и лабиодентальная)	F_2 в районе 500-1500 гц в зависимости от положения языка для соответствующего сегмента гласного. / ϕ / - низкая интенсивность шума (1500-10000 гц)
8	Междудентальная артикуляция	F_2 - 1400-1800 гц Фрикативный шум для / θ / более слабый, чем для / s / и имеет большую продолжительность
9	Зубная и преальвеолярная артикуляция	F_2 в районе 1400-1800 гц; F_3 - высокая Фрикативный шум интенсивный и высокой частоты;

10 Велярная (задненёбная)
и фарингальная артику-
ляция

Для /s/ - основная энергия
сосредоточена в районе 4-10
кГц.

F_2 - средняя или низкая.
Большую часть фрикативного шу-
ма несет F_2 .

F - картина относительно
отчетлива (искл. F_1)

При обработке спектрограмм задача сегментации значительно
облегчается тем, что исследователю уже заранее известен текст,
который следует сегментировать. Намного сложнее задача обработ-
ки спектрограмм без знания текста. Эта задача может быть сфор-
мулирована как задача декодирования акустического изображения
речи. Значение решения данной задачи велико, т.к. связано с
проблемой автоматического распознавания речи, а так же имеет
прямое отношение к проблеме восприятия речи.

Лекция 4. КООРТИКУЛЯЦИЯ В ПРОЦЕССЕ РЕЧЕПРОИЗВОДСТВА И ЕЕ СПЕКТРАЛЬНЫЕ КОРРЕЛЯТЫ

Исследование спектральной картины звуков речи включает два аспекта: изучение относительно устойчивых качеств реализации звука, взятого изолированно вне фонетического контекста, и изучение динамических качеств реализации звука, взятого в определенном фонетическом контексте в потоке речи.

Предметом первого аспекта является акустическое описание звуков в статике. Предметом второго аспекта — акустическое описание звуков в динамике.

В плане речеобразования последовательность артикуляторных жестов, порождающая (генерирующая) последовательность звуков, называется коартикуляцией. Эффект коартикуляции может быть описан в терминах спектрального анализа путем исследования характера изменения переходов формант.

Изучение коартикуляции велось и ведется в двух направлениях. Первое направление характеризуется тем, что в основу изучения эффекта коартикуляции положена теория "локуса". Сущность этой теории заключается в том, что с помощью психоакустических опытов с синтезированной речью подбираются такие характеристики формантного перехода, при которых согласный хорошо опознается только на основании этих данных даже в том случае, когда шум

взрыва отсутствует.

Теория локуса разработана сотрудниками Хаскинской лаборатории (США). Основное исходное положение этой теории было сформулировано А. Либерманом.

Для каждого согласного имеются характерные точки, в которых начинаются формантные переходы или к которым примыкают последние. На этом основании переходы могут рассматриваться как движение формант от характерной для каждого согласного точки к частотам последующего звука.

Конкретные данные, характеризующие "локус" согласных, были предложены Делаттром, Либерманом и Купером^{*)}. Эти данные были получены в результате проведения спектрального анализа (CT^{xx}) - сочетаний в американском варианте английского языка и были проверены путем эксперимента с синтезом и далее путем проведения психоакустических опытов по восприятию.

Эмпирически найденный "локус", например, для F_2 смычных взрывных звонких b, d, g составляет следующие значения:

b - 700 гц, d - 1800 гц, g - 3000 гц.

Дальнейшая разработка теории "локуса" привела к нахождению универсальной точки в районе 1400 гц.

Второе направление изучения эффекта коартикуляции ведется на основе трехпараметрической модели Г. Фанта:

*) Liberman A., Delattre P., Cooper F. The role of selected stimulus - variables in the perception of the unvoiced stop consonants. Amer. J. Psych., 65, 1952.

xx) CT - согласный + гласный

Напомним сущность трехпараметрической модели Г.Фанта:

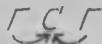
Три параметра:

- 1) положение минимального просвета x_{min} между языком и нёбом;
- 2) площадь поперечного сечения A_{min} в этой точке;
- 3) степень округления и вытянутости губ, измеренная в отношении к ℓ_1/A_1 , где ℓ_1 - длина секции, моделирующей губы, а A_1 - площадь сечения в районе губ.

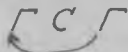
Исходя из того, что каждый участок резонаторной системы голосового тракта влияет на положение формант, можно утверждать, что с точки зрения эффекта коартикуляции формантные переходы в интервалах, прилегающих к согласному, есть не что иное, как переход от F -картины, характерной для моментов максимального сужения при артикуляции согласного, к F -картине гласного. Г.Фант (1964, стр.38) считает, что локус должен быть закреплен как термин для F -картины любого звука как характеристика предельного положения артикуляторных органов.

Итак, фонетическая реализация звука (или сегмента звука) зависит от непосредственного фонетического контекста. Проиллюстрируем это положение:

1. Изменение F -картины взрывного сегмента согласного в зависимости от F -картины предшествующего и последующего гласного:



2. Некоторое влияние конечной гласной на последнюю часть первой гласной:



3. Небольшое влияние согласной на F -картину квазистационарной части гласных (как начального, так и конечного в сочетании ГСГ):

Г С Г

Основной вывод - незначительный эффект коартикуляции в пределах квазистационарной части гласных, значительный эффект коартикуляции - на переходах от Г к С и от С к Г.

В работе Омана^{х)} исследуется с помощью спектрального анализа эффект коартикуляции в сочетаниях типа ГСГ, где С - смычный звонкий согласный. Исследование проводится на материале трех языков: шведского, американского варианта английского языка и русского. На основе данных предлагается модель коартикуляции гласных и согласных в сочетаниях ГСГ.

Что касается случаев английской и шведской речи, то конечные формантные частоты в сочетаниях ГСГ зависят не только от согласных, но и от контекста гласных. Локусы смычных согласных не постоянны. Причиной этого явления служит, по-видимому, то, что при образовании согласного вовлекаются сопутствующие артикуляторные движения, которые частично превосходят конфигурацию последующего гласного. Следовательно, первые импульсы - команды для согласного и последующего гласного должны происходить одновременно. Однако некоторые компоненты конца гласного, видимо, задерживаются еще в период произнесения последующего согласного (рис.3).

Таким образом, С.Оман в результате своего исследования приходит к выводу о том, что сами значения локусов очень переменны и не могут быть описаны тремя, двумя, а тем более одним зна-

х) S. Ohman. Coarticulation in VCV - Utterances. JASA, 1969,

г.39, 111.

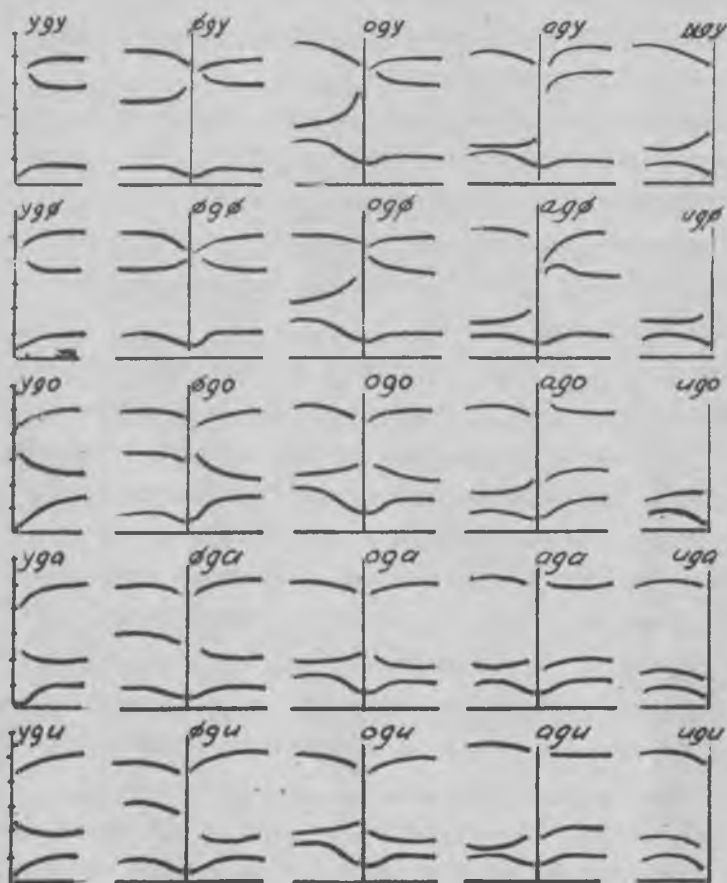


РИС. 3. ИЗМЕНЕНИЯ ФОРМАНТ В СОЧЕТАНИЯХ ГСГ (ШВЕДСКИЙ ЯЗЫК)

чением, как это предполагалось сотрудниками Хаскинской лаборатории. Оман предлагает пересмотреть теорию локуса и считать локусом "среднее число от нескольких измерений, которое получается при учете всех возможных контекстов в ГСТ. На основе полученных данных Оман предлагает модель коартикуляции типа ГСТ. Входом в эту систему является последовательность "нервных" команд, отбирающих два гласных, согласный и сигналы для согласования во времени согласных и гласных звуков. Выход этой системы - это непрерывно меняющийся коэффициент формы голосового тракта.

Языковая коартикуляция сочетания ГСТ (где С - смычный) включает два одновременных движения: а) движение языка, характеризующее гласный и б) наложенное на эту артикуляцию, сжимающее движение апикальных, дорсальных, а также губных артикуляторов, что вызвано наличием смычного согласного в артикулируемом сочетании. В модели Омана процесс образования гласной включает артикуляцию гласной от начала до конца плюс наложенная на нее артикуляция согласного. Степень выраженности эффекта коартикуляции зависит от характера изменения голосового тракта и от времени реализации.

По-видимому, эти наблюдения могут быть изложены в терминах модели синтеза артикулируемой речи, в которой учтены:

- 1) апикальные сближения,
- 2) дорсальные сближения,
- 3) положения тела языка.

Данные свидетельствуют о том, что существуют значительные отклонения в формантах при переходах.

Максимальные значения зоны отклонения для F_1, F_2, F_3 равны 130, 540, 330 кгц. Средние значения - 62, 280 и 153 кгц. Интересно, что наибольшая зона отклонений наблюдается у F_2 перехода от -гу и т.д. к угу = 540 кгц. Подобное отклонение сохраняется у F_2 в ~~φ~~гу - и -φφ. 50% значений зоны отклонений у F_3 и 50% - у F_1 больше, чем критическая (нормируемая) величина.

В случае с/а/ - зоны отклонений пограничных частот в переходах у всех трех формант значительны (при всех окружениях).

Таким образом, данные таблицы 5 свидетельствуют о том, что конечные частоты в формантных переходах в сочетаниях типа ГС и СТ должны рассматриваться в речевом материале не изолированно, не как самостоятельная величина, а в зависимости от F -модели гласного.

Таблица 5. Значения формантных переходов.

Сочетания		Г b Г		Г d Г		Г o Г	
Гласные		ГС	СТ	ГС	СТ	ГС	СТ
у	у	-450	-350	-155	-130	-15	+140
	φ	-515	-370	-290	-200	-30	+55
	а	-775	-370	-370	-370	-280	+115
	и	-650	-390	-200	-190	-325	+100
φ	у	-375	-65	-90	+75	0	+430
	φ	-415	-175	-200	-5	-25	+375

	<i>a</i>	-465	-90	-340	-25	-275	+340
	<i>u</i>	-515	-140	-240	+40	-365	+315
<i>a</i>	<i>y</i>	+100	+205	+265	+520	+175	+330
	<i>ø</i>	+25	+75	+240	+365	+125	+225
	<i>a</i>	-100	-100	+165	+290	+30	+165
	<i>u</i>	-150	-55	+250	+315	-65	+80
<i>u</i>	<i>y</i>	+200	+255	+575	+940	+190	+115
	<i>ø</i>	+80	+325	+530	+640	+65	+255
	<i>a</i>	+15	+200	+490	+370	+15	+190
	<i>u</i>	+5	-5	+490	+690	-50	-15

Таблица № 6. Разность $\bar{F}_2$ - пограничной и

$\bar{F}_2$ - стационарной части гласного ($\bar{F}_{2t} - \bar{F}_{2v}$)
(усредненно) в ГС-комбинациях (в КЦ).

$\epsilon b'$	ϵb	$\epsilon d'$	ϵd	$\epsilon q'$	ϵq
120	90	80	60	45	570

Во всех случаях, за исключением /-g/, отклонения в зависимости от палатальных - непалатальных, слишком малы, чтобы их учитывать. Тип коартикуляции, характерной для шведского и английского языков, не был характерен для русского языка.

Была изучена коартикуляция в сочетаниях типа СГУ в зависимости от темпа произнесения и силы ударения. Чем длительнее гласная, тем больше значение $\bar{F}_2$ будет приближаться к фиксированному значению, инвариантному относительно окружающих согласных. Чем менее длительна гласная, тем меньше $\bar{F}_2$ будет отличаться от ее значения в $\bar{F}$ -картине согласного.

Из приведенных исследований следует:

1) эффект коартикуляции первичен на артикуляторном уровне и вторичен - на акустическом уровне;

2) эффект коартикуляции может быть достаточно точно описан в терминах спектрального анализа речи;

3) акустическими коррелятами эффекта коартикуляции являются измерения формантных переходов: F_1 , F_2 , F_3

4) Наиболее показательной характеристикой коартикуляции на акустическом уровне является изменение ;

5) формантный переход может быть измерен от начала изменения значения форманты до момента относительной стабилизации значения той же форманты на квазистационарном участке звука;

6) изменения значений формант (F_1, F_2, F_3) на переходных участках могут быть описаны в терминах "зоны отклонения"

($F_{it} - F_{iv}$) с указанием направления изменения (+ или -);

7) степень выраженности формантных переходов в значительной мере зависит от длительности реализации звука, что находится в прямой зависимости от темпа произнесения и степени ударения.

Исследователи речи отмечают, что фонетическая изменчивость на артикуляторном и акустическом уровнях не отрицает существования неизменных фонематических команд на нейронном уровне речеобразования. Есть все основания верить, что факторы, обуславливающие фонетическую изменчивость, могут быть полностью описаны в виде моделей, что даст возможность предсказания физического речевого явления по транскрипции лингвистического сообщения.

ЛЕКЦИЯ 5. СИНТЕЗ РЕЧЕВЫХ СЕГМЕНТОВ ПО ДАННЫМ СПЕКТРАЛЬНОГО АНАЛИЗА

Фонетические исследования посвящены преимущественно выявлению и классификации акустических моделей, играющих большую роль в сфере коммуникации между говорящим и слушающим. С самого начала экспериментальная фонетика изучала, в основном, артикуляцию, затем большое внимание было уделено акустической структуре речевых сигналов.

С помощью психологов, физиологов, специалистов по теории коммуникации и лингвистов проводятся многочисленные объединенные исследования, направленные на более полное изучение процесса восприятия. С технической точки зрения подобные эксперименты были до недавнего времени затруднены, т.к. естественные стимулы с трудом поддаются контролю. В наши дни можно использовать в качестве материала для исследований в области восприятия различные синтезирующие устройства. Их применение в современной фонетике достаточно широко.

Синтез речи представляет большой интерес для фонетистов ввиду возможности установления точного контроля над речевыми параметрами.

Разработанная в настоящее время методика синтеза речи позволяет сделать описание на фонологическом уровне. С этой точки зрения фонология становится фактически сводом правил для синтеза с точной процедурой следования от последовательности фонем к их актуализации в качестве звуков речи.

Процедура синтеза, которая ведет к фонетическому описанию в терминах акустических правил для синтеза, позволяет получить и другие данные, представляющие интерес для лингвистов, хотя

и не являющиеся для него вопросами первостепенной важности. Так, например, можно выявить, что различные классы фонем обладают различными перцептивными свойствами; можно полагать, что благодаря этому они несут различную нагрузку в передаче информации, что и определяет их лингвистическую роль.

Мы различаем два типа сегментных единиц и, таким образом, два уровня описания: фонетический и фонологический. Сугубо фонетические правила определяют группу сегментов (фонев) на основе их артикуляционных и/или акустических характеристик; фонологические правила организуют фонев в единицы более высокого порядка (фонемы) на основе как фонетических, так и дистрибуционных принципов. В конечном итоге, правила последнего рода касаются определенных реляционных свойств фонев, рассматриваемых как абстрактные элементы какого-либо ряда.

Возможный метод, позволяющий осуществить замену сегментов (фонев) имеет дело не с речью как таковой, а с приближением к речи — синтетической речью. Речевые синтезаторы позволяют строго контролировать изменения акустического сигнала и, таким образом, дают возможность определить, какие именно акустические характеристики служат для идентификации сигнала. Мы можем, например, задержаться на определенном временном участке сигнала, изменить эту часть сигнала определенным образом и получить при этом серию сигналов-стимулов, среди которых акустические варианты воспринимаются как более или менее постепенный сдвиг в фонетическом качестве, и потом подавать эти стимулы носителям языка для идентификации в качестве высказываний этого языка.

В случае, если говорящие на данном языке различают стимулы, их отреты дают возможность указать на определенные признаки в акустической модели, которые выполняют различительную функцию.

Используя устройства, искусственно генерирующие речь, мы можем не только синтезировать речь, в достаточной мере приближающуюся к естественной; мы можем получать синтезированную продукцию, которая в известном смысле превосходит человеческую речь.

Информация, подаваемая на синтезатор, может быть сформулирована в виде свода правил, и эти правила можно подавать на синтезатор в любом желаемом порядке для того, чтобы получить на выходе разборчивую, похожую на речь последовательность. Такие правила, по-существу, являются функциональными. В процессе преобразования правила синтеза порождают субправила, которые предусматривают, что каждая фонема синтезированного высказывания будет репрезент^{на}рована ее соответствующим аллофоном. Таким образом, техника синтеза речи впервые дает возможность использовать тексты на заменяемость, как возможный способ определения отношений между фонемами, и тем самым проверить фонологические описания.

Практическое значение метода синтезирования состоит в том, что он может быть использован для проверки некоторых аспектов, фонологических описаний. Таким образом, гипотеза о роли определенных признаков (свойств) (например, лабиальности или назальности) в установлении "фонетического сходства" для естественных носителей языка может быть подтверждена экспериментальным путем: синтетическая спектрограмма может быть составлена так, чтобы она включала (или опускала) характерный признак, сохраняя все остальные признаки. Звуки, производимые с помощью этих

правил, могут быть затем поданы носителям языка для идентификации. Очевидно, что этот метод зависит от возможности преобразовать субфонемный признак в эксплицитное различие в схеме-спектрограмме. Это возможно только в той степени, в которой объем описания правил для синтеза соответствует объему фонологического описания, которое проверяется на практике; соответствие, близкое, но неполное.

Другой аспект внутренней структуры правил выявляется тогда, когда мы интересуемся не отношениями между правилами, а процессом преобразования транскрипции в звуки через схему-спектрограмму и синтезатор. Тогда мы можем классифицировать индивидуальные инструкции, составляющие правила в соответствии с их функцией в процессе синтеза. Инструкции характеризуют такие понятия как частота локуса и частота форманты. Другая группа специфических признаков нужна для того, чтобы обеспечить непрерывность элементов моделей и, таким образом, соответствовать непрерывному характеру артикуляционных движений; эти "соединительные инструкции", по существу настолько просты, что они оказываются имплицитно включенными в правила. Другие группы специфических признаков, которые мы назовем "позиционными определителями", определяются контекстом, т.е. определенной последовательностью, которая передается транскрипцией.

В действительности, функция этих позиционных модификаторов, действующих наряду с основными инструкциями, состоит в выполнении требования, заключающегося в том, чтобы правила для синтеза, подобно фонемам, могли бы быть свободно заменяемыми.

Из трех видов инструкций, которые входят в правила синтеза, основные инструкции можно рассматривать, как акустические признаки, которые соответствуют фонемам. Намного сложнее показать, каким образом другие типы инструкций соотносятся с аспектами фонологического описания. В частности, позиционные модификаторы могут быть связаны с теми фонетическими различиями, которые характеризуют аллофоны фонем; мы можем сказать, что позиционные модификаторы вместе с главными инструкциями обеспечивают акустические специфические характеристики — "субправила", которые относятся к аллофону, как правило к фонеме. Более точно позиционный модификатор находит свой аналог в описании, которое касается субфонемной вариативности как результата слияния (перекрывания) между соседними фонемами. Что же касается соединительных инструкций, то мы находим, что их функция заключается в обеспечении непрерывности модели. Она не соответствует ни одному из аспектов фонетического описания, так как фонетическое описание не должно эксплицировать непрерывность артикуляторного движения, которую физическая природа голосового тракта вызывает говорящему. Соединительные инструкции необходимы просто потому, что мы выбрали акустический способ описания для извлечения речи, а отсюда вытекает, что синтезатор, который использует описание речи, должен отвечать акустическому "языку на входе". Если, например, мы бы выбрали артикуляторный способ описания, мы должны были бы использовать синтезатор типа речевого аналога. Такой тип синтезатора может быть создан с встроенными в него ограничениями подобно тем, которые существуют в человеческом артикуляционном механизме, и "языку на входе" такого

синтезатора не потребуются соединительные инструкции, которые необходимы как часть акустического описания. Дело в том, что соединительные и позиционные правила характеризуют, в основном, отношения не между индивидуальными правилами, а между классами правил.

Небезинтересно сравнить фонематическую репрезентацию высказывания с последовательностью правил для этого же высказывания. Точные соответствия между фонемами и правилами для синтеза не могут быть определены априорно, для этого мы должны предположить, что степень соответствия будет зависеть от критерия, на основании которого каждый из элементов выделяется при анализе, а также, возможно, от языка, который является предметом описания. Возникает, однако, несколько довольно общих вопросов, на которые следует ответить: 1) что вовлечено в процесс исследования от фонематической транскрипции к акустическому сигналу; 2) существует ли однозначное соответствие между сегментами двух уровней. Когда мы рассматриваем преобразование транскрипции в слышимое высказывание, оказывается, что существует две стадии в этом процессе, независимо от того, осуществляется ли он живым существом или синтезатором. В первом случае происходит замена фонемы на соответствующий аллофон; во втором случае синтезатор должен применить правила (особенно позиционные модификаторы), которые изменяют процесс синтеза, учитывая специфику контекста, в котором это правило было сформулировано. Итак, в функциональном смысле оба вида преобразования символа в звук являются операцией из двух ступеней, которые параллельны между собой.

Выводы относительно правил синтеза можно суммировать следующим образом: экономия описания и использование правил, которые являются дискретными, свободно заменяемыми и эксплицитно преобразуемыми в звук. Правила находятся в однозначном соответствии с фонемами, описание которых известно. Соответствие на субфонемном уровне предоставляет полезный метод для проведения опытов по восприятию. Отсутствие однозначного соответствия и иногда требование "позиционных модификаторов" указывает на некоторую более далекую связь между акустическим описанием и идеальной фонемой. Мы должны добавить, что никакие описания не достигают идеала, т.к. засвидетельствовано отсутствие прямого одноступенчатого процесса преобразования фонологических единиц в речь.

В области синтеза речи по правилам был достигнут значительный прогресс. Сконструированный электронный синтезатор сам по себе способен синтезировать речь высокого качества (высокой степени разборчивости). Программа для ЭВМ, разработанная с целью устранения громоздкой операции расчета значений параметров, управляет синтезатором и осуществляет синтез больших отрезков речи. Программа включает различные технические приемы для вычисления параметров, которые встречаются в высказываниях и являются достаточно разнообразными. Фонетическая информация, необходимая для программы, включается в определенную серию таблиц "на входе", в которые быстро и легко могут быть внесены изменения, являющиеся результатом аудирования. При синтезе речи по правилам на вход синтезирующего устройства

подается фонематическая транскрипция высказывания. Успешный синтез высказывания заключается не в том, чтобы оно явилось адекватным воспроизведением реального высказывания носителя языка, но в том, чтобы оно воспринималось носителями данного языка (диалекта) как возможное в этом языке (диалекте).

Программа для ЭВМ — это контрольная лента с двумя входами. На один из них подается предложение, манифестирующее собой высказывание, которое должно быть синтезировано. Предложение состоит из серии буквенных символов, обозначающих "фонематические сегменты"; фонематический сегмент — операционный эквивалент фонемы — является основным элементом программы "на вход".

Другая программа "на вход" — это серия таблиц, причем для каждого отдельного элемента существует своя таблица. В таблицу входит длительность сегмента за каждые 10 мсек, значения параметров (кроме f_0) стационарного участка этого сегмента и характеристики переходов.

При составлении таблиц используют данные спектрального анализа. Таблицы могут существенно модифицироваться, о чем свидетельствует слуховая оценка высказываний, полученных с помощью синтеза.

Соответствие между фонемами и фонетическими сегментами неполное. В некоторых случаях ~~последовательность~~ из двух или трех сегментов обозначает одну фонему, а аллофоны одной фонемы могут

быть обозначены различными сегментами. Три сегмента вообще не соответствуют фонемам: пауза, длительность которой равна 100мсек; конец предложения, поданного "на вход"; элемент нулевой длительности, используемый при синтезе звонких фрикативных. Сегментация речевого потока, предусматривающая заданные длительности для дискретных сегментов, есть всего лишь удобное для программирования представление.

Каждый программный цикл порождает последовательность значений параметров, необходимых для синтеза одного сегмента, количество таких последовательностей равно длительности сегмента. Значения параметров каждого сегмента (кроме F_0) зависят от информации, заданной для этого сегмента и от соседства с другими (предшествующим и последующим сегментами в высказывании). Частота основного тона (F_0), как правило, определяется один раз для каждого элемента. F_0 - есть величина на границе между предшествующим и данным сегментом. Программа определяет значение F_0 между пограничными величинами при помощи интерполяции^{*)}.

Значения для каждого из параметров $F_1, F_2, F_3; A_1, A_2, A_3$ (где F_i - значения формант, A_i - значения амплитуд) должны быть определены для начального перехода, для стационарной части и конечного перехода. Программа может создавать многообразие переходов, зависящее от сегментов, созданных в таблице "на вход".

*) В математике интерполяцией называется всякий способ, с помощью которого по таблице, содержащей некоторые числовые данные, можно найти промежуточные результаты, которые непосредственно не даны в таблице. Латинское слово "интерполяция" в переводе означает "вставка внутрь". Нахождение промежуточных значений функции по ее графику называется графической интерполяцией.

Выше говорилось о том, что некоторые фонемы рассматриваются как последовательности сегментов. Например, дифтонг представляется последовательностью из двух гласных сегментов. В таблице "на вход" для первого сегмента, который является доминантой, точно определяется длительность перехода. Взрывной представляет последовательность нескольких сегментов, соответствующих периоду смычки, периоду возрастания энергии в начале отступа и периоду спада энергии в конце отступа.

При синтезе звонких ффрикативных эти звуки рассматриваются как последовательности из трех сегментов. Первый сегмент, длительность которого равна половине длительности всего сегмента, всегда звонкий. Второй сегмент нулевой длительности. Третий сегмент реализуется специальным действием программы, если следующий за ним глухой согласный; в противном случае он идентичен первому сегменту. Глухая аффриката представлена последовательностью из двух сегментов. Звонкая аффриката представлена 4-х сегментной последовательностью. В варианте, синтезированном по правилам, значения стационарного участка наиболее далеки от средних и малых значений, длительность стационарных участков имеет тенденцию к увеличению, а переходы вследствие этого становятся более резкими. В результате получается, что этот вариант, хотя вполне разборчивый, звучит не столь естественно. Причина этого заключается в том, что значения стационарных периодов сегментов в таблицах "на вход" получены, в конечном счете, в результате экспериментов с изолированными словами или словами из специальных контекстов. В звучащей речи формантные частоты имеют тенденцию к централизации благодаря инерции артикуляторов. Централизация значений формантных частот должна способствовать повышению

надежности результатов.

Информация для подачи "на вход":

1. Символ, обозначающий фонему.
2. Символ, обозначающий сегмент.
3. Ранг сегмента.
4. Стандартная длительность сегмента.
5. Длительность сегмента ударной позиции (для гласных).
6. Условные переключения, которые относятся либо к шумовому генератору, либо генератору импульсов.

7. Частотные значения для F_1 , F_2 , F_3 .

8. Установленная величина для пограничного значения соответственно F_1 , F_2 , F_3 .

9. Часть стационарного перехода соседнего сегмента, который прибавляется к установленной величине для получения пограничной величины соответственно для F_1 , F_2 , F_3 .

10. Длительность внешнего перехода F_1 , F_2 , F_3 .

11. Длительность внутреннего перехода F_1 , F_2 , F_3 .

12. Амплитуда для A_1 , A_2 , A_3 , A_w .

Установленная пограничная величина для A_1 , A_2 , A_3 , A_w .

13. Часть стационарного перехода соседнего сегмента, которая прибавляется к установленной для того, чтобы вычислить пограничную величину для A_1 , A_2 , A_3 , A_w .

14. Длительность внешнего перехода A_1 , A_2 , A_3 , A_w .

15. Длительность внутреннего перехода A_1 , A_2 , A_3 , A_w .

ЛЕКЦИЯ 6. ИСПОЛЬЗОВАНИЕ ДАННЫХ СПЕКТРАЛЬНОГО АНАЛИЗА В РЕЧЕВЫХ ИССЛЕДОВАНИЯХ

Проведение спектрального анализа позволило исследователям получить ценнейшие данные, которые легли в основу акустической теории речеобразования. Исследования, проводимые с помощью спектрального анализа, дифференцировались по двум направлениям: исследования речевого сегмента в статике и исследования речевых сегментов в динамике. На смену количественному описанию спектра с позиций наблюдателя пришло активное экспериментирование, проверка данных анализа с помощью синтеза.

При исследовании гласных и согласных на материале различных языков были определены типичные спектральные контуры с характерной локализацией формант, проводились работы по поиску и определению значимости различных спектральных параметров для распознавания фонем, осуществлялся поиск полезных признаков сегментации речевого высказывания, велись исследования с целью определения ключевых признаков декодирования речи, что особенно важно для решения проблемы автоматического распознавания речи.

Благодаря результатам спектрального анализа удалось установить, что вследствие вариативности речевого сигнала вряд ли возможно надежно выделить и опознать фонемы без обращения к более крупным сегментам речи — слогам, ритмическим структурам.

Следует отметить, что спектральный анализ вкупе со слуховым анализом и синтезом позволил определить специфику сегментного строя различных языков. Приведем некоторые примеры.

Известен тот огромный интерес, который проявляют лингвисты

к изучению проблемы назализации. Известно также, что попытки изучения назализованных гласных были довольно-таки многочисленными. В приводимом примере анализ назализованных гласных проводился с помощью спектрографа "Видимая речь". Авторы исследования задались целью проверить изменения на спектрограммах, возникающие при произнесении японских гласных без эффекта назализации. В результате эксперимента удалось установить, что на спектрограммах отражены следующие признаки назализации:

а) увеличение интенсивности резонансных частот в районе 250 гц.

б) ослабление интенсивности резонансных частот в районе 500 гц,

в) появление относительно слабых, но рассеянных по всему району от 1000 до 2500 гц дополнительных резонансных частот.

Целый ряд экспериментов в отечественных лабораториях был так же проведен с помощью "Видимой речи". Например, исследования сотрудников Лаборатории экспериментальной фонетики им.Л.В. Щербы при ЛГУ им.Жданова.

В частности, ими исследовались спектральные характеристики русских безударных гласных. В результате этих исследований были установлены общие качественные отличия безударного вокализма русской речи:

а) безударный /а/ более закрытый, чем ударный /а/, что определялось по положению первой форманты;

б) безударный /и/ более открытый и менее продвинул вперед, чем ударный, что определялось по более высокому положению первой форманты и более низкому положению второй форманты;

в) безударный /у/ более открытый, чем ударный, что опреде-

лялось по более высокому положению первой форманты, и более продвинуто вперед, что определялось по относительно высокому положению второй форманты;

г) безударный /н/ /значительных различий в своей формантной структуре по сравнению с формантной структурой ударного /н/ не обнаружил.

Кроме того, было определено, что чем больше редукция, тем больше шумовых составляющих в спектре гласных.

Исследование модификаций составляющих слога в русской речи было выполнено с помощью "Видимой речи". Модификация элементов слога рассматривались с точки зрения изменения некоторых акустических характеристик.

В результате исследования обнаружено следующее:

а) гласные существенно изменяются в зависимости от тех или иных качеств начального согласного;

б) наиболее значительны изменения гласных после носовых согласных;

в) значительно изменяется переход от согласного к стационарной части гласного;

г) наиболее ярко выражен переход от мягкого согласного к гласному;

д) переходный участок гласного изменяется также в зависимости от места образования согласного^{*)}.

*) Бондарко Л.В. Слоговая структура речи и дифференциальные признаки фонем. Дд, Ленинград, 1969.

Все перечисленные изменения были характерны для слога типа СТ. Что же касается слогов типа ГС, то здесь заимордление элементов носило несколько иной характер.

Подводя итог вышесказанному, можно сделать вывод, что "Видимая речь" является значительной вехой в развитии фонетических знаний.

Число фонетических исследований, проведенных с помощью этого прибора, огромно. Несмотря на известную огрубленность получаемых исследователем характеристик, данные "Видимой речи" широко используются не только в чисто теоретических целях, но и в прикладных областях. Так, например, для разработки практических методов автоматического распознавания речи, ввода в синтезатор и т.д.

В качестве примера следует привести результаты исследования гласных польского языка. Исследовались первые две форманты польского языка и восприятие синтезированных гласных стимулов, составленных на базе данных спектрального анализа. Известно, что в польском языке 6 гласных. Гласные начитывались 14 дикторами:

а) изолированно ($t = 1$ сек.);

б) в сочетании СГС $c_1 = b$
 $c_2 = t$

Исследование проводилось поэтапно:

I этап - анализ;

II этап - синтез по полученным значениям формантных областей;

III этап - аудиторский анализ синтезированных гласных стимулов.

Синтезированный стимул предъявлялся с $t = 0,6$ сек и с

комбинациями F_1 и F_2 . Всевозможные значения F_1 :

250, 350, 450, 550, 650, 750, 850 гц

Всевозможные значения F_2 :

800, 1000, 1200, 1400, 1600, 1800, 2000, 2200, 2400,

2600 гц.

Всего было получено 70 комбинаций F_1 и F_2 , но в эксперименте применялось 69, т.к. практически невозможна комбинация 850 гц и 800 гц.

Форманты F_3 и F_4 были заготовлены заранее, они использовались как константы: $F_3 = 2800$ гц, $F_4 = 3700$ гц.

F_5 варьировала в пределах 110-130 гц.

Ответы аудиторов фиксировались в качестве идентифицируемых гласных польского языка. (См. таблицы I, 2).

Таблица I. Определение аудиторами синтезированных гласных стимулов.

Стимулы.		Опознаваемые гласные					
F_1	F_2	i	ɨ	e	α	o	u
250	2600	80	-	-	-	-	-
350	2600	46	14	2	-	-	-
450	2600	-	24	27	-	-	-
550	2600	-	2	59	-	-	-
650	2600	-	-	67	-	-	-
750	2600	-	-	62	-	-	-
850	2600	-	-	51	-	-	-
250	2400	79	-	-	-	-	-
350	2400	23	45	-	-	-	-

450	2400	2	16	42	-	-	-
550	2400	-	-	60	-	-	-
650	2400	-	-	72	-	-	-
750	2400	-	-	70	-	-	-
850	2400	-	-	58	-	-	-
250	2200	41	-	-	-	-	-
350	2200	7	62	-	-	-	-
450	2200	-	25	25	2	-	-
550	2200	-	2	67	-	-	-
650	2200	-	-	76	-	-	-
750	2200	-	-	81	-	-	-
850	2200	-	-	63	-	-	-

Наибольшее число однозначно опознаваемого гласного дает 100%-ную идентификацию гласного.

Таблица 2. Характеристики F_1 и F_2 польских гласных, определенные при категоризации синтезированных гласных стимулов.

F_n	<i>i</i>	<i>ɨ</i>	<i>e</i>	<i>a</i>	<i>o</i>	<i>u</i>
F_1 (гц)	250	380	610	(820)	570	320
F_2 (гц)	(2550)	2000	2070	1140	(840)	800

Вывод: сравнение данных спектрального и аудитивного анализа показало, что существенных различий между данными нет. Некоторые расхождения возникают вследствие влияния F_3 .

Спектральный анализ голландских дифтонгов показал, что

дальнейший перцептивный анализ свидетельствует о разграничении двух компонентов без наличия непосредственного глайда (перехода), постулируемого в традиционной литературе. Синтез дифтонгов был успешным: при использовании спектральных данных двух устойчивых частей гласных при условии небольшого перекрытия во времени. В результате исследования был получен вырод- приемлемый синтез дифтонгов возможен без использования частот переходов.

Далее следует привести примеры использования метода спектрального анализа при исследовании согласных. Сравнительно недавно был проведен анализ твердых и мягких согласных русского языка в связной речи с помощью "Видимой речи". Авторы исследования сформулировали три основные гипотезы о качестве твердых и мягких согласных русской речи:

а) положение F_2 для мягких согласных всегда выше, чем для твердых;

б) соседство различных гласных /фонетический контекст/ не влияет на более высокое положение F_2 у мягких согласных по сравнению с твердыми;

в) F_2 у мягких согласных всегда выше, чем F_2 у предшествующих и последующих гласных, т.е. переход от гласного к согласному всегда позитивный, а переход от согласного к гласному всегда негативный.

В качестве анализируемого материала использовались спектрограммы, записанные с помощью широких фильтров. Измерялись частоты F_2 и F_3 на участке в пределах согласных в месте максимального уровня интенсивности. Результаты проведенного эксперимента позволили отбросить третью гипотезу как несостоятельную. Было

установлено, что направление перехода от гласного к согласному и от согласного к гласному далеко не инвариантно. В частности, было обнаружено, что переходы второй *ф*орманты от гласного к согласному, в основном, положительны, но могут быть случаи наличия отрицательных переходов. Соответственно не однозначны переходы и от согласного к гласному.

Что же касается двух первых гипотез, то здесь эксперимент подтвердил их состоятельность. Следовательно, было установлено, что вторая *ф*орманта мягких согласных всегда выше второй *ф*орманты твердых согласных вне зависимости от *ф*онетического контекста.

Другим примером использования спектрального анализа является исследование сонантов русского языка с последовательным разделением их на твердые — мягкие и с использованием различного *ф*онетического контекста. В результате этого исследования были выявлены характерные области концентрации энергии для твердых и мягких сонантов русского языка применительно к различным *ф*онетическим контекстам.

В качестве примера прикладного спектрального исследования приведем работу в области патологии речи.^{*)} Основная задача данного исследования заключалась в разработке объективного метода оценки качественных нарушений речепроизводства в результате патологических изменений органов артикуляции. Исследование было направлено на решение так же целого ряда частных задач, связанных с анализом речевого сигнала у больных и здоровых испытуемых, анализом акустической /спектральной/ картины согласных до и после лечения и т.д.

Результаты исследования показали, что при определении спектральных данных у больных до и после операции основным парамет-

*) См. работы В.Ф. Останина

ром является изменение энергетических характеристик спектра. Для оценки этих изменений используется:

- а) энергия формантных полос;
- б) энергия межформантных полос;
- в) отношение значений энергии формантных и межформантных

полос;

- г) ширина межформантных полос;
- д) средняя амплитуда формантных и межформантных полос.

Необходимо остановиться еще на одной области изучения речевого сигнала. Речь идет об изучении переходных процессов на материале конкретных языков. В качестве иллюстрации можно привести исследование на материале современного немецкого литературного языка. Основная задача исследования заключалась в определении тех изменений, которые гипотетически происходят на переходах от одного звука к другому в сочетаниях различного типа. Спектральный анализ составляющих слога позволил определить те изменения, которые претерпевает формантная структура гласных в зависимости от типа стыка.

В ходе исследования, проведенного автором, были отмечены спектральные отличия, свойственные гласным в открытой и закрытой позиции типа СТ, С#Г^{*)}. Так, например, было установлено, что формантная структура гласных в закрытой позиции значительно отличается от формантной структуры соответствующих гласных в открытой позиции, где вследствие наличия стыка между согласным и гласным происходит ослабление коартикуляции звуков. Фор-

*) Знак /#/ обозначает наличие стыка между звуковыми сегментами.

мантная структура гласных в открытой позиции максимально близка к формантной структуре тех же гласных в абсолютном начале или в изолированной позиции. При наличии примыкающего непосредственно к гласному согласного формантная структура гласных претерпевает существенные изменения. Причем наиболее индикативным параметром в данном случае является вторая форманта - F_2 .

Большое значение имеют данные спектрального анализа при различных опытах по восприятию речи. Так, например, следует упомянуть опыты по взаимной маскировке формант в спектрах стационарных синтетических гласных, проведенные под руководством Л.А.Чистович.

Одноформантный гласный с частотой форманты ниже 700 гц воспринимается как /у/ или /о/. Подавление верхних составляющих в спектрах естественных /и, ы, э, а/ превращает эти гласные в /у/ или /о/. Отсюда следовало ожидать, что синтетический двухформантный гласный с фиксированной по уровню первой формантой, расположенной в области 700 гц, будет восприниматься как /у/ или /о/ только при том условии, если уровень интенсивности второй форманты не превышает некоторого порога. При большом уровне второй форманты гласный будет приниматься за /а, э, ы/ или /и/.

Оказалось, что каждый из полученных таким образом "пороговых двухформантных" спектров отличается от соответствующего "одноформантного" только тем, что в нем увеличен уровень одной гармоники, той, на которой устанавливалась частота F_2 .

Данные свидетельствуют в пользу формантной гипотезы восприятия гласных и плохо совместимы с представлениями о том, что фонемные решения основываются на сравнении энергий в некоторых фиксированных полосах или на сравнении спектральной огибающей

входного сигнала с набором эталонных огибающих.

Приведем еще пример. Для исследования восприятия звонких смычных были использованы стимулы-синтезированные /b/, /d/, /g/ на базе рисованных спектрограмм. Если звонкие смычные соединялись с гласным на моделях, то большинство аудиторов идентифицировали их как слог, в котором за /æ/ следовали звонкие смычные. Если переход F_2 -повышающийся, то стимул воспринят как /b/. Если переход F_2 был прямым или слегка понижающимся, то согласный воспринимался как /d/. Если переход F_2 падал от высокой точки на частотной шкале, согласный воспринимался как /g/. Многие исследователи определяли F_2 как признак, достаточный для различения звонких смычных. Исследование показало, что F_3 так же важна для различения звонких смычных.

Данные позволяют утверждать, что отдельные акустические явления (переход F_2 и F_3) при восприятии относительно независи-мы.

Лекция 7. ПРИНЦИПЫ ПОДБОРА ЭКСПЕРИМЕНТАЛЬНОГО МАТЕРИАЛА

В настоящее время имеется несколько принципов подбора экспериментального материала. Однако каждый из этих принципов имеет своей целью анализ отдельных сторон исследуемого объекта и поэтому не дает полного представления об объеме всего экспериментального материала, определяемого морфофонологической системой конкретного языка.

Поскольку речевой континуум может быть последовательно разбит на отдельные единицы, образующие условно иерархический ряд, то одним из возможных подходов к изучению звукового строя языка является исследование, построенное на переходе от более крупных к более мелким единицам звукового потока

Остановимся на основных принципах составления экспериментального материала с использованием матричной системы, так как этот способ представляется наиболее строгим в экспериментальном отношении и наиболее экономичным.

Допустим, что исследователю необходимо получить все возможные для данного языка сочетания типа СГС. Для этого составляется матрица, где компонентами векторов* выступают все согласные исследуемого языка. Число компонентов вертикального вектора (вектора А) или иными словами, длина вектора А будет равняться длине вектора В (горизонтального вектора), т.е. $A=B$. (См. матрицу I).

* Вектор — составной объект, имеющий определенное число компонентов или составляющих. Обычно его записывают в виде $(a_1, a_2, \dots, a_n)$; это означает, что первая составляющая имеет значение a_1 , вторая — значение a_2 и т.д.

Вектор A

Вектор B

K	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	21	22	23
1.	p	t	k	f	s	r	h	x	g	pf	ts	tf	b	d	g	v	z	j	r	l	m	n	q
p																							
t																							
k																							
f																							
s																							
r																							
h																							
x																							
g																							
pf																							
ts																							
tf																							
b																							
d																							
g																							
v																							
z																							
j																							
r																							
l																							
m																							
n																							
q																							

МАТРИЦА 1. ВЕКТОРЫ - СОГЛАСНЫЕ НЕМЕЦКОГО ЯЗЫКА

Если исследователю необходимо получить экспериментальный материал, включающий сочетания типа СГ, то, в большинстве случаев, длина вектора А не будет равна длине вектора В, т.е. число компонентов вектора А (число согласных фонем исследуемого языка) не будет равно числу компонентов вектора В. В таком случае мы говорим, что $A \neq B$.

Итак, для составления матрицы сочетаний типа СГС* на материале, например, немецкого языка, мы получаем следующий исходный материал:

$$A = \langle a_1, a_2, a_3 \dots a_n \rangle$$

где a_1^{**} -	p	a_9 -	ç	a_{17} -	z
a_2 -	t	a_{10} -	pʃ	a_{18} -	j
a_3 -	k	a_{11} -	ts	a_{19} -	ʀ
a_4 -	f	a_{12} -	tʃ	a_{20} -	l
a_5 -	s	a_{13} -	b	a_{21} -	m
a_6 -	ʃ	a_{14} -	d	a_{22} -	n
a_7 -	h	a_{15} -	g	a_{23} -	ŋ
a_8 -	x	a_{16} -	v		

$$B = \langle a_1, a_2 \dots a_{23} \rangle$$

$$\langle a_1, a_2 \dots a_{23} \rangle = \langle a_1, a_2 \dots a_{23} \rangle$$

$$\bullet \quad A = B$$

* Естественно, что для составления экспериментального материала с применением матричного метода могут быть взяты сочетания не только типа СГС, но СГ, ГС, ССГ, СГГ и т.д.

** При обозначении компонентов вектора используются знаки международной фонетической транскрипции.

Далее, принимая компоненты вектора А за начало сочетания, а компоненты вектора В за исход сочетания вводим постоянный компонент – исследуемый гласный, который будет выступать во всех возможных сочетаниях между двумя согласными (исследуется сочетание типа СГС).

Таким образом, исследователь получает полный перечень сочетаний типа СГС, возможных для данной звуковой системы. Однако, как правило, в языке используется только определенная часть из всех возможных сочетаний. Все невозможные для данного языка сочетания исключаются и обозначаются на матрице знаком минус (-). Все возможные сочетания обозначаются знаком плюс (+).

Например, для немецкого языка исключаются комбинации:

- 1) с конечным звонким согласным;
- 2) ξ и χ в начальной позиции;
- 3) ξ после гласных среднего и заднего ряда;
- 4) с h в конечной позиции;
- 5) с ξ в конечной позиции (после гласных среднего и заднего ряда) ;
- 6) с s в начальной позиции ;
- 7) с j в конечной позиции ;

Далее из полученного материала по порядку исключаются все сочетания, не являющиеся однокорневыми лексемами. Оставшиеся однокорневые лексемы отмечаются любым символом, например, X(см. матрицу 2).

Предположим, что исследователю необходимо получить все единицы, значимые для данного языка (лексемы, морфемы). В таком случае объем материала увеличится за счет прибавления

H.	K.	P	T	K	F	S	I	A	X	S	R	T	S	L	B	D	G	V	Z	J	C	L	M	N	O
P	X	X	X	X	X		X															X			
K	X	X	X	X	X																	X			
F	X	X	X	X	X																	X			
S	X	X	X	X	X																				
I	X	X	X	X	X																				
A	X	X	X	X	X																				
X	X	X	X	X	X																				
E	X	X	X	X	X																				
X	X	X	X	X	X																				
E	X	X	X	X	X																				
X	X	X	X	X	X																				
E	X	X	X	X	X																				
X	X	X	X	X	X																				
E	X	X	X	X	X																				
X	X	X	X	X	X																				
E	X	X	X	X	X																				
X	X	X	X	X	X																				
E	X	X	X	X	X																				
X	X	X	X	X	X																				
E	X	X	X	X	X																				
X	X	X	X	X	X																				
E	X	X	X	X	X																				
X	X	X	X	X	X																				
E	X	X	X	X	X																				
X	X	X	X	X	X																				
E	X	X	X	X	X																				
X	X	X	X	X	X																				
E	X	X	X	X	X																				
X	X	X	X	X	X																				
E	X	X	X	X	X																				
X	X	X	X	X	X																				
E	X	X	X	X	X																				
X	X	X	X	X	X																				
E	X	X	X	X	X																				
X	X	X	X	X	X																				
E	X	X	X	X	X																				
X	X	X	X	X	X																				
E	X	X	X	X	X																				
X	X	X	X	X	X																				
E	X	X	X	X	X																				
X	X	X	X	X	X																				
E	X	X	X	X	X																				
X	X	X	X	X	X																				
E	X	X	X	X	X																				
X	X	X	X	X	X																				
E	X	X	X	X	X																				
X	X	X	X	X	X																				
E	X	X	X	X	X																				
X	X	X	X	X	X																				
E	X	X	X	X	X																				
X	X	X	X	X	X																				
E	X	X	X	X	X																				
X	X	X	X	X	X																				
E	X	X	X	X	X																				
X	X	X	X	X	X																				
E	X	X	X	X	X																				
X	X	X	X	X	X																				
E	X	X	X	X	X																				
X	X	X	X	X	X																				
E	X	X	X	X	X																				
X	X	X	X	X	X																				
E	X	X	X	X	X																				
X	X	X	X	X	X																				
E	X	X	X	X	X																				
X	X	X	X	X	X																				
E	X	X	X	X	X																				
X	X	X	X	X	X																				
E	X	X	X	X	X																				
X	X	X	X	X	X																				
E	X	X	X	X	X																				
X	X	X	X	X	X																				
E	X	X	X	X	X																				
X	X	X	X	X	X																				
E	X	X	X	X	X																				
X	X	X	X	X	X																				
E	X	X	X	X	X																				
X	X	X	X	X	X																				
E	X	X	X	X	X																				
X	X	X	X	X	X																				
E	X	X	X	X	X																				
X	X	X	X	X	X																				
E	X	X	X	X	X																				
X	X	X	X	X	X																				
E	X	X	X	X	X																				
X	X	X	X	X	X																				
E	X	X	X	X	X																				
X	X	X	X	X	X																				
E	X	X	X	X	X																				
X	X	X	X	X	X																				
E	X	X	X	X	X																				
X	X	X	X	X	X																				
E	X	X	X	X	X																				
X	X	X	X	X	X																				
E	X	X	X	X	X																				
X	X	X	X	X	X																				
E	X	X	X	X	X																				
X	X	X	X	X	X																				
E	X	X	X	X	X																				
X	X	X	X	X	X																				
E	X	X	X	X	X																				
X	X	X	X	X	X																				
E	X	X	X	X	X																				
X	X	X	X	X	X																				
E	X	X	X	X	X																				
X	X	X	X	X	X																				
E	X	X	X	X	X																				
X	X	X	X	X	X																				
E	X	X	X	X	X																				
X	X	X	X	X	X																				
E	X	X	X	X	X																				
X	X	X	X	X	X																				
E	X	X	X	X	X	</																			

МАТРИЦА II. ОДНОКОРНЕВНЕ ЛЕКСЕМЫ НЕМЕЦКОГО ЯЗЫКА С КОРНЕВЫМ
ГЛАСНЫМ /a/

сочетаний, являющихся морфемами исследуемого языка. Обозначим эти сочетания символом (X) (См. матрицу 3).

Если цель исследования заключается в анализе сочетания типов СГС в минимальном объеме, т.е. при том условии, что согласные в начальной и исходной позициях будут выступать только по одному разу, то в данном случае в качестве экспериментального материала берутся сочетания по диагонали матрицы. Естественно, что полученный таким образом экспериментальный материал будет включать наряду с сочетаниями – лексемами, сочетаниями-морфемами, искусственно составленные сочетания, возможные для данного языка. (См. матрицу 4).

В том случае, если предметом исследования являются все гласные, то на каждую гласную при наличии тех же согласных составляется отдельная матрица. В целях экономии можно разбить гласные по группам, например, гласные переднего ряда, среднего ряда, заднего ряда. В таком случае матрица составляется с переменной величиной, включающей сразу одну из названных групп гласных.

Но здесь возможно некоторое усложнение самого процесса отбора материала, поэтому все же рациональнее для каждой гласной фонемы составлять отдельную матрицу.

При строгом отборе минимального материала возможно еще большее сужение круга исследуемых единиц. Например, требуется, чтобы в экспериментальном материале были:

- а) сочетания-лексемы, сочетания-морфемы;
- б) сочетания с неповторяющимися гласными;
- в) сочетания с неповторяющимися согласными и т.д.

к.	р	т	к	ф	с	ш	х	з	ц	б	д	г	в	з	ч	ш	н	г
р	х	х	х	х	х	х	х	х	х	х	х	х	х	х	х	х	х	х
т	х	х	х	х	х	х	х	х	х	х	х	х	х	х	х	х	х	х
к	х	х	х	х	х	х	х	х	х	х	х	х	х	х	х	х	х	х
ф	х	х	х	х	х	х	х	х	х	х	х	х	х	х	х	х	х	х
с	х	х	х	х	х	х	х	х	х	х	х	х	х	х	х	х	х	х
ш	х	х	х	х	х	х	х	х	х	х	х	х	х	х	х	х	х	х
х	х	х	х	х	х	х	х	х	х	х	х	х	х	х	х	х	х	х
з	х	х	х	х	х	х	х	х	х	х	х	х	х	х	х	х	х	х
ц	х	х	х	х	х	х	х	х	х	х	х	х	х	х	х	х	х	х
б	х	х	х	х	х	х	х	х	х	х	х	х	х	х	х	х	х	х
д	х	х	х	х	х	х	х	х	х	х	х	х	х	х	х	х	х	х
г	х	х	х	х	х	х	х	х	х	х	х	х	х	х	х	х	х	х
в	х	х	х	х	х	х	х	х	х	х	х	х	х	х	х	х	х	х
з	х	х	х	х	х	х	х	х	х	х	х	х	х	х	х	х	х	х
ч	х	х	х	х	х	х	х	х	х	х	х	х	х	х	х	х	х	х
ш	х	х	х	х	х	х	х	х	х	х	х	х	х	х	х	х	х	х
н	х	х	х	х	х	х	х	х	х	х	х	х	х	х	х	х	х	х
г	х	х	х	х	х	х	х	х	х	х	х	х	х	х	х	х	х	х

Таблица III. ОДНОКОРНЕВЫЕ ЛЕКСЕМЫ И МОРФЕМЫ НЕМЕЦКОГО ЯЗЫКА
С КОРНЕВЫМ ГЛАСНЫМ /а//

В таком случае из полученного материала по диагонали составленной матрицы исследователь отбирает нужные сочетания. Например, на материале немецкого языка для ГГС с кратким гласным /a/ и при различном фонетическом контексте мы получаем следующий ряд примеров:

patasch, Tanaz, kaff, Fail, Schach, lat, Pfarr(er), Zan(der),
 Ball, Damm, Gang, Papp, satt, Jatz, rasch, Lack, Mann,
 nass, wasch, tapf(er), Satz

Возможны случаи, когда в экспериментальный материал следует включить не только одиночные согласные в начальной и исходной позициях, а две, три и более согласных фонемы, что зависит от фонологической системы данного языка. Тогда на основании составленной матрицы к каждому из сочетаний в начальной и исходной позициях прибавляются все консонантные группы, возможные для данного языка. Таким образом, мы получаем все возможные вышеперечисленные варианты экспериментального материала, но уже с консонантным наращением в начале и исходе сочетаний.

Например, материал, полученный для немецкого языка в сочетаниях - ГССС - применительно только к конечной позиции:

*herpst, ʔernst, hilfst, hegsts,
 herpstʰ, ʔo:psts, di:nsts,
 tælstʰ, fælfst, pa:psts и т. д.*

Таким образом, на основании использования матричного метода мы имеем возможность получить такой экспериментальный материал, который бы гарантировал:

✓ а) всесторонний охват согласных и гласных фонем исследуемого языка;

б) возможность получения нескольких вариантов экспериментального материала, что зависит от задачи исследования.

В заключение перечислим все возможные варианты экспериментального материала, получаемого в результате применения матричного метода:

1. Вариант-максимум, включающий все возможные комбинации исследуемых гласных и согласных, сюда входят как значимые единицы, так и искусственно составленные звуко сочетания.

2. Вариант, включающий: а) лексемы, б) морфемы, в) отдельные сочетания многосложных слов.

3. Вариант включающий: а) лексемы; б) морфемы.

4. Вариант включающий: а) сочетания – однокорневые лексемы.

5. Вариант-минимум включающий: а) значимые единицы,

б) искусственно составленные сочетания при условии только одного употребления начального и исходного звука.

6. Вариант, включающий: а) сочетания с одним согласным (или гласным) в начальной и исходной позициях;

б) сочетания с группой согласных (или гласных) в начальной и исходной позициях.

Таким образом, матричный метод составления экспериментального материала дает исследователю возможность провести строгий учет всего экспериментального материала, необходимого для иссле-

дования звуковой системы данного языка. С помощью матричного метода можно установить частоту встречаемости той или иной фонемы (группы фонем) в той или иной позиции. Например, применение матричного способа позволило нам определить ранговое распределение глухих согласных в немецком языке для начальной и конечной позиции в односложных морфемах:

а) распределение глухих согласных в начальной позиции:

<i>k</i>	<i>h</i>	<i>t</i>	<i>p</i>	<i>ʃ</i>	<i>ts</i>	<i>ʀʃ</i>	<i>x</i>	<i>ʒ</i>	<i>s</i>	<i>ʈ</i>
30%	26%	23%	22%	19%	18%	8%	6%	0	0	0

б) распределение глухих согласных в конечной позиции:

<i>s</i>	<i>t</i>	<i>p</i>	<i>f</i>	<i>ʒ</i>	<i>k</i>	<i>x</i>	<i>ts</i>	<i>ʃ</i>	<i>ʈ</i>	<i>pʃ</i>	<i>h</i>
36%	35%	33%	29%	25%	20%	16%	14%	12%	6%	5%	0

Отвлекаясь от чисто утилитарного значения матричного метода, можно утверждать, что предлагаемые типы матриц применимы и в целях формального описания морфологического уровня исследуемого языка. При данном подходе мы определяем единицы более высокого уровня в иерархической системе языка через единицы низшего уровня чисто индуктивным, аналитическим путем в отличие, скажем от прямо противоположного метода Харриса (т. наз. метод " — *n* + I"), состоящего в подсчете числа всех возможных фонем, способных появиться после *n* фонемы в высказывании. При этом автор метода " *n* + I" ставит своей задачей чисто формальным путем, не принимая во внимание смысловую сторону, расчленить цепь звуков, представляющую отдельное высказывание, на сегменты, которые соответствуют морфемам и словам.

Наряду с тем, что метод " $n + 1$ " далек от строгой систематизации получаемого исследователем материала, он несколько неудобен еще и в том отношении, что все данные исследователь получает от информантов-носителей языка.

Таким образом, цель установления взаимодействия двух рассматриваемых уровней (фонологического и морфологического) достигается как при применении матричного метода, так и при применении метода " $n + 1$ ". Однако, как признает Харрис, метод " $n + 1$ " не имеет практического значения, оставляя за собой право чисто формальным путем вскрыть структурное своеобразие того или иного конкретного языка на уровне морфологического анализа.

В противовес этому методу матричный метод может быть использован в тройном плане: 1) для составления экспериментального материала; 2) для подсчета частотности согласных и гласных фонем в дистрибутивном аспекте; 3) для получения порождающей модели при переходе от фонологического к морфологическому уровню. Что касается последнего пункта, то здесь каждая матрица может быть представлена в качестве конечного (для всякого конкретного языка) множества комбинаций (сочетаний согласных и гласных фонем).

Так, можно представить набор согласных и гласных фонем в каждой матрице как некое устройство (систему правил), способное порождать морфемы и лексемы, принцип построения которых будет отвечать нормам данного языка.

В качестве критериев оценки данной порождающей модели предлагаются следующие:

1) полнота объема (полученные модели должны охватывать все возможные и потенциально возможные варианты сочетаний);

2) систематизация по ступеням (порождение только морфем, только лексем, тех и других вместе. Естественно, что, используя сочетания типа СГС, мы будем получать только моносиллабические единицы (слова, морфемы). Если же изменить сочетания, т.е. изменить правила матричного устройства, то на выходе мы будем получать иные комплексы, также отвечающие морфологическим нормам данного языка. Таким образом, модель вышеописанного типа можно представить как один из этапов так называемого марковского процесса с конечным числом состояний^{*}, т.е. как некоторого рода устройство, способное принимать конечное число состояний и при определенном переходе порождать искомую единицу. Поскольку инвентарь гласных и согласных фонем определенного языка является конечным числом состояний, количество образований тоже будет конечным.

Для определения числа звукосочетаний, получаемых в результате комбинаций гласных и согласных в пределах каждой матрицы, мы можем предположить, что всем реализациям звукосочетания (например СГС) будет соответствовать какое-то множество A комбинаций гласных и согласных. Если обозначить все реализации

* Процесс называется марковским, если для каждого момента времени вероятность любого состояния системы в будущем зависит только от состояния системы в настоящий момент и не зависит от того, каким образом система пришла в это состояние.

их комбинаций через A_1, A_2, A_3 , то сумма всех возможных реализаций, полученных от n -реализаций гласных со всеми согласными (n равно числу гласных исследуемого языка) будет равна:

$$= \bigcup^n A_k$$

В данном случае берется не сумма $\sum^n A_k = A_1 + A_2 + \dots + A_k$, а используется теоретико-множественная сумма $\bigcup^n A_k = A_1 \cup A_2 \cup \dots \cup A_n$, где $\cup$ - символ теоретико-множественного объединения, A - символ множества,*
 n - символ числа реализаций,
 k - символ перечисления от 1 до n

Использование теоретико-множественной суммы продиктовано в свою очередь существом исследуемого явления, так как всякая комбинация гласных и согласных рассматривается нами как изменение множества A в зависимости от фонологических критериев заданного конкретного языка.

Далее этот анализ можно продолжить в плане определения для каждой конкретной матрицы (т.е. для определения множества комбинаций, например, всех согласных языка с одним гласным) релевантных, иррелевантных и редундантных составляющих в плане реализации полученных сочетаний в речи.

I. Если одна или несколько составляющих примыкают к релевантной составляющей (например, гласной) относительно свободно, т.е. могут сопровождать ее или нет, то эти составляющие обоз-

Понятие "множества" является одним из главных понятий в математике. Под множеством понимают обычно совокупность, составленную из каких-либо объектов, называемых элементами множества.

начаются как иррелевантные.

Отношение релевантного составляющего А к иррелевантному составляющему (В) может быть выражено следующим образом:

$$\left\{ \begin{array}{c} A \\ (B) \end{array} \right\}$$

Например, вокализованный вибрант в конечной позиции на материале немецкого языка.

II. Если какая-либо составляющая является частью всего комплекса, а ее выпадение не изменяет всего комплекса, то такая составляющая может быть определена в данном случае как редундантная. Отношение релевантной составляющей А к избыточной составляющей <С> выражается следующим образом:

$\left\{ \begin{array}{c} A \\ <C> \end{array} \right\}$ Например, редуцированная гласная /ə/ в немецком языке.

III. Если каждая из составляющих комплекса не может быть заменена другой или вообще отброшена, то в данном случае можно утверждать, что все составляющие релевантны:

$\left\{ \begin{array}{c} A \\ B \end{array} \right\}$

Имеется в виду большинство существующих в языке комбинаций гласных и согласных, обеспечивающих правильную смысловую идентификацию высказывания.

На основании вышеописанных принципов и на материале заранее составленных матриц можно определить фонологические возможности при переходе от фонологического уровня к морфологическому, характеризующие конкретную фонологическую систему рассматриваемого языка. Данный анализ может быть использован и в типологическом плане, в результате чего мы получим матрицу, построенную уже на основе функциональной зависимости морфологического уровня от фонологического. Этот анализ можно провести и далее, используя построение матриц на основе разбиения материала, полученного на I этапе, с учетом позиции звуков в слове, принадлежностью слов, составленных из анализируемых гласных, к

различным грамматическим категориям и т.д.

Таким образом, исходя из того, что фонетическое описание должно быть селективным, т.е. мы не можем описать речительно всего, а должны выбрать элементы описания и определить степень их необходимости для эксперимента, в основу матричного метода составления экспериментального материала был положен принцип релевантности. Таким образом, каждая матрица определяет конечное для каждого конкретного языка множество комбинаций (сочетаний). Полученные с помощью матриц комбинации должны удовлетворять требованиям:

- а) по полноте объема;
- б) по систематизации материала;
- в) по наличию перехода от одного языкового уровня к другому.

С помощью матричного метода возможно определение и отбор сегментов:

1. Актуальных, т.е. функционирующих в языке и соответствующих фонотактическим моделям языка;
 2. Маргинальных, т.е. функционирующих, но не соответствующих фонотактическим моделям данного языка;
 3. виртуальных, т.е. не функционирующих на данном этапе развития языка, но в перспективе возможных по фонотактическим правилам языка;
 4. недопустимых, т.е. не функционирующих и невозможных.
- экспериментальный материал, подобранный при помощи матричного метода, может служить как основной

выборкой * в исследовании, так и контрольной, рассматриваемой параллельно с экспериментальным материалом, источником которого являются словари, тексты и т.д.

Примечание:

В лекциях вышеизложенного спецкурса использованы материалы большого числа работ отечественных и зарубежных авторов, изложены результаты собственных многолетних исследований, проведенных на материале различных языков. Обилие материала не снижает необходимости указания на основные работы, в которых освещены в той или иной степени вышеизложенные проблемы.

Для того, чтобы составить представление о распределении той или иной величины и ее важнейших характеристиках, можно обследовать некоторую **выборку**, а не обращаться ко всей генеральной совокупности в целом. В математике принято постулировать, что число составляющих генеральной совокупности велико, а число составляющих выборки ограничено. Поэтому изучается лишь часть материала, которую принято называть **выборкой из генеральной совокупности**.

Основная литература:

1. Артемов В.А. Экспериментальная фонетика, М., 1956.
2. Блохина Л.П., Потапова Р.К. К вопросу о методике исследования фонемного состава языка. Сб. "Материалы коллоквиума по экспериментальной фонетике и психологии речи", М., 1966.
3. Бондарко Л.В. Слоговая структура речи и дифференциальные признаки фонем, Дд., Л., 1969.
4. Вентцель Е.С. Теория вероятностей. М., 1964.
5. Златоустова Л.В. Фонетические единицы русской речи. Дд., М., 1970.
6. Потапова Р.К. Основные современные способы анализа и синтеза речи, М., 1971.
7. Прокопова Л.И. Структура слога в немецком языке. Киев, 1973.
8. Сапожков М.А. Речевой сигнал в кибернетике и связи. М., 1963.
9. Фант Г. Акустическая теория речеобразования. М., 1964(перевод).
10. Фант Г. Анализ и синтез речи. Новосибирск, 1970 (перевод).
11. Фланган Д. Анализ, синтез и восприятие речи. М., 1968 (перевод).
12. Чистович Л.А. и др. Речь, артикуляция и восприятие. М.-Л., 1965.