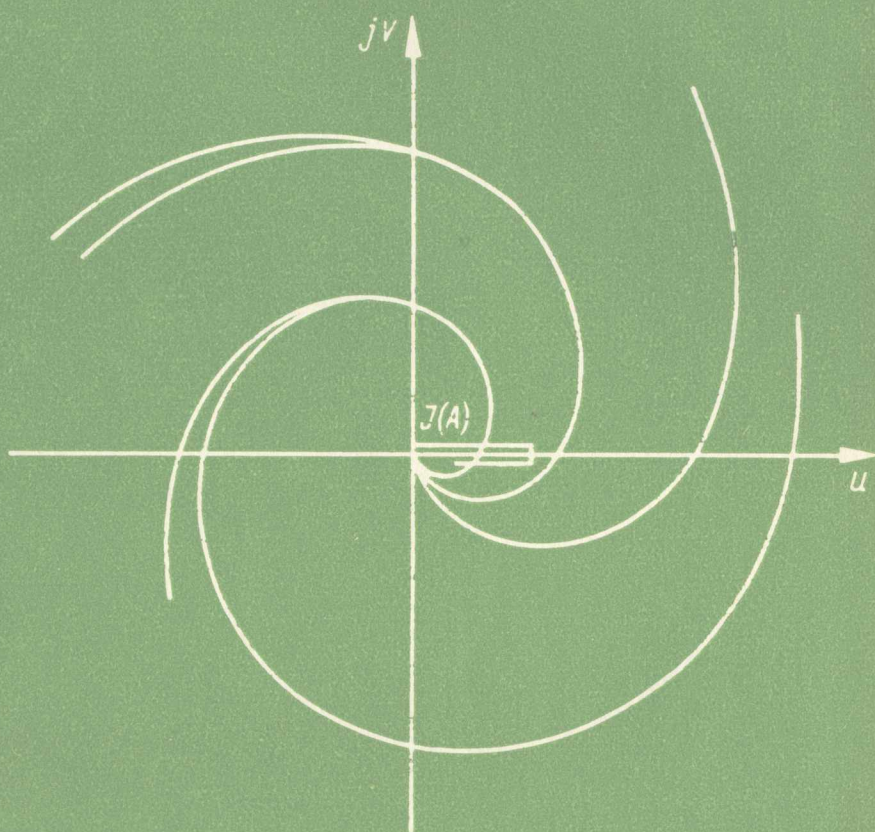
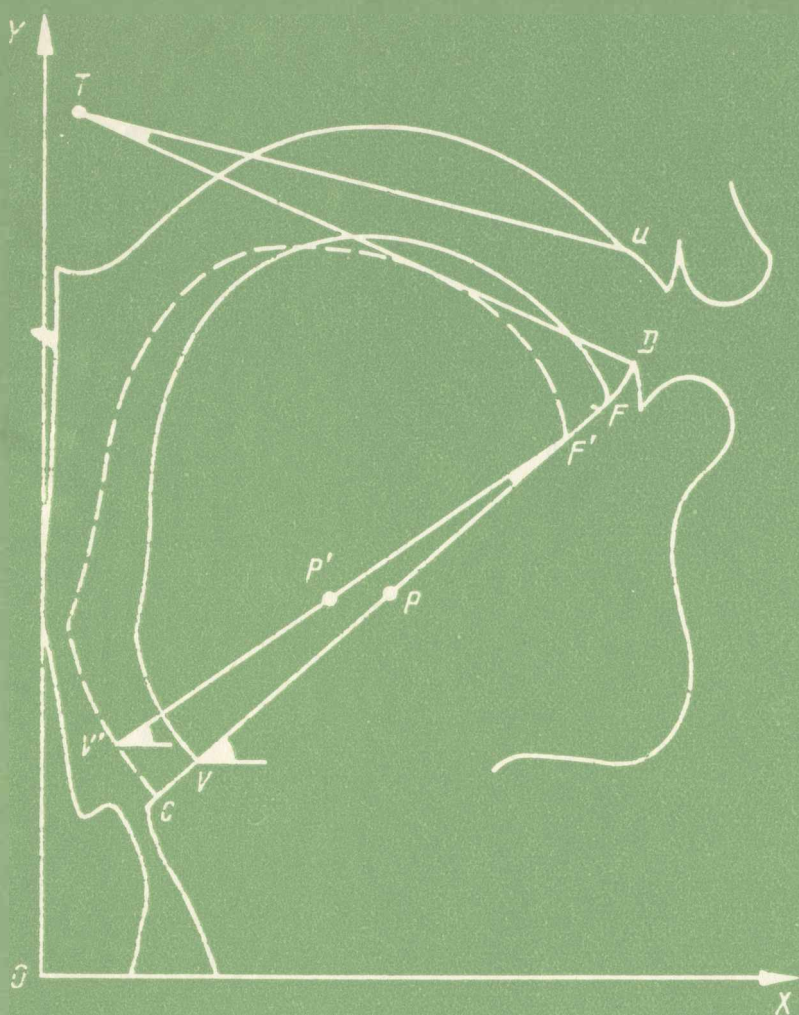


4030

Теория _____ В. Н. Сорокин
РЕЧЕ-
ОБРАЗОВАНИЯ





В. Н. Сорокин

Теория РЕЧЕ- ОБРАЗОВАНИЯ



Москва
«Радио и связь»
1985

В. Н. Сорокин. Теория речеобразования.— М.: Радио и связь, 1985.— 312 с., ил.

В монографии последовательно рассматривается влияние различных уровней процесса речеобразования на характеристики речевого сигнала. На основе экспериментальных и теоретических исследований разработаны количественные модели движений артикуляторных органов и их управления, колебаний голосовых складок и акустических колебаний в речевом тракте, причем основное внимание уделяется динамике процессов и параметрическим явлениям. Показано, что в основе многих процессов лежат волновые механизмы, что позволяет использовать для их описания единый математический аппарат. В монографии излагаются наиболее важные экспериментальные результаты, полученные как автором, так и зарубежными исследователями с применением кинорентгенографии, сверхскоростной съемки и электромиографии. Описываются существенные физические явления, влияющие на характер процессов речеобразования.

Монография рассчитана на научных работников, занимающихся вопросами речеобразования, автоматическим распознаванием, синтезом и передачей речи, и будет полезна специалистам смежных областей науки и техники (кибернетикам, акустикам, логопедам, лингвистам).

Табл. 16. Ил. 129. Библиогр. 400 назв.

Рецензенты: М. А. САПОЖКОВ, В. С. ГУРФИНКЕЛЬ

Редакция литературы по электросвязи

ВИКТОР НИКОЛАЕВИЧ СОРОКИН

ТЕОРИЯ РЕЧЕОБРАЗОВАНИЯ

Редактор В. К. Старикова

Переплет художника М. В. Бакурина

Художественный редактор Р. А. Клочков

Технический редактор И. Л. Ткаченко

Корректор Л. С. Глаголева

ИБ № 954

Сдан в набор 07.08.84

Подписано в печать 15.10.84

Т-21106

Формат 60×90/16

Бумага тип. № 2

Гарнитура литературная

Печать высокая Усл. печ. л. 19,5

Усл. кр.-отт. 19,5 Уч.-изд. л. 21,99

Тираж 2900 экз.

Изд. № 20766

Зак. № 80

Цена 3 р. 60 к.

Издательство «Радио и связь». 101000 Москва, Почтамт, а/я 693

Московская типография № 5 ВГО «Союзучетиздат»

101000 Москва, ул. Кирова, д. 40

С 2402030000-004
046(01)-85 76-85

© Издательство «Радио и связь», 1985

Теория речеобразования имеет свой предмет исследования, однако она в значительной степени носит междисциплинарный характер, используя методы и сведения из физиологии высшей нервной деятельности, нейрофизиологии управления движениями, анатомии, слухового восприятия, механики, аэродинамики, акустики, теории автоматического управления, теории упругости и теории информации. Это создает субъективные трудности в развитии речевых исследований из-за необходимости усвоения большого объема разнородной информации. В больших коллективах, состоящих из специалистов разных областей, это приводит к сложностям взаимопонимания и координации научной деятельности. В малых же коллективах возникает опасность дилетантства, недостаточно глубокого знания некоторых разделов смежных дисциплин, что может привести к недопустимым упрощениям или ошибочным оценкам. Все это осложняется различием в терминологии и методологических подходах в смежных областях. Поэтому назрела необходимость в монографии, в которой были бы систематизированы результаты экспериментальных и теоретических исследований на всех уровнях процесса речеобразования.

Автор придерживался терминов и обозначений, принятых в соответствующих смежных областях, методы которых использовались в описываемых исследованиях. Поэтому некоторые обозначения в разных разделах могут иметь разный смысл, о чем специально оговаривается.

При написании книги были использованы результаты экспериментальных исследований, проведенных автором в различных научных коллективах. Моделирование влияния подсвязочной области на характеристики речевого сигнала было выполнено в Лаборатории передачи речи Королевской технологической школы (Стокгольм) в сотрудничестве с Г. Фантом. Кинорентгенограммы были сняты в поликлинике г. Линчепинга (Швеция) при помощи Г. Эриксона и Д. Стинпорта. Измерения скорости распространения возбуждения в подбородной мышце были проведены в поликлинике г. Упсала (Швеция) вместе с Э. Столбергом. Электромиография лицевых мышц была проведена в Стокгольмском университете с участием Д. Лабкера, Я. Карлсона и Б. Фритцеля. Эксперименты по стимуляции жевательной мышцы и измерению амплитудно-частотных характеристик нижней челюсти были выполнены в Коннектикутском университете (США) вместе с Т. Геом и В. Эваном, и там же были сделаны записи нейромоторных потенциалов некоторых лицевых мышц синхронно с движениями губ и нижней челюсти. Измерения движений небной занавески и языка проводились на микролучевом рентгенооскопе Токийского

университета с участием Х. Хирозе, Ш. Киритани и других сотрудников Института логопедии и фониатрии. Кинорентгенограммы записей для эстонского языка предоставил А. Ээк (Институт кибернетики АН ЭССР).

Автор почерпнул много полезных сведений из бесед с А. Либманом, К. Стивенсом, Д. Перкеллом, Д. Клаттом, П. Ладефогедом, П. Макниладжем, М. Хирано и Х. Фуджисаки. Вопросы исследования и моделирования процессов речеобразования неоднократно обсуждались с В. С. Гурфинкелем, Р. С. Персон, Е. И. Пальцевым, А. Г. Фельдманом и Н. Д. Введенской.

Большую помощь в разработке программного обеспечения для моделирования на ЭВМ оказали сотрудники лаборатории № 3 Института проблем передачи информации АН СССР В. А. Суханов, А. П. Вайншток, В. А. Потапов, Е. И. Кабанова, а также Е. В. Власов и Е. Е. Белицкая.

Автор благодарен В. С. Гурфинкелю, Н. А. Дубровскому, М. А. Сапожкову и В. Н. Телепневу за ценные замечания по содержанию монографии.

Все замечания по книге просьба направлять в издательство «Радио и связь» по адресу: 101000, Москва, Почтамт, а/я 693.

Автор

Речевые исследования имеют долгую историю, включая рассуждения Платона (V век до н. э.) о происхождении языка в его книге «Кратилл». Уже в древности эти исследования носили не только отвлеченный характер, о чем свидетельствует интерес к речи таких выдающихся медиков, как Гиппократ, Гален, Авицена. Особое внимание уделялось психологическому воздействию речи, которое, как становится ясным в наши дни, имеет глубокую физиологическую основу — например, известно, что состав крови и биохимия организма человека, находящегося в гипнотическом состоянии, могут быть изменены с помощью слова. Начиная с XVIII века неоднократно предпринимаются попытки механического синтеза речи. Первый синтезатор был построен Кемпеленом, который в 1791 г. написал о нем книгу [B.8]. С тех пор усилия по созданию синтезаторов речи не прекращаются.

По мере развития физики объективные методы измерений применяются и к речи, и вскоре появляются теории образования и восприятия гласных звуков (Виллис, 1829, и Гельмгольц, 1875).

Осознание свойств речевого сигнала как изменения давления в звуковой волне привело к одному из наиболее революционных изобретений — телефону Г. Белла. Гельмгольц обнаружил максимум энергии в частотном спектре гласных и назвал его «формантом». Расширение возможностей измерительной техники помогло сначала найти второй максимум энергии (Крендэлл, 1927), а затем получить и другие максимумы, соответствующие резонансам речевого тракта. Изобретение динамического спектрографа дало возможность наблюдать за изменением во времени частотных характеристик речевого сигнала в форме так называемой «видимой речи» [B.14]. С 30-х годов к исследованию процессов речеобразования привлекается рентгенография, а позднее — электромиография мышц. Изобретение полосных вокодеров позволило во много раз сжать требуемую полосу частот для передачи речи. Это стало возможным после того, как возникло представление о модуляционной природе речевого сигнала и было разрешено противоречие между медленными артикуляционными движениями и относительно широкой полосой речевого сигнала. Монографии [B1, B3, B4] внесли существенный вклад в исследование процессов речеобразования.

В настоящее время совершенствование экспериментальных методик исследования речи и разработка теории речеобразования, с одной стороны, и развитие электронно-вычислительной техники и микроэлектронной технологии — с другой чрезвычайно расширили область применения результатов речевых исследований. Разнообразные устройства уже в настоящее время создают положитель-

ный экономический эффект в различных областях. По-видимому, на определенном этапе технического развития общества применение автоматизированных речевых систем окажется жизненно необходимым и будет обеспечивать доступ к информационным ресурсам.

Помимо ставшего уже традиционным применения речевых исследований для передачи речи сформировались следующие направления: синтез, распознавание речи и медицинские приложения. Методы сжатия полосы для экономной передачи речевого сигнала позволяют во много раз увеличить использование пропускной способности каналов связи, стоимость которых и в ближайшем будущем будет составлять значительную долю стоимости систем связи. Кроме того, сжатие речевого сигнала создает возможность для обмена частотной полосы на мощность, что особенно важно для космической и дальней связи, а также для перекодирования речи применительно к конкретным каналам связи и видам помех. Параметрическое представление речевого сигнала обеспечивает засекречивание разговоров. В последнее время разрабатываются каналы связи, по которым одновременно передаются цифровые данные и речь, и эффективность этих каналов зависит от степени использования свойств речевых сигналов.

Синтезаторы речи находят применение в управлении действиями человека, когда по каким-либо причинам другие способы доставки информации не могут использоваться или закрыт зрительный канал. Это относится, например, к управлению сборкой сложной электронной аппаратуры, где перевод инструкций с алфавитно-цифрового печатающего устройства на речевое управление позволяет монтажникам повысить производительность труда, уменьшить утомляемость и число ошибок [В.10]. Созданы системы оповещения для операторов подводных лодок [В.9]. Синтезаторы речи используются в бытовой технике в виде говорящих калькуляторов, инструкций к пользованию, для объявлений остановок в поездах, номеров этажей в лифтах, для информации в аэропортах, музеях и крупных универсальных магазинах. Одно из важнейших применений синтезаторов заключается в протезировании — изготовлении говорящих устройств для людей с нарушениями речи и читающих машин для слепых. В первом случае используется клавишный набор сообщений, а применение читающих машин облегчается применением в полиграфии новых методов набора с одновременным созданием магнитных лент, на которых текст кодируется в цифровой форме. Только число людей, не способных пользоваться зрением для чтения, составляет, по данным для США, около 400 тыс. человек. Если принять, что и в остальных странах относительная доля таких инвалидов примерно та же, и добавить к ним немых и людей с тяжелыми расстройствами речи, то общее число нуждающихся в синтезаторах-протезах окажется близким к 10 млн. человек.

Системы автоматического распознавания речи применяются в промышленности для сортировки, инвентаризации, контроля ка-

чества деталей на конвейере, описании объектов на картах и аэрофотоснимках, геологических разрезах и т. д. При этом производительность труда повышается (в ряде случаев до 30—40%), а число ошибок уменьшается во много раз, улучшаются условия труда за счет снижения напряжения и утомляемости рабочих. Имеется опыт программирования ЭВМ голосом [В.6]. Для систем автоматического распознавания характерна быстрая окупаемость вложений — системы стоимостью около 50 000 дол. окупаются примерно за год. Другие применения автоматического распознавания состоят в вводе данных в ЭВМ, управления радио- и телеприемником, включении света, наборе номера телефона и т. д. Разрабатываются варианты автоматической пишущей машинки.

Наибольший эффект, однако, ожидается от одновременного использования системы распознавания и синтеза речи, что обеспечит двусторонний обмен информацией с удаленными источниками, например, в информационно-справочных службах. В настоящее время такие системы используются в США и Японии для банковских операций и заказа авиа- и железнодорожных билетов. Преимущества таких систем — круглосуточная готовность, исключение человека из промежуточных операций, ускорение движения информационных потоков. В будущем, с появлением систем искусственного интеллекта, для общения с ними целесообразно использовать естественный язык. Возможно применение систем такого класса для автоматизации определенной части процесса преподавания.

Автоматическая идентификация и верификация дикторов находят применение в криминалистике и системах индивидуального доступа [В.2].

Механика процессов артикуляции и управление артикуляционными движениями являются разделом физиологии движений человека, и их исследование вносит существенный вклад в понимание общих принципов управления движениями. Особый интерес вызывает связь речи и мышления. Результаты речевых исследований применяются в хирургии расщелин мягкого неба и в логопедии для обучения правильному произнесению звуков. Эти исследования открывают новое направление в диагностике некоторых заболеваний. Например, по акустическим характеристикам речевого сигнала удается идентифицировать такие виды патологии гортани, как полип, анемия, паралич и, самое главное, рак [В.7, В.11]. С тех пор как была показана возможность диагностики асфиксии, мозговых повреждений, гипербилирубина и синдрома Дауна по крику новорожденных с использованием спектрографических и других измерений [В.12, В.13, В.15], открылись перспективы лечения некоторых генетических болезней и уменьшения последствий родовых травм.

Специфика процессов речеобразования затрудняет их экспериментальное исследование. Подавляющее большинство артикуляционных и акустических процессов протекает внутри речевого тракта, а колебания голосовых складок имеют к тому же очень широ-

кий частотный спектр. Поэтому для измерения формы артикуляционных органов и голосовых складок применяются кинорентгенография и скоростная киносъемка. Поскольку доза рентгеновского облучения очень ограничена, рентгеновские исследования не могут длиться продолжительное время. Решение этой проблемы частично найдено на пути создания микролучевых рентгеноскопов, управляемых ЭВМ, но эти установки очень дороги, и пока во всем мире их существует только две. Доступ к мышцам, управляющим артикуляцией и фонацией, иногда очень затруднен, а сами мышцы зачастую малы и близко расположены, что усложняет электромиографию. Почти во всех случаях объем измерений настолько велик, что единственным способом усвоения этого потока информации является непосредственный его ввод во время эксперимента и последующая обработка на ЭВМ. Все эти сложности превращают постановку каждого эксперимента в трудоемкое и дорогостоящее предприятие и придают особую ценность каждому достоверному экспериментальному факту. Особую трудность создают ошибки измерений, которые не позволяют исследовать процессы при малых значениях каких-либо параметров, например площади прохода в носовую полость или минимальной площади поперечного сечения речевого тракта. Быстрые изменения параметров речевого тракта в процессе артикуляции создают принципиальные трудности в измерении акустических характеристик речевого сигнала. Вследствие этого в речевых исследованиях особо важное значение приобретают теоретический анализ и разработка математических моделей.

Речевой сигнал предназначен для передачи информации от одного человека к другому, поэтому система управления артикуляцией должна формировать такие параметры речевого сигнала, которые сохраняли бы всю необходимую информацию для понимания аудитором, несмотря на помехи в канале связи и индивидуальные различия в геометрии и механических параметрах речевых трактов разных дикторов. Такое предназначение речевого сигнала порождает существенные отличия в характеристиках системы управления артикуляцией по сравнению с системами управления произвольными движениями человека, причем механика и акустика речеобразования, свойства периферических процессов управления сокращением мышц и свойства восприятия речи оказываются связанными в единую систему управления процессами речеобразования.

Характеристики речевого сигнала определяются процессами разной физической природы. Структура процесса речеобразования, начиная с его конечного результата — речевого сигнала, такова: акустические параметры речевого сигнала определяются формой речевого тракта, импедансом его стенок и граничными условиями на его концах, а также формой возбуждающих импульсов, порождаемых голосовым и другими источниками. Форма речевого тракта и ее изменения создаются вследствие перемещения артикуляторных органов: языка, губ, небной занавески, нижней

челюсти, а длина речевого тракта может изменяться путем подъема или опускания гортани. Перемещения артикулярных органов происходят вследствие сокращения мышц, причем язык и губы сами состоят из мышц и их форма определяется активностью не только внешних, но и внутренних мышц. Степень и последовательность сокращения мышц, а также частично динамика их сокращения определяются характеристиками периферической и центральной систем управления. Взаимодействия между элементами системы управления очень многообразны. Например, вязкоупругие и инерционные характеристики тканей стенок речевого тракта обуславливают потери на низких частотах и частоту первого радиального резонанса. Те же характеристики тканей голосовых складок определяют форму их упругих колебаний. Жесткость мышц влияет на скорость переходных процессов в движениях артикуляторных органов и частоту основного тона. Релаксационные процессы в изменении жесткости мышц являются причиной некоторых особенностей движения артикуляторных органов. Небольшие относительные изменения натяжения голосовых мышц создают очень важные эффекты в колебаниях голосовых складок. Жесткость мышц, в свою очередь, может быть изменена путем воздействия со стороны центральной нервной системы либо непосредственно, либо рефлекторно в ответ на активность мышц-антагонистов. В некоторых случаях, как, например, для нижней челюсти, динамические характеристики артикуляторного органа в значительной степени определяются упругими характеристиками управляющих мышц. Система управления артикуляцией, обеспечивающая требуемое положение и форму артикуляторных органов в определенные моменты, должна считаться с механическими характеристиками речевого аппарата и кодовыми свойствами речи.

Несмотря на различие физической природы, многие процессы речеобразования обладают одним и тем же качественно важным свойством — и акустические колебания, и деформации тканей, и сокращение мышц, и даже процессы управления протекают в непрерывных средах, где преобладают волновые явления. Другое свойство, присущее всем уровням системы речеобразования, — доминирование нестационарных процессов и существенное влияние параметрических эффектов на характеристики речевого сигнала и его восприятие. Параметрические эффекты проявляются, например, в изменении жесткости мышц в процессе сокращения, в изменении частотных характеристик речевого тракта на интервале открытой голосовой щели, в податливости стенок речевого тракта, в изменении сопротивления постоянному току в легких, голосовой щели и речевом тракте.

В настоящее время имеется незначительное число объективных данных о центральных механизмах речеобразования, поэтому в книге им уделено мало места. Остальные вопросы освещены в соответствии с их важностью в процессе речеобразования, объемом экспериментальных данных и степенью достоверности математических моделей.

МЫШЦА

1.1. МИКРОМЕХАНИКА МЫШЕЧНОГО СОКРАЩЕНИЯ

Движение артикуляторных органов происходит под воздействием различных мышц, а некоторые органы, например язык и губы, сами по себе в значительной степени состоят из мышц. Характеристики мышц влияют на динамические свойства артикулярных органов, и это заставляет включить в рассмотрение математическую модель мышечного сокращения.

Мышцы, принимающие участие в процессах артикуляции, относятся к типу так называемых поперечнополосатых. Они состоят из множества продольно расположенных мышечных волокон, каждое из которых содержит около 10^4 последовательно соединенных сократительных единиц — саркомеров [1.25]. В свою очередь, в саркомер входит около 10^6 толстых и тонких нитей, образующих в поперечном сечении гексагональную структуру (рис. 1.1). Тонкие нити саркоме-

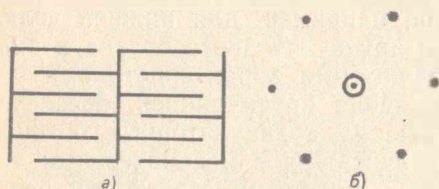


Рис. 1.1. Схема строения саркомера:

а — последовательное соединение сократительных единиц; б — поперечное сечение

ров посередине прикреплены к поперечным мембранам, а их концы свободны и обладают возможностью сдвигаться относительно толстых нитей. Толстые нити содержат белок миозин, а тонкие — белок актин, которые способны взаимодействовать друг с другом через миозиновые мостики, развивая тянущие усилия, заставляющие тонкие нити вдвигаться внутрь саркомера [1.20]. Миозиновые мостики образуются под влиянием ионов кальция, которые выделяются из хранилищ, расположенных вблизи мембран. Энергия активных миозиновых мостиков создается за счет гидролиза аденозинтрифосфорной кислоты (АТФ).

Соотношения между максимальной скоростью сокращения и силой, развиваемой мышцей, было получено Хиллом [1.15] на основе измерения тепловыделения сокращающейся мышцы. Это соотношение получило название уравнения Хилла:

$$v = \frac{b(P_0 - P)}{P + a}, \quad (1.1)$$

где v — скорость изменения длины l укорачивающейся мышцы; P — сила, развиваемая мышцей; P_0 — максимальная сила при нулевой скорости сокращения; a и b — константы. Обычно принимается, что $a \approx 0,25 P_0$, но это соотношение справедливо лишь для больших усилий P , тогда как при малых усилиях $a \approx 0,16 P_0 + 0,18 P$ [1.21]. Уравнение Хилла было проверено на различных мыш-

дах, оно удовлетворительно описывает установившееся соотношение между силой и скоростью сокращения, хотя предпринимаются попытки уточнить это уравнение (например, в [1.14]).

Для описания динамики сокращения мышцы был разработан ряд вязкоупругих моделей, из которых наибольшую популярность получила модель, включающая два упругих и один демпфирующий элемент [1.28] (рис. 1.2). И хотя Хилл, впервые выдвинувший вязкоупругую теорию, впоследствии подверг ее критике, этот подход сохраняет свою силу, а модель мышцы с тремя упругими и двумя демпфирующими элементами используется и в реальных исследованиях [1.29]. Один из важных недостатков вязкоупругих моделей состоит в том, что они игнорируют свойства мышцы как активной среды, представляя ее пассивными элементами. Зависимость между характеристиками сокращения мышцы и внутренним механизмом сокращения исследуется в кинетических моделях сокращения саркомеров, основанных на гипотезе скользящих нитей [1.24].

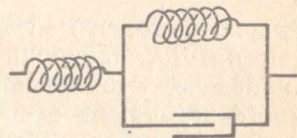


Рис. 1.2. Трехпараметрическая вязкоупругая модель саркомера

В модели, развитой Дешеревским, принимается, что миозиновые мостики могут находиться в одном из двух состояний — свободном и замкнутом, а замкнутый мостик создает либо тянущее, либо тормозящее усилие [1.3]. Для изотонического (с постоянной нагрузкой) сокращения, в частности, записывается следующая система уравнений:

$$\left. \begin{aligned} \frac{dn}{dt} &= k_1 [\alpha (\Delta l) - n - m] + V \frac{n}{\delta} ; \\ \frac{dm}{dt} &= V \frac{n}{\delta} - k_2 m ; \\ \frac{dV}{dt} &= \frac{1}{2NT} [f(n-m) - P], \end{aligned} \right\} \quad (1.2)$$

где n — число тянущих мостиков; α — зависимость числа активных мостиков в половине саркомера от укорочения Δl ; m — число тормозящих мостиков; V — скорость укорочения; k_1 — среднее значение константы скорости замыкания свободных мостиков; k_2 — константа скорости размыкания активных мостиков; δ — интервал активной перестройки мостиков; N — число саркомеров; f — сила, развиваемая мостиком; T — масса нагрузки; P — нагрузка; t — время. Сопоставление поведения модели Дешеревского с поведением реальных объектов показывает, что эта модель удовлетворяет уравнению Хилла в соответствующих режимах работы и позволяет точнее описать динамику сокращения.

Экспериментально установлено, что статическая зависимость между напряжением (усилием, развиваемым мышцей, или нагрузкой) и изменением длины мышцы Δl имеет характер, показанный на рис. 1.3 (по [1.22]). Из этого рисунка видно, что кривая «удлинение — напряжение» существенно нелинейна в области малых напряжений, причем в этой области справедливо эмпирическое соотношение

$$P = C_1 (e^{C_2 \Delta l} - 1) / C_2, \quad (1.3)$$

где C_1 и C_2 — некоторые параметры [1.18, 1.31]. Легко видеть, что зависимость (1.3) является решением обыкновенного дифференциального уравнения

$$\frac{dP}{d\Delta l} - C_2 P = C_1, \quad (1.4)$$

означающего, что мгновенная жесткость мышцы, определяемая как $dP/d\Delta l$, пропорциональна развиваемому ею усилию в этом диапазоне. Это свойство хорошо установлено и подтверждается многочисленными экспериментами [1.3, 1.13—1.16]. Известно также, что жесткость мышцы в максимально напряженном состоянии может отличаться на порядок от жесткости расслабленной мышцы [1.21, 1.26, 1.31]. В области больших напряжений, примерно в диапазоне $0,5 P_0 \leq P \leq P_0$ (где P_0 — максимальное напряжение, развиваемое мышцей), зависимость между напряжением и удлинением хорошо аппроксимируется линейным законом, т. е. укорочение становится пропорциональным напряжению.

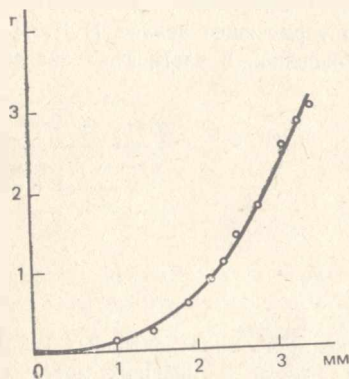


Рис. 1.3. Зависимость «удлинение — напряжение» для голосовых складок человека

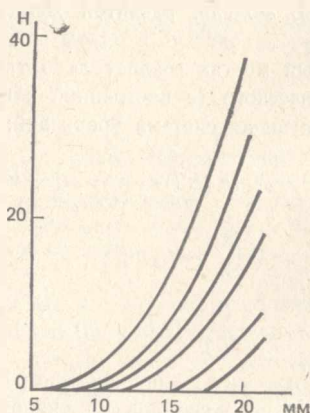


Рис. 1.4. Зависимость силы мышцы децеребрированной кошки от укорочения

Упругие характеристики мышцы не одинаковы для различных режимов ее работы, а могут изменяться под влиянием управляющих команд из центральной нервной системы. Таким управляемым параметром в системе (1.2) является α — число активных мостиков. Исследование роли этого управляемого параметра в [1.13] позволило установить, что он определяется с учетом действия механизма отрицательной обратной связи, который вырабатывает в мышце сигналы, пропорциональные рассогласованию между заданной и фактической длиной мышцы, ее скоростью и напряжением (подробнее об этом механизме см. в гл. 5). В результате влияния центральной нервной системы, управляющей значениями длины и напряжения мышцы в покое (т. е. положением рабочей

точки статической характеристики «удлинение — напряжение»), мышца обладает не одной характеристикой, а целым семейством (рис. 1.4 по [1.13]).

Вместе с тем для каждого значения управляющего параметра вид характеристики «удлинение — напряжение» сохраняет одну и ту же закономерность — жесткость пропорциональна напряжению при малых напряжениях и постоянна при больших. Аналогичными свойствами обладает мышца и относительно эквивалентного коэффициента демпфирования, пропорционального скорости укорочения мышцы. Хотя в основе потерь энергии в мышце лежат не только процессы вязкого трения, но и химические преобразования, связанные с перестройкой миозиновых мостиков, в некотором практически важном диапазоне скоростей сокращения их можно трактовать как вязкость.

1.2. УПРУГИЕ СВОЙСТВА МЫШЕЧНОЙ ТКАНИ

Модели, описывающие микромеханику сокращения мышцы, более или менее полно характеризуют процессы развития напряжений в саркомерах, но нуждаются в дополнении при описании мышцы как макрообъекта, поскольку обычно при ее сокращении активизируется лишь часть саркомеров, тогда как остальные саркомеры и соединительные ткани ведут себя как пассивно деформирующаяся среда. Вследствие того, что мышца состоит из очень большого числа элементов, при количественном описании ее свойств, с одной стороны, допустим статистический подход, учитывающий разброс параметров мышцы посредством вероятностных распределений, а с другой — вполне законно перейти от множества дискретных элементов, составляющих мышцу, к представлению ее в виде непрерывной среды. Такой подход доказал свою плодотворность во многих областях физики и в технических приложениях. Сведения о микромеханике мышечного сокращения при этом не отбрасываются, а берутся в основу описания дифференциальных свойств мышечной ткани как непрерывной среды. Помимо этих локальных свойств необходимо принять во внимание процессы развития напряжений и упругих деформаций в мышцах, рассматриваемых как объекты конечных размеров и обладающих массой и протяженностью во всех трех измерениях. В речевых исследованиях наиболее подходящим для этой цели является аппарат теории упругости, который позволяет не только дать количественный анализ процессов мышечного сокращения, но и описать формы языка и губ, которые они принимают под воздействием сокращения внешних и внутренних мышц. Адекватная модель колебаний голосовых связок может быть построена только с учетом их упругих деформаций.

Рассмотрим элемент объема мышцы, достаточно малый по сравнению с ее размерами, но такой, чтобы в него входило достаточно большое число саркомеров и упругие свойства этого элементарного объема можно было бы рассматривать как некоторые

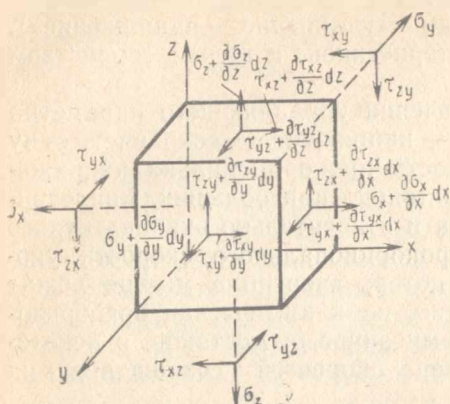


Рис. 1.5. Нормальные и касательные напряжения в элементе упругой среды

оси объемных сил (массы, силы инерции), отнесенных к единице объема, через X, Y, Z . Тогда дифференциальные уравнения равновесия, или уравнения Навье, записываются в следующем виде [1.8]:

$$\left. \begin{aligned} \frac{\partial \sigma_x}{\partial x} + \frac{\partial \tau_{xy}}{\partial y} + \frac{\partial \tau_{xz}}{\partial z} + X &= 0; \\ \frac{\partial \tau_{yx}}{\partial x} + \frac{\partial \sigma_y}{\partial y} + \frac{\partial \tau_{yz}}{\partial z} + Y &= 0; \\ \frac{\partial \tau_{zx}}{\partial x} + \frac{\partial \tau_{zy}}{\partial y} + \frac{\partial \sigma_z}{\partial z} + Z &= 0. \end{aligned} \right\} \quad (1.5)$$

Для изотропного тела справедлив закон парности касательных напряжений: $\tau_{yx} = \tau_{xy}$; $\tau_{zy} = \tau_{yz}$; $\tau_{zx} = \tau_{xz}$. Система (1.5) дополняется так называемыми условиями на поверхности, позволяющими определить составляющие напряжений на любой наклонной площадке с нормалью n через составляющие напряжений на площадках, параллельных координатным осям:

$$\left. \begin{aligned} X_n &= \sigma_x \cos(x, n) + \tau_{xy} \cos(y, n) + \tau_{xz} \cos(z, n); \\ Y_n &= \tau_{yx} \cos(x, n) + \sigma_y \cos(y, n) + \tau_{yz} \cos(z, n); \\ Z_n &= \tau_{zx} \cos(x, n) + \tau_{zy} \cos(y, n) + \sigma_z \cos(z, n), \end{aligned} \right\} \quad (1.6)$$

где $\cos(x, n)$, $\cos(y, n)$, $\cos(z, n)$ — направляющие косинусы для нормали n .

При деформации тела каждая точка с координатами (x, y, z) смещается, проекции этого смещения на оси x, y, z являются функциями исходных координат $u(x, y, z)$, $v(x, y, z)$, $w(x, y, z)$. Линейными деформациями по направлениям координатных осей называют относительное удлинение вдоль этих осей:

$$\epsilon_x = \frac{\partial u}{\partial x}; \quad \epsilon_y = \frac{\partial v}{\partial y}; \quad \epsilon_z = \frac{\partial w}{\partial z}. \quad (1.7)$$

Наряду с линейными деформациями рассматривают и угловые деформации, которые в некоторой плоскости равны сумме частных производных составляющих перемещения в этой плоскости по переменным в перпендикулярных направлениях:

$$\gamma_{xy} = \frac{\partial u}{\partial y} + \frac{\partial v}{\partial x}; \quad \gamma_{yz} = \frac{\partial v}{\partial z} + \frac{\partial w}{\partial y}; \quad \gamma_{zx} = \frac{\partial w}{\partial x} + \frac{\partial u}{\partial z}.$$

Объемная деформация θ есть сумма линейных деформаций по каждой из координатных осей: $\theta = \epsilon_x + \epsilon_y + \epsilon_z$. Условие неразрывности тела при упругих деформациях добавляет систему уравнений, связывающих линейные и угловые деформации:

$$\begin{aligned} \frac{\partial^2 \epsilon_x}{\partial y^2} + \frac{\partial^2 \epsilon_y}{\partial x^2} &= \frac{\partial^2 \gamma_{xy}}{\partial x \partial y}; \\ \frac{\partial^2 \epsilon_y}{\partial z^2} + \frac{\partial^2 \epsilon_z}{\partial y^2} &= \frac{\partial^2 \gamma_{yz}}{\partial y \partial z}; \\ \frac{\partial^2 \epsilon_z}{\partial x^2} + \frac{\partial^2 \epsilon_x}{\partial z^2} &= \frac{\partial^2 \gamma_{zx}}{\partial z \partial x}; \\ \frac{\partial}{\partial x} \left(\frac{\partial \gamma_{zx}}{\partial y} + \frac{\partial \gamma_{xy}}{\partial z} - \frac{\partial \gamma_{yz}}{\partial x} \right) &= 2 \frac{\partial^2 \epsilon_x}{\partial y \partial z}; \\ \frac{\partial}{\partial y} \left(\frac{\partial \gamma_{xy}}{\partial z} + \frac{\partial \gamma_{yz}}{\partial x} - \frac{\partial \gamma_{zx}}{\partial y} \right) &= 2 \frac{\partial^2 \epsilon_y}{\partial z \partial x}; \\ \frac{\partial}{\partial z} \left(\frac{\partial \gamma_{yz}}{\partial x} + \frac{\partial \gamma_{zx}}{\partial y} - \frac{\partial \gamma_{xy}}{\partial z} \right) &= 2 \frac{\partial^2 \epsilon_z}{\partial x \partial y}. \end{aligned} \quad (1.8)$$

На следующем этапе анализа упругих деформаций тел в теории упругости рассматривают связь между напряжением, приложенным к элементарному объему, и смещениями его точек. В этой области органические ткани сильно отличаются от обычных конструкционных материалов структурой тканей и физикой процессов удлинения и укорочения мышцы. В мышечных тканях процессы укорочения и удлинения связаны с принципиально иными энергетическими процессами, чем те, которые протекают при упругих деформациях пассивных сред. Тем не менее если интересоваться лишь внешними характеристиками упругих деформаций мышц, то в некотором диапазоне практически важных режимов работы мышцы вполне применим аппарат теории упругости.

Важной характеристикой мышечных тканей является их анизотропность, т. е. различие свойств ткани вдоль мышечных волокон и в плоскости, перпендикулярной этому направлению. В этой плоскости мышцу приближенно можно считать изотропной. Другое отличительное свойство тканей состоит в их практической несжимаемости. В диапазоне малых напряжений необходимо принимать во внимание нелинейную зависимость укорочения от напряжения, описываемую (1.4). Все эти особенности создают значительные трудности при анализе упругих деформаций тканей. Обычный путь преодоления этих трудностей состоит в использовании упрощенных моделей, причем степень и характер упрощения зависят

от конкретной задачи и справедливость ее решения ограничивается принятыми допущениями.

В предыдущем параграфе было установлено, что при достаточно больших напряжениях жесткость мышцы есть величина постоянная, т. е. в этом диапазоне справедлив закон Гука, описывающий линейную зависимость между приложенным усилием σ и деформацией ε : $\varepsilon = E\sigma$, где E — модуль упругости материала (модуль Юнга). Закон Гука является одной из основных аксиом, используемых в математическом аппарате теории упругости, поэтому соответствие поведения мышечных тканей этому закону в некотором диапазоне параметров определяет возможность применения методов теории упругости. Если же диапазон изменения напряжений в мышечной ткани невелик по сравнению со средней точкой, то характеристика «напряжение — укорочение» может быть линеаризована и на участке малых напряжений, позволяя распространить действие закона Гука и на этот диапазон. Такая линейная модель мышечных тканей применима во многих практических важных режимах работы мышцы, хотя существуют процессы, для которых нелинейность упругих характеристик мышц является доминирующим фактором [1.19].

Упругие постоянные живых тканей определяют экспериментально по степени растяжения препарата мышцы под воздействием грузов различной массы или с помощью измерения импеданса тканей на различных частотах. Здесь необходимо различать значение модуля упругости тканей при всестороннем сжатии, которое вследствие несжимаемости чрезвычайно велико — порядка 10^{10} Па [1.30], и то значение эквивалентного показателя жесткости мышцы, которое определяет связь между укорочением мышцы и напряжением, развиваемым при сдвиге нитей саркомеров относительно друг друга. Именно этот эквивалентный показатель жесткости определяет упругие характеристики живой мышечной ткани, и мы в дальнейшем будем уже без оговорок называть его модулем упругости. Эксперименты с нагружением препаратов тканей показывают, что модуль Юнга не одинаков для разных тканей и находится в диапазоне $5 \cdot (10^3 - 10^4)$ Па (для напряженной мышцы — на порядок больше) для мышечных тканей, $10^4 - 10^5$ Па для соединительных тканей и $10^5 - 10^6$ Па для поверхностных тканей (эпителия) [1.22, 1.26].

Следующая упругая постоянная характеризует связь между линейными деформациями в продольном (ε_x) и поперечном (ε_y , ε_z) направлениях. Эта постоянная называется коэффициентом Пуассона:

$$\nu = \varepsilon_y / \varepsilon_x. \quad (1.9)$$

Для практически несжимаемых материалов типа резины коэффициент Пуассона примерно равен 0,5 [1.5]. Мышечные ткани во многом подобны резине, однако количественные показатели для них несколько различаются. В отношении коэффициента Пуассона это показано в [1.26] путем сравнения модулей Юнга E и

сдвига G . Последний определяется как отношение между касательным напряжением τ и угловой деформацией γ :

$$G = \tau / \gamma. \quad (1.10)$$

По оценкам [1.30] G близко к $5 \cdot 10^3$ Па, а по оценкам [1.26] — к $5 \cdot 10^4$ Па, и, поскольку между названными модулями существует связь

$$E = 2(1 + \nu)G \quad (1.11)$$

при $\nu = 0,5$ (т. е. если бы свойства мышц относительно несжимаемости были аналогичны свойствам резины), модуль Юнга должен был бы превосходить модуль сдвига в 3 раза. Однако по данным [1.26] отношение E/G значительно превышает 3. Вместе с тем согласно [1.5] коэффициент Пуассона для пассивных сред находится в диапазоне $0 \leq \nu \leq 0,5$, и либо эксперименты, описанные в [1.26], страдают какой-то погрешностью, либо соотношение (1.11) не применимо к мышечным тканям из-за их анизотропии и иной физической природы соотношения между линейными и угловыми деформациями мышечных тканей.

Анизотропия мышечных тканей проявляется в том, что модули упругости в направлении вдоль мышечных волокон и в любом направлении в перпендикулярной плоскости отличаются примерно на порядок, что связано со структурой мышечных тканей и различием молекулярных связей вдоль и поперек волокон. Принимая во внимание анизотропию мышечных тканей, направим ось x вдоль мышечных волокон и обозначим модуль Юнга в этом направлении через E_x , а модули Юнга в плоскости, перпендикулярной оси x , обозначим через E_y и E_z , причем примем, что в этой плоскости свойства мышцы изотропны и $E_y = E_z = E'$. Допуская справедливость закона Гука для упругих деформаций мышечных тканей, найдем линейные деформации вдоль каждой из координатных осей с учетом соотношения (1.9):

$$\left. \begin{aligned} \epsilon_x &= \frac{\sigma_x}{E_x} - \frac{\nu}{E'} (\sigma_y + \sigma_z); \\ \epsilon_y &= \frac{\sigma_y}{E'} - \nu \left(\frac{\sigma_z}{E'} + \frac{\sigma_x}{E_x} \right); \\ \epsilon_z &= \frac{\sigma_z}{E'} - \nu \left(\frac{\sigma_x}{E_x} + \frac{\sigma_y}{E'} \right). \end{aligned} \right\} \quad (1.12)$$

Поскольку модуль сдвига в поперечной плоскости G' также отличается от модуля сдвига вдоль мышечных волокон G_x , из (1.10) следует:

$$\gamma_{xy} = \tau_{xy} / G_x; \quad \gamma_{yz} = \tau_{yz} / G'; \quad \gamma_{zx} = \tau_{zx} / G'. \quad (1.13)$$

Уравнения (1.12) и (1.13) называют обобщенным законом Гука, учитывающим в случае мышечных тканей их анизотропию. Из (1.12) найдем объемную деформацию

$$\theta = \epsilon_x + \epsilon_y + \epsilon_z = \frac{1 - 2\nu}{E_x} \sigma_x + \frac{1 - 2\nu}{E'} (\sigma_y + \sigma_z).$$

Добавим и вычтем величину $\nu \sigma_x / E_x$ в первом уравнении системы (1.12), получив при этом $\varepsilon_x = \frac{1+\nu}{E_x} \sigma_x - \nu \frac{\theta}{1-2\nu}$. Отсюда можно найти связь между напряжением и линейной деформацией

$$\sigma_x = \frac{E_x}{1+\nu} \varepsilon_x + \frac{E_x \nu}{(1+\nu)(1-2\nu)} \theta. \quad (1.14)$$

Обозначив

$$\lambda_x = \frac{E_x \nu}{(1+\nu)(1-2\nu)}; \quad \mu_x = \frac{E_x}{1+\nu}, \quad (1.15)$$

перепишем (1.14) как

$$\sigma_x = \lambda_x \theta + 2 \mu_x \varepsilon_x, \quad (1.16)$$

где упругие постоянные λ_x и μ_x называются коэффициентом Ламе. Очевидно, что $\mu_x = 2G_x$, что следует из сопоставления (1.15) и (1.11). Аналогичным образом найдем выражения, связывающие деформации и напряжения в других направлениях:

$$\sigma_y = \lambda_y \theta + 2 \mu_y \varepsilon_y; \quad \sigma_z = \lambda_z \theta + 2 \mu_z \varepsilon_z, \quad (1.17)$$

причем в силу изотропности мышечных тканей в поперечной плоскости имеем $\lambda_y = \lambda_z = \lambda'$ и $\mu_y = \mu_z = \mu'$. К уравнениям (1.16), (1.17) добавляется очевидная система:

$$\tau_{xy} = \mu_x \gamma_{xy}; \quad \tau_{yz} = \mu_y \gamma_{yz}; \quad \tau_{zx} = \mu_z \gamma_{zx}. \quad (1.18)$$

В большинстве задач, связанных с упругими деформациями мышечных тканей, нас будут интересовать не напряжения, а перемещения. Поэтому воспользуемся системами (1.17), (1.18) и (1.18) для того, чтобы исключить напряжения из системы (1.5). Тогда для направления вдоль оси x получим

$$-\frac{\partial}{\partial x} (\lambda_x \theta + 2 \mu_x \varepsilon_x) + \frac{\partial}{\partial y} \mu_x \gamma_{xy} + \frac{\partial}{\partial z} \mu_z \gamma_{zx} + X = 0.$$

Действуя по аналогии для направлений вдоль осей y и z и учитывая (1.7), в конечном итоге находим:

$$\left. \begin{aligned} (\lambda_x + \mu_x) \frac{\partial \theta}{\partial x} + \mu_x \nabla^2 u + X &= 0; \\ (\lambda_y + \mu_y) \frac{\partial \theta}{\partial y} + \mu_y \nabla^2 v + Y &= 0; \\ (\lambda_z + \mu_z) \frac{\partial \theta}{\partial z} + \mu_z \nabla^2 w + Z &= 0. \end{aligned} \right\} \quad (1.19)$$

В силу изотропности мышц в плоскости, перпендикулярной направлению мышечных волокон, систему (1.19) можно записать так:

$$\left. \begin{aligned} (\lambda_x + \mu_x) \frac{\partial \theta}{\partial x} + \mu_x \nabla^2 u + X &= 0; \\ (\lambda' + \mu') \frac{\partial \theta}{\partial y} + \mu' \nabla^2 v + Y &= 0; \\ (\lambda' + \mu') \frac{\partial \theta}{\partial z} + \mu' \nabla^2 w + Z &= 0. \end{aligned} \right\} \quad (1.20)$$

В системах (1.19) и (1.20) символ ∇^2 представляет собой лапласиан

$$\nabla^2 = \frac{\partial^2}{\partial x^2} + \frac{\partial^2}{\partial y^2} + \frac{\partial^2}{\partial z^2}. \quad (1.21)$$

В таком виде система (1.20) совпадает с формой, известной в теории упругости для изотропного тела [1.12], но с учетом различий в механических свойствах мышечных тканей по разным направлениям. Принимая во внимание несжимаемость тканей, в результате чего $\lambda \gg \mu$ [1.30], можно далее упростить систему (1.20):

$$\left. \begin{aligned} \lambda_x \frac{\partial \theta}{\partial x} + \mu_x \nabla^2 u + X &= 0; \\ \lambda' \frac{\partial \theta}{\partial y} + \mu' \nabla^2 v + Y &= 0; \\ \lambda' \frac{\partial \theta}{\partial z} + \mu' \nabla^2 w + Z &= 0. \end{aligned} \right\} \quad (1.22)$$

Преобразуя аналогичным образом систему (1.6) для граничных условий, получаем три условия на поверхности, выраженные через перемещения вдоль каждой из координатных осей.

Рассмотрим теперь динамику сокращения мышцы, для чего добавим в систему (1.22) инерционные члены, не принимая пока во внимание вязкость и пренебрегая объемными силами:

$$\left. \begin{aligned} \lambda_x \frac{\partial \theta}{\partial x} + \mu_x \nabla^2 u - \rho \frac{\partial^2 u}{\partial t^2} &= 0; \\ \lambda' \frac{\partial \theta}{\partial y} + \mu' \nabla^2 v - \rho \frac{\partial^2 v}{\partial t^2} &= 0; \\ \lambda' \frac{\partial \theta}{\partial z} + \mu' \nabla^2 w - \rho \frac{\partial^2 w}{\partial t^2} &= 0, \end{aligned} \right\} \quad (1.23)$$

где ρ — плотность тканей, равная примерно 1,057—1,059 г/см³.

Известно, что в упругих телах распространяются волны двух типов: продольные (волны расширения) и поперечные (волны искажения) [1.5, 1.12], причем появление одного типа волн всегда сопровождается появлением другого. Продольные волны соответствуют деформациям без вращения элементов объема, а поперечные волны — деформациям сдвига и вращения. Если определить вращение объема как

$$\omega_x = \frac{1}{2} \left(\frac{\partial w}{\partial y} - \frac{\partial v}{\partial z} \right); \quad \omega_y = \frac{1}{2} \left(\frac{\partial u}{\partial z} - \frac{\partial w}{\partial x} \right); \quad \omega_z = \frac{1}{2} \left(\frac{\partial v}{\partial x} - \frac{\partial u}{\partial y} \right).$$

то условия отсутствия вращений записываются в виде:

$$\frac{\partial v}{\partial x} - \frac{\partial u}{\partial y} = 0; \quad \frac{\partial w}{\partial y} - \frac{\partial v}{\partial z} = 0; \quad \frac{\partial u}{\partial z} - \frac{\partial w}{\partial x} = 0.$$

Тогда перемещения u , v , w можно представить с помощью одной функции Φ [1.12] следующим образом: $u = \frac{\partial \Phi}{\partial x}$; $v = \frac{\partial \Phi}{\partial y}$; $w = \frac{\partial \Phi}{\partial z}$. Отсюда $\theta = \nabla^2 \Phi$ и

$$\left. \begin{aligned} \frac{\partial \theta}{\partial x} &= \frac{\partial}{\partial x} \nabla^2 \Phi = \nabla^2 u; \\ \frac{\partial \theta}{\partial y} &= \frac{\partial}{\partial y} \nabla^2 \Phi = \nabla^2 v; \\ \frac{\partial \theta}{\partial z} &= \frac{\partial}{\partial z} \nabla^2 \Phi = \nabla^2 w. \end{aligned} \right\} \quad (1.24)$$

Подставляя (1.24) в (1.23), находим:

$$\left. \begin{aligned} (\lambda_x + \mu_x) \nabla^2 u - \rho \frac{\partial^2 u}{\partial t^2} &= 0; \\ (\lambda' + \mu') \nabla^2 v - \rho \frac{\partial^2 v}{\partial t^2} &= 0; \\ (\lambda' + \mu') \nabla^2 w - \rho \frac{\partial^2 w}{\partial t^2} &= 0. \end{aligned} \right\} \quad (1.25)$$

Если один из геометрических размеров упругого тела значительно больше или значительно меньше остальных и действующие силы лежат в плоскости, перпендикулярной оси, направленной вдоль этого геометрического размера (например, оси z), но такое состояние носит название плоской деформации и для его анализа часто используется так называемая функция Эри $\psi(x, y)$, связанная с компонентами напряжения следующим образом:

$$\sigma_x = \frac{\partial^2 \psi}{\partial y^2} - \rho g y; \quad \sigma_y = \frac{\partial^2 \psi}{\partial x^2} - \rho g y; \quad \tau_{xy} = -\frac{\partial^2 \psi}{\partial x \partial y},$$

где g — ускорение свободного падения [1.12]. Подставив эти выражения в (1.5), найдем, что функция $\psi(x, y)$ должна удовлетворить уравнению

$$\frac{\partial^4 \psi}{\partial x^4} + 2 \frac{\partial^4 \psi}{\partial x^2 \partial y^2} + \frac{\partial^4 \psi}{\partial y^4} = 0.$$

Использование функции Эри в некоторых случаях позволяет применить эффективные методы решения задачи упругих деформаций, например метод разделения переменных (метод Фурье).

Если предположить, что вдоль оси x распространяется продольная волна и все частицы движутся параллельно распространению этой волны (т. е. $v=w=0$), то смещение u будет функцией лишь одной координаты x и соответствующее уравнение колебаний выглядит как

$$(\lambda_x + \mu_x) \frac{\partial^2 u}{\partial x^2} = \rho \frac{\partial^2 u}{\partial t^2}.$$

Отсюда получаем скорость распространения продольной волны

$$c_{пр} = \sqrt{\frac{\lambda_x + \mu_x}{\rho}} \approx \sqrt{\frac{\lambda_x}{\rho}}. \quad (1.26)$$

В поперечных волнах, когда частицы колеблются в плоскости zoy , перпендикулярной оси x , смещения $u=w=0$, если принять, например, что существует лишь один вид поперечных колебаний, а именно колебаний вдоль оси y . Тогда уравнение колебаний есть

$$(\lambda' + \mu') \frac{\partial^2 v}{\partial x^2} = \rho \frac{\partial^2 v}{\partial t^2},$$

и соответствующая скорость распространения поперечных колебаний находится как

$$c_{поп} = \sqrt{\frac{\lambda' + \mu'}{\rho}}. \quad (1.27)$$

Для мышечных тканей $\lambda' \ll \lambda_x$, и, следовательно, скорость распространения поперечных волн в них значительно меньше скорости распространения продольных волн.

Если представить мышцу в виде цилиндра с произвольной формой поперечного сечения, то окажется, что условие $v=w=0$, необходимое для распространения продольных волн, приводит к таким требованиям на компоненты боковых напряжений σ_y и σ_z , которые трудно удовлетворить для тел ограниченного объема [1.12]. Вообще проблема отыскания решений системы (1.25), удовлетворяющих конкретным граничным условиям реальной задачи, за редкими исключениями оказывается аналитически неразрешимой. Это заставляет искать упрощенные подходы, обеспечивающие содержательные по физическому смыслу решения без формального анализа системы (1.25). В отношении мышечных тканей упрощенные модели оправдываются тем фактом, что необходимые для точного решения системы (1.25) геометрические параметры, граничные условия и упругие постоянные никогда не бывают известны без погрешности и точные решения точной системы (1.25) с неточными значениями параметров не имели бы никакого смысла (во всяком случае такого смысла, который оправдал бы затраченные на решение усилия). Вместе с тем проведенный в данном разделе анализ упругих свойств мышечных тканей необходим для построения упрощенных моделей, и в дальнейшем будут использованы отдельные элементы приведенного здесь математического аппарата.

1.3. ОДНОМЕРНАЯ МОДЕЛЬ МЫШЦЫ

Рассмотрим модель мышцы, представленной в виде цилиндра длиной l и постоянной площадью поперечного сечения S , причем $l \gg \sqrt{S/\pi}$. Примем, что свойства тканей изотропны по длине мышцы. Допустим также, что деформации происходят только вдоль мышечных волокон, а деформациями в остальных

направлениях пренебрежем. Пусть в результате возникновения внутренних сил, связанных с сокращением саркомеров, в некотором сечении x цилиндра возникло натяжение P_x . Тогда под действием этого натяжения в слое, заключенном между координатами x и $x + \Delta x$, произойдет сжатие на величину $\Delta u = P_x / ES$ (предполагая, что мышца сокращается в диапазоне напряжений, позволяющем линеаризовать характеристику «напряжение — укорочение»). Обозначим напряжение в сечении $x + \Delta x$ как $P_{x+\Delta x}$. Тогда результирующая сила, действующая на слой ткани, заключенный между координатами x и $x + \Delta x$, в дифференциальной форме запишется как $P_{x+\Delta x} - P_x = ES \frac{\partial^2 u}{\partial x^2} \Delta x$. Эта сила уравновеши-

вается силой ускорения слоя $\rho S \frac{\partial^2 u}{\partial t^2} \Delta x$, где ρ — объемная плотность тканей. Не принимая пока во внимание силы, пропорциональной скорости сжатия, запишем уравнение вынужденных продольных колебаний цилиндра как

$$\frac{E}{\rho} \frac{\partial^2 u}{\partial x^2} = \frac{\partial^2 u}{\partial t^2} - \frac{F(x, t)}{\rho}, \quad (1.28)$$

где $F(x, t)$ — распределение усилий по длине мышцы, а отношение E/ρ является квадратом скорости распространения продольных волн, которая несколько меньше скорости, определенной по более точной формуле (1.26).

Экспериментальные исследования показывают, что вязкость оказывает большое влияние на процессы мышечного сокращения, причем, как упоминалось в § 1.2, демпфирование колебаний происходит также и за счет химических преобразований мышечных белков. По оценкам [1.26] эквивалентный коэффициент вязкого трения для мышечных тканей лежит в диапазоне значений от 10 Пуаз на частоте 0,1 Гц до 5 Пуаз на частоте 100 Гц. Экспериментальные оценки коэффициента затухания для скелетных мышц, приведенные в [1.7], дают значения в диапазоне 5—10 с⁻¹. Форма, в которой вязкое трение должно быть учтено в уравнении движения мышцы, зависит от модели внутренней структуры мышцы. Если допустить, что вязкое трение связано с объемными деформациями тканей, то соответствующий член принимает вид $r \partial^3 u / \partial x^2 \partial t$. Если же принять, что вязкое трение возникает главным образом при скольжении нитей саркомеров относительно друг друга, то сила трения может быть представлена как $r \partial u / \partial t$. Последняя модель как будто более соответствует действительной структуре мышцы и к тому же существенно упрощает анализ волнового уравнения (1.28).

Таким образом, учет сил, эквивалентных вязкому трению, приводит к следующему виду волнового уравнения:

$$\frac{E}{\rho} \frac{\partial^2 u}{\partial x^2} = \frac{\partial^2 u}{\partial t^2} + \frac{r}{\rho} \frac{\partial u}{\partial t} - \frac{F(x, t)}{\rho}.$$

Произведя, однако, замену переменной

$$u(x, t) = \bar{u}(x, t) e^{-\frac{r}{2\rho} t} \quad (1.29)$$

и пренебрегаем членом с $(r/2\rho)^2$, вновь получаем уравнение в форме (1.28). Мы будем в дальнейшем пользоваться обозначением $u(x, t)$ вместо $\bar{u}(x, t)$, памятуя о необходимости учета затухания по (1.29) после решения (1.28). Уравнение (1.28) должно быть дополнено начальными и граничными условиями. Начальные условия записываются как

$$u(x, 0) = f_1(x); \quad \frac{\partial u(x, 0)}{\partial t} = f_2(x),$$

где $f_1(x)$ и $f_2(x)$ — некоторые функции, определяющие смещения и скорости каждого сечения мышцы в начальный момент.

Сокращение мышцы обычно приводит к перемещению управляемого ею органа. Но к этому же органу прикреплены и другие мышцы — как синергисты, действующие в том же направлении, так и антагонисты, обеспечивающие движение органа в противоположном направлении. Кроме того, каждый орган связан с пассивными тканями, обладающими некоторой вязкостью и упругостью. В результате всякое перемещение органа встречает упругое сопротивление F_y и противодействующую силу со стороны мышц-антагонистов или внешней среды $F_l(t)$. Примем, что упругое сопротивление пропорционально укорочению мышцы $F_y = c_l u(l, t)$. В общем случае масса управляемого органа и присоединенных тканей M проявляет себя в виде статической нагрузки, растягивающей мышцу под действием сил массы и в виде инерционного члена, пропорционального ускорению конца мышцы. Если направление силы тяжести совпадает с направлением мышечных волокон, то граничные условия на конце мышцы при $x=l$ запишутся как

$$-ES \frac{\partial u(x, t)}{\partial x} \Big|_{x=l} = -F_l(t) + M \frac{\partial^2 u(x, t)}{\partial t^2} \Big|_{x=l} + c_l u(l, t) - Mg, \quad (1.30)$$

где g — ускорение силы тяжести. При $x=0$ примем, что конец мышцы жестко закреплен, что соответствует граничному условию

$$u(0, t) = 0. \quad (1.31)$$

Рассмотрим однородное уравнение (т. е. при отсутствии внутренних напряжений, возникающих за счет сокращения саркомеров, $F(x, t) = 0$). Применим метод разделения переменных (метод Фурье) [1.4], при котором неизвестная переменная $u(x, t)$ представляется в виде произведения двух функций, каждая из которых зависит только от одного аргумента — пространственной координаты или времени:

$$u(x, t) = U(x)T(t). \quad (1.32)$$

Подставляя (1.32) в (1.28), получаем

$$\frac{\rho T''(t)}{ET(t)} = \frac{U''(x)}{U(x)} = -\lambda^2, \quad (1.33)$$

где λ — некоторая константа, означающая, что первый и второй члены (1.33), будучи функциями различных аргументов, могут быть равны друг другу только в том случае, если они тождественно равны постоянной величине. Отсюда получаем систему:

$$U'' + \lambda^2 U = 0; \quad T'' + \lambda^2 \frac{E}{\rho} T = 0. \quad (1.34); (1.35)$$

Решение (1.34) есть $U(x) = C_1 \cos \lambda x + C_2 \sin \lambda x$, откуда, принимая во внимание граничные условия (1.31), при $x=0$ получаем $C_1=0$ и собственные функции представляются как

$$U_k(x) = C_{2k} \sin \lambda_k x. \quad (1.36)$$

Коэффициент C_{2k} определяется из условия ортонормированности собственных функций $\int_0^l U_k^2(x) dx = 1$, откуда

$$C_{2k} = \frac{1}{\sqrt{\int_0^l \sin^2 \lambda_k x dx}} = \frac{2 \sqrt{\lambda_k l}}{\sqrt{2 \lambda_k l - \sin 2 \lambda_k l}}.$$

Решение (1.35) есть

$$T(t) = D_{1k} \cos \omega_k t + D_{2k} \sin \omega_k t, \quad (1.37)$$

$$\text{где } \omega_k = \lambda_k \sqrt{E/\rho}. \quad (1.38)$$

Для упрощения примем, что сила $F_l(t)$ на конце мышцы отсутствует, и определим влияние инерционной и упругой нагрузок на характеристики движения этого конца. Нагрузка, выражаемая членом Mg , создает неоднородность в граничном условии (1.30). Чтобы избавиться от этой неоднородности, введем новую переменную $\tilde{u}(x, t)$ и вспомогательную функцию $R(x)$, связав их с $u(x, t)$ следующим образом:

$$u(x, t) = \tilde{u}(x, t) + R(x). \quad (1.39)$$

Перепишем граничное условие (1.30) с учетом (1.39):

$$\begin{aligned} ES \frac{\partial \tilde{u}(x, t)}{\partial x} \Big|_{x=l} + c_l \tilde{u}(l, t) + M \frac{\partial^2 \tilde{u}(x, t)}{\partial t^2} \Big|_{x=l} = \\ = Mg - ES \frac{\partial R(x)}{\partial x} \Big|_{x=l} - c_l R(l). \end{aligned} \quad (1.40)$$

Потребуем, чтобы правая часть (1.40) была равна нулю для всех x , что приводит к обыкновенному дифференциальному уравнению первого порядка:

$$R'(x) + \frac{c_l}{ES} R(x) - \frac{Mg}{ES} = 0. \quad (1.41)$$

Общее решение (1.41) есть

$$R(x) = e^{-\int_0^x \frac{c_l}{ES} dx} \left[R(0) + \int_0^x \frac{Mg}{ES} e^{\int_0^x \frac{c_l}{ES} dx} dx \right].$$

Для сохранения однородности граничных условий при $x=0$ необходимо, чтобы $R(0)=0$, откуда получаем при $ES=\text{const}$ $R(x)=$

$= \frac{Mg}{c_l} \left(1 - e^{-\frac{c_l}{ES} x} \right)$. Теперь волновое уравнение для новой переменной $\tilde{u}(x, t)$ есть

$$E \frac{\partial^2 \tilde{u}}{\partial x^2} - \rho \frac{\partial^2 \tilde{u}}{\partial t^2} = \tilde{F}(x, t), \quad (1.42)$$

где новая возбуждающая функция $\tilde{F}$ получается путем подстановки в (1.28) выражения (1.39) и переноса в правую часть членов, не содержащих $\tilde{u}$ или ее производных:

$$\tilde{F}(x, t) = \frac{Mg c_l}{ES^2} e^{-\frac{c_l}{ES} x} - F(x, t).$$

Граничное условие при $x=l$ для (1.42) теперь однородно:

$$ES \frac{\partial \tilde{u}(x, t)}{\partial x} \Big|_{x=l} + c_l \tilde{u}(l, t) + M \frac{\partial^2 \tilde{u}(x, t)}{\partial t^2} \Big|_{x=l} = 0.$$

Для поиска собственных чисел разделим переменные $\tilde{u}(x, t) = \tilde{U}(x) \tilde{T}(t)$, что дает для однородного уравнения ($\tilde{F}(x, t)=0$) систему, аналогичную (1.34), (1.35):

$$\tilde{u}'' + \lambda^2 \tilde{u} = 0; \quad \tilde{T}'' + \lambda^2 \frac{E}{\rho} \tilde{T} = 0.$$

Разделяя переменные в новых граничных условиях при $x=l$, получаем $ES \tilde{U}'(l) \tilde{T} + c_l \tilde{U}(l) \tilde{T} + M \tilde{U}(l) \tilde{T}'' = 0$ и, следовательно,

$$ES \tilde{U}'(l) + c_l \tilde{U}(l) - \lambda^2 \frac{ME}{\rho} \tilde{U}(l) = 0, \text{ откуда имеем}$$

$$SE \tilde{U}'(l) = \left(\lambda^2 \frac{ME}{\rho} - c_l \right) \tilde{U}(l). \quad (1.43)$$

Граничное условие (1.31) позволяет определить собственные функции для $\tilde{U}(x)$, которые совпадают по форме с (1.36). Затем подставляя (1.36) в (1.43), получаем трансцендентное уравнение для поиска собственных чисел

$$\left(\lambda \frac{M}{\rho S} - \frac{1}{\lambda} \frac{c_l}{SE} \right) \operatorname{tg} \lambda l = 1. \quad (1.44)$$

Обозначим массу мышцы как $m = \rho S l$ и перепишем (1.44):

$$\left(\lambda l \frac{M}{m} - \frac{c_l}{\lambda SE} \right) \operatorname{tg} \lambda l = 1. \quad (1.45)$$

Если масса нагрузки значительно больше массы мышцы ($M \gg m$), то можно использовать приближенное соотношение $\operatorname{tg} \lambda^{(1)} l \approx \lambda^{(1)} l$, и тогда $(\lambda^{(1)} l)^2 \frac{M}{m} - \frac{c_l l}{SE} = 1$, откуда

$$\lambda^{(1)} = \frac{1}{l} \sqrt{\frac{m}{M} \left(1 + \frac{c_l l}{SE} \right)}. \quad (1.46)$$

Если масса мышцы сопоставима с массой нагрузки, то можно взять следующий член разложения тангенса в степенной ряд: $\operatorname{tg} \lambda^{(2)} l \approx \lambda^{(2)} l + \frac{(\lambda^{(2)} l)^3}{3}$, и тогда, например, при отсутствии упругой нагрузки собственные числа определяются как

$$\lambda^{(2)} = \frac{1}{l} \sqrt{\frac{1}{\frac{M}{m} + \frac{1}{3}}}. \quad (1.47)$$

При $m=M$ значения $\lambda^{(2)}$, определенные по (1.47), дают по сравнению с точным решением ошибку менее 1% [2.10], и во всех случаях, когда масса мышцы меньше массы нагрузки, ошибка второго приближения пренебрежимо мала. В большинстве же случаев справедливо условие $m \ll M$, и определение собственных чисел по (1.46) вполне удовлетворяет требованиям на погрешность вычислений. К тому же значения геометрических параметров и упругих постоянных мышцы редко бывают известны с точностью, лучшей, чем 5—10%, поэтому нецелесообразно требовать большей точности вычислений для собственных чисел. Поскольку $\operatorname{tg} \lambda l$ есть периодическая функция, значения остальных корней трансцендентного уравнения (1.45) в первом приближении находятся как $\lambda_k l = \lambda^{(1)} l + k\pi$, $k=0, 1, 2, \dots$ Таким образом, решение однородного уравнения представляется в виде бесконечного ряда

$$\tilde{u}(x, t) = \sum_{k=1}^{\infty} C_{2k} \sin \lambda_k x (D_{1k} \cos \omega_k t + D_{2k} \sin \omega_k t). \quad (1.48)$$

Начальные условия для (1.42) также претерпевают изменения:

$$\left. \begin{aligned} \tilde{u}(x, 0) = u(x, 0) - R(x) = f_1(x) - \frac{Mg}{c_l} \left(1 - e^{-\frac{c_l}{ES} x} \right); \\ \left. \frac{\partial \tilde{u}(x, t)}{\partial t} \right|_{t=0} = f_2(x). \end{aligned} \right\} \quad (1.49)$$

Коэффициенты D_{1k} и D_{2k} найдем из начальных условий (1.49), приравняв им $\tilde{u}(x, 0)$ и $\frac{\partial \tilde{u}(x, 0)}{\partial t}$ и используя ряд (1.48):

$$D_{1k} = \frac{2 C_{2k}}{l} \int_0^l \left[f_1(x) - \frac{Mg}{c_l} \left(1 - e^{-\frac{c_l}{ES} x} \right) \right] \sin \lambda_k x dx;$$

$$D_{2k} = \sqrt{\frac{\rho}{E}} \frac{2C_{2k}}{l} \int_0^l f_2(x) \sin \lambda_k x dx.$$

Возвращаясь к неоднородному уравнению, находим парциальные возбуждения для каждой гармоники, представляя временную часть решения волнового уравнения следующим образом: $\ddot{T}_k + \lambda_k^2 \frac{E}{\rho} \tilde{T}_k = \tilde{F}_k(t)$.

Подставляя ряд (1.48) в (1.42), имеем $\tilde{F}(x, t) = \sum_{k=1}^{\infty} C_{2k} \sin \lambda_k x (\ddot{T}_k + \lambda_k^2 \frac{E}{\rho} \tilde{T}_k)$, а поскольку функция $\tilde{F}$ может быть разложена в ряд Фурье на интервале $0 \leq x \leq l$, $\tilde{F}(x, t) = \sum_{k=1}^{\infty} \tilde{F}_k(t) \sin \lambda_k x$, и, таким образом, найдем парциальные силы

$$\tilde{F}_k(t) = \frac{2C_{2k}}{\rho l} \int_0^l \tilde{F}(x, t) \sin \lambda_k x dx. \quad (1.50)$$

Если предположить, что сокращения саркомеров создают усилия, равномерно распределенные по длине мышцы, т. е. $F(x, t) = F(t)$, то в (1.50) берется интеграл и

$$\tilde{F}_k(t) = -\frac{2C_{2k}}{\rho l} \left[\frac{1 - \cos \lambda_k l}{\lambda_k} F(t) - \frac{Mg c_l}{ES^2} \frac{\lambda_k - \left(\frac{c_l}{ES} \sin \lambda_k l + \lambda_k \cos \lambda_k l \right) e^{-\frac{c_l l}{ES}}}{\left(\frac{c_l}{ES} \right)^2 + \lambda_k^2} \right].$$

Теперь можно представить решение уравнения (1.42) в виде суммы свободных колебаний $\tilde{T}_{k \text{ св}}(t)$, определяемых начальными условиями, и вынужденных колебаний $\tilde{T}_{k \text{ вын}}(t)$, найденных из решений (1.44). С учетом вязкого трения получим $\tilde{u}(x, t) = \sum_{k=1}^{\infty} C_{2k} \sin \lambda_k x [\tilde{T}_{k \text{ св}}(t) + \tilde{T}_{k \text{ вын}}(t)] e^{-\eta t}$, где $\tilde{T}_{k \text{ св}}$ находим из (1.37), а

$$\tilde{T}_{k \text{ вын}}(t) = \frac{1}{\omega_k} \int_0^t e^{\eta \xi} \tilde{F}_k(\xi) \sin \omega_k (t - \xi) d\xi, \quad (1.51)$$

где $\eta = r/2\rho$. Возвращаясь к старой переменной, находим $u(x, t) = \tilde{u}(x, t) + \frac{Mg}{c_l} \left(1 - e^{-\frac{c_l}{ES} x} \right)$.

При анализе движений управляемых систем часто пользуются исследованием их переходных процессов, которые определяются

как отклик системы на возбуждающую функцию в виде ступенчатого воздействия:

$$F(t) = \begin{cases} 0, & t < 0; \\ A = \text{const}, & t \geq 0 \end{cases}$$

при нулевых начальных условиях. В этом случае интеграл в (1.51) берется и при нулевых начальных условиях переходные процессы полностью определяются (1.51).

Обращаясь вновь к оценке собственных чисел по (1.46), видим, что наличие упругого сопротивления увеличивает их значения и, следовательно, ускоряет переходные процессы, хотя и требует, естественно, больших усилий для достижения заданного положения управляемого органа.

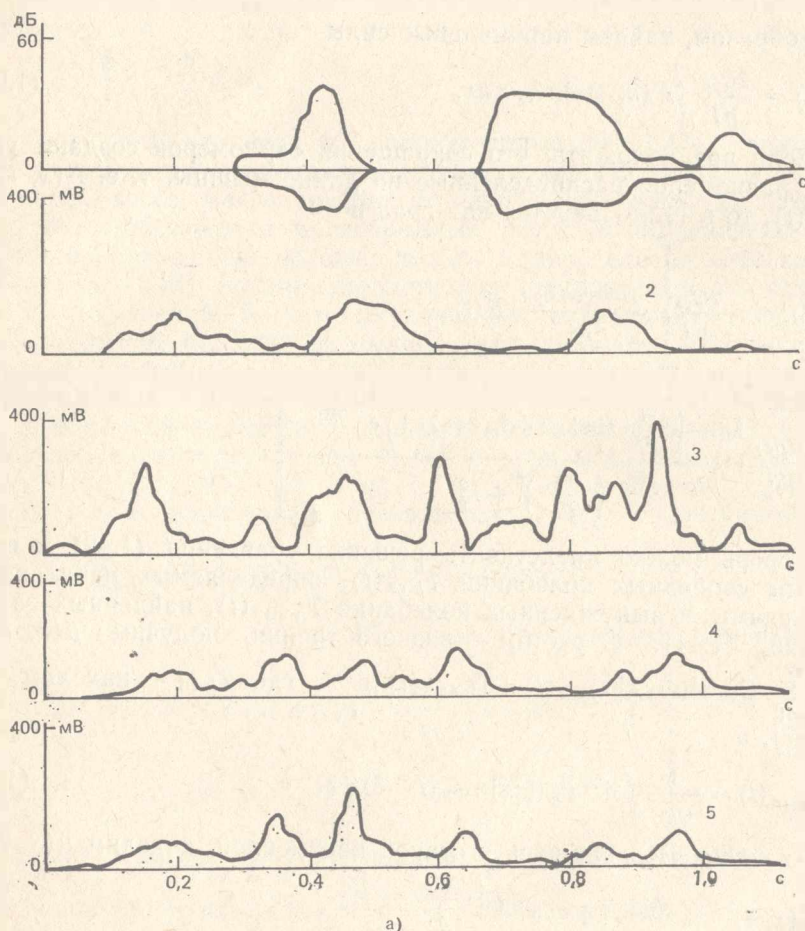
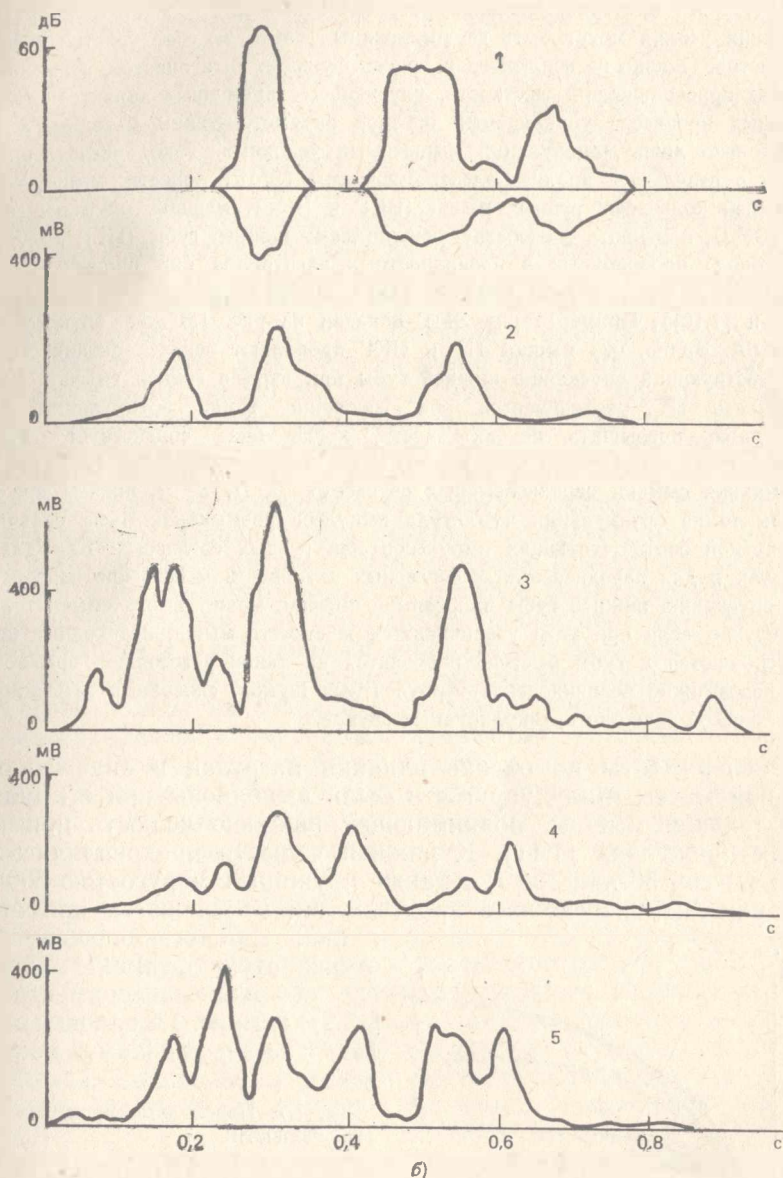


Рис. 1.6. Звукосочетание БАПАБА:
а — норма; б — ускоренный темп:

В § 1.1 обсуждалось явление изменения коэффициента жесткости мышц (как синергистов, так и антагонистов) в широких пределах под влиянием импульсации, направляемой из центральной нервной системы. В [1.13] было показано, что совместная активизация агонистов и антагонистов улучшает линейность статической характеристики и может смещать ее параллельно самой себе.



1 — огибающая речевого сигнала; 2 — верхняя ветвь кольцевой мышцы губ (ЭМГ); 3 — нижняя ветвь кольцевой мышцы губ (ЭМГ); 4 — мышца, опускающая угол губы (ЭМГ); 5 — подбородочная мышца (ЭМГ)

Электрмиографические измерения процессов артикуляции в ряде случаев обнаруживают такую электрическую активность мышц-антагонистов на начальных участках движения какого-либо органа, которая как будто направлена на торможение этого движения. Возможно, что такая активность мышц-антагонистов необходима для создания управляющего воздействия, оптимального в смысле некоторого критерия, поскольку отрицательные для данного направления движения усилия могут быть сформированы только за счет участия мышц-антагонистов. Однако не исключена и другая функция антагонистов, а именно увеличение присоединенной жесткости, улучшающее временные характеристики переходных процессов управляемых органов параметрическим методом, а не за счет только лишь манипуляций мышечными усилиями. Такое впечатление создается в результате анализа электромиограмм (ЭМГ) лицевых мышц, верхней и нижней кольцевых губных мышц (ВКГ и НКГ), мышцы, опускающей угол рта (ОУР), и мышцы, способствующей подъему нижней губы (ПГ), полученных с помощью проволочных и поверхностных электродов при произнесении бессмысленных звукосочетаний (об экспериментальной методике этих измерений см. § 5.7 и [1.10]). Пример таких ЭМГ показан на рис. 1.6 для звукосочетания БАПАБА. Видно, что мышцы ПГ и НКГ проявляют всплеск активности, противодействующий опусканию нижней губы при взрыве губной смычки. Всплеск активности ПГ увеличивается при ускорении темпа артикуляции и может даже превышать ее активность, когда она содействует подъему нижней губы при смычке. Обнаружена была также активность НКГ на фазе размыкания смычки переднеязычных согласных Д, Н, З, Л, причем при ускоренном темпе артикуляции амплитуда импульса размыкания была сравнима с амплитудой соответствующих импульсов для губных согласных. Как для губных, так и для переднеязычных согласных активность мышц, противодействующих опусканию нижней губы при взрыве смычки, может иметь смысл только в том случае, если при этом увеличивается жесткость мышц, переходные процессы ускоряются и губы быстрее расходятся на такое расстояние, при котором уже невозможно возникновение турбулентных шумов, способных замаскировать фонетическое качество артикулируемых звуков.

Эксперименты по оценке влияния нагрузки в виде дополнительной массы и/или упругости были выполнены при стимуляции жевательной мышцы, поднимающей нижнюю челюсть при речевых движениях [1.11]. К нижней челюсти прикреплялись грузики массой 50 или 100 г, а также пружина с упругостью 500 г/см и измерялись переходные процессы нижней челюсти при стимуляции, представляющей собой

ступенчатую функцию (подробнее об этих экспериментах см. § 3.2). На рис. 1.7 сплошными линиями с кружочками показаны

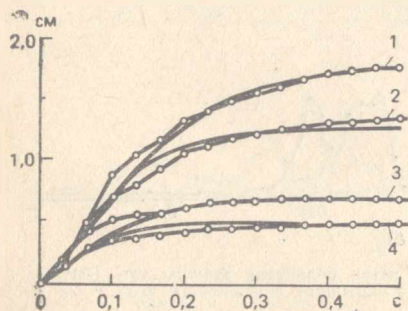


Рис. 1.7. Переходные процессы нижней челюсти:

— результаты измерений;
—○— результаты расчетов

1 — без нагрузки; 2 — нагрузка 50 г; 3 — пружина (жесткость 500 г/см); 4 — пружина и нагрузка 50 г

измеренные переходные процессы. Видно, что добавление внешней нагрузки в виде пружины уменьшает коэффициент усиления системы «мышца — нижняя челюсть».

Вычисленные переходные процессы показаны на рис. 1.7 непрерывными линиями. При расчетах принималось, что длина мышцы $l=7,5$ см, площадь ее поперечного сечения $S=2$ см², масса m — около 15 г, начальное значение модуля упругости без нагрузки $E_0=2 \cdot 10^2$ Па, а начальное значение коэффициента вязкого трения $\eta=8,5$ с⁻¹. Эти данные были получены частично путем обмера испытуемого, а частично из источников [1.6, 1.26, 1.30], затем они были уточнены для данного эксперимента, чтобы измеренные и вычисленные переходные процессы нижней челюсти при отсутствии внешней нагрузки совпадали наилучшим образом. Среднестатистическая масса нижней челюсти примерно равна 200 г, а упругость присоединенных тканей, идентифицированная в [1.11], принималась равной $c_l=315$ г/см. Поскольку отношение масс мышцы и нижней челюсти даже без нагрузки очень мало ($\approx 0,075$), поиск собственных чисел по (1.46) гарантирует хорошую точность. Эта мышца рассчитана на развитие больших усилий при пережевывании пищи, поэтому внешняя нагрузка массой 50 или 100 г соответствует нелинейному участку на характеристике «напряжение — удлинение», где модуль упругости и вязкость примерно пропорциональны развиваемому усилию. В соответствии с этим для каждого варианта нагрузки при расчетах изменялись модуль упругости и коэффициент вязкости и затем осуществлялась небольшая корректировка для достижения наилучшего сходства между измеренными и вычисленными переходными процессами. Сопоставляя вычисленные и измеренные переходные процессы, нетрудно заметить полное их совпадение в конце при отсутствии внешней нагрузки, близость на самых начальных участках и некоторое отличие, которое может быть объяснено параметрическими явлениями, такими как изменение модуля упругости и коэффициента вязкости в течение переходных процессов. В этих расчетах принималось, что начальные условия вносят свой вклад в переходные процессы при расслабленной мышце, т. е. при $E=E_0$, тогда как с момента начала переходного процесса модуль упругости скачком увеличивается до E_{max} . Сила сокращения мышц, развиваемая под воздействием стимуляции, была определена равной 1,5 Н, что вытекает из предположения преимущественного влияния упругости присоединенных тканей и оценки степени сокращения мышцы примерно на 0,5 см. (Амплитуда перемещения нижней челюсти составляла 1,8 см, но, учитывая место прикрепления жевательной мышцы, получаем амплитуду мышечного сокращения примерно 1 см.)

В экспериментах по стимуляции жевательной мышцы испытуемый максимально расслаблял все мышцы, управляющие положением нижней челюсти, и следил за сохранением расслабленности во время эксперимента, так что влияние мышц-антагонистов отсутствовало или было очень мало, т. е. $F_i(t) \approx 0$. В общем же

случае антагонисты могут проявлять активность, изменяя характеристики движения управляемого органа.

Рассмотрим роль внешних усилий, принимая для простоты, что внешние нагрузки в виде массы и упругости отсутствуют. Тогда граничные условия следующие: $u(0, t) = 0$; $ES \frac{\partial u(x, t)}{\partial x} \Big|_{x=l} = F_1(t)$.

Для решения краевой задачи необходимо избавиться от неоднородности в граничных условиях, для чего снова введем вспомогательную функцию $R(x, t)$ и заменим переменные: $\tilde{u}(x, t) = u(x, t) - R(x, t)$. Определим вспомогательную функцию как $R(x, t) = F_1(t) \frac{x^2}{2lES}$. Тогда граничные условия для новой переменной

$$\tilde{u}(x, t) \Big|_{x=0} = 0, \quad \frac{\partial \tilde{u}(x, t)}{\partial x} \Big|_{x=l} = 0$$

Одновременно с этим изменяются и начальные условия:

$$\tilde{u}(x, 0) = u(x, 0) - R(x, 0) = f_1(x) - F_1(0) \frac{x^2}{2lES};$$

$$\begin{aligned} \frac{\partial \tilde{u}(x, t)}{\partial t} \Big|_{t=0} &= \frac{\partial u(x, t)}{\partial t} \Big|_{t=0} - \frac{\partial R(x, t)}{\partial t} \Big|_{t=0} = \\ &= f_2(x) - \frac{\partial F_1(t)}{\partial t} \Big|_{t=0} \frac{x^2}{2lES}, \end{aligned}$$

а уравнение относительно новой переменной теперь выглядит так:

$$E \frac{\partial^2 \tilde{u}}{\partial x^2} - \rho \frac{\partial^2 \tilde{u}}{\partial t^2} = -\tilde{F}(x, t),$$

где $\tilde{F}(x, t)$ — новая возбуждающая функция:

$$\tilde{F}(x, t) = F(x, t) - \frac{F_1(t)}{lS} - \frac{\rho x^2}{2lES} \frac{\partial^2 F_1(t)}{\partial t^2}. \quad (1.52)$$

В (1.52) важным элементом является появление второй производной по времени от внешней силы $F_1(t)$, которая подчеркивает моменты изменения скорости нарастания или спада этой силы. В результате мышцы-антагонисты приобретают возможность влияния на характеристики движений управляемого органа не только непосредственно за счет развития силы $F_1(t)$, но и за счет ее изменений, что может оказаться более выгодным энергетически и для обеспечения точности процессов управления. Эффект влияния второй производной по времени от силы, приложенной на конце мышцы, проявляется при сопоставлении поведения реальной мышцы и волновой модели при растяжении или отпускании с постоянной скоростью [1.6]: оказалось, что существует большое сходство в поведении модели и реальной мышцы, причем характерной особенностью являются скачки усилий на конце $x=l$, соответствующие моментам изменения скорости изменения приложенного усилия.

В [1.6] рассматривались также типичные для исследования мышц ситуации, такие как изменение напряжения на конце мыш-

цы при ее быстром укорочении, зависимость скорости укорочения мышцы от величины подтягиваемого груза (уравнение Хилла) и быстрое начальное укорочение мышцы при ее нагрузке. Во всех этих случаях было обнаружено близкое сходство в поведении модели и реального объекта и присутствовали качественные эффекты, играющие большую роль в деятельности реальных мышц. Таким образом, очевидно, что даже упрощенная одномерная волновая модель оказывается полезным инструментом в исследованиях мышечного сокращения.

Тот факт, что мышца представляет собой систему с распределенными параметрами, создает принципиально новую возможность управления динамикой ее сокращения со стороны центральной системы. Эта возможность состоит в изменении соотношения между амплитудами гармоник посредством соответствующего выбора участка мышцы, который подвергается сокращению. Пусть сократительные усилия не распределены равномерно по длине мышцы, как это принималось ранее, а сосредоточены в одной точке с координатой x_0 . Математически это условие может быть представлено с помощью δ -функции: $F(x_0, t) = \delta(x - x_0)F(t)$. Подставляя эту силу в (1.50), находим парциальные возбуждающие функции $F_h(t) = \frac{2C_{2h}}{\rho l} F(t) \sin \lambda_h x_0$.

Теперь видно, что, изменяя координату приложения сосредоточенного усилия x_0 , можно изменить и амплитуду соответствующей гармоники в колебаниях мышцы. Например, за счет соответствующего выбора можно сделать амплитуду некоторой гармоники равной 0 или ± 1 (поскольку $-1 \leq \sin \lambda x \leq 1$). Таким образом, центральная нервная система может управлять динамическими характеристиками мышечного сокращения за счет простого перераспределения возбуждения по длине мышцы. Пока еще нет экспериментальных данных о том, реализуется ли эта возможность на самом деле, однако известны примеры, когда наличие возбуждения, приложенного к отдельному участку мышцы, не подлежит сомнению. В частности, изгиб вверх или вниз кончика языка, очевидно, осуществляется за счет сокращения мышечных волокон верхней и нижней продольных мышц языка, находящихся лишь вблизи кончика языка, а остальные участки этих мышц, распространяющихся вплоть до корня языка, остаются невозбужденными.

1.4. ПАРАМЕТРИЧЕСКИЕ ЯВЛЕНИЯ

Обсуждая в § 1.1 зависимость укорочения мышцы от напряжения, мы установили, что жесткость мышцы есть некоторая функция от развиваемого ею напряжения. Причина такой зависимости лежит в изменении структурных связей в мышце при сдвиге тонких нитей саркомера относительно толстых. Однако и при фиксированном уровне напряжения жесткость мышцы устанавливается не сразу, а имеет некоторый переходный процесс. В основе этого переходного процесса лежат релаксационные явле-

ния, которые представляют собой отклик молекулярной структуры мышечных тканей на нарушение состояния внутреннего равновесия, возникающее в результате локальной деформации тканей. В простейшем виде релаксационные процессы описываются обыкновенным дифференциальным уравнением первого порядка [1.9]:

$$\tau_{\varepsilon} \frac{d\varepsilon}{dt} + \varepsilon = \kappa_{\varepsilon} \sigma(t), \quad (1.53)$$

где ε — линейная деформация; κ_{ε} — коэффициент усиления (точнее говоря, ослабления); σ — локальное напряжение. Если, например, локальное напряжение изменяется скачком от 0 до σ_0 , то решение (1.53) есть

$$\varepsilon(t) = \kappa_{\varepsilon} \sigma_0 \left(1 - e^{-\frac{t}{\tau_{\varepsilon}}} \right). \quad (1.54)$$

Некоторые материалы, например газы, обладают лишь одним типом релаксации, тогда как другие имеют несколько типов, а материалы, подобные резине, обнаруживают непрерывный спектр времен релаксации от очень больших до очень малых.

Анализ системы уравнений, описывающих развитие напряжения в кинетической модели Дещеревского, также указывает на существование переходных процессов с конечным временем в механизме сокращения саркомеров. Все эти явления можно формально учесть путем записи модуля упругости мышцы как функции времени. Тогда уравнение мышечного сокращения принимает вид

$$\frac{E_0 + E_1(t)}{\rho} \frac{\partial^2 u}{\partial x^2} + \frac{\partial^2 u}{\partial t^2} = \frac{F(x, t)}{\rho}, \quad (1.55)$$

где E_0 — некоторая начальная жесткость тканей до появления локальных деформаций. Установившееся значение $E_{1\infty}$ примерно пропорционально усилию F_{∞} , развиваемому мышцей, однако зависимость $E_1(t)$ иная, чем зависимость F от времени.

Пусть $E_{1\infty}$ — некоторое установившееся значение переменной компоненты модуля упругости для некоторого постоянного значения усилия F_{∞} . Тогда модуль упругости изменяется во времени таким образом, что в любой момент справедливо неравенство $E_0 \leq E(t) \leq E_0 + E_{1\infty}$, которое означает, что любой переходный процесс мышечного сокращения происходит при меньшем значении модуля упругости на начальном этапе и при большом значении E на конечном.

Из (1.54) видно, что нарастание модуля упругости происходит быстрее на начальных участках переходного процесса и медленнее на конечных. Это означает, что для времен релаксации τ_{ε} , сравнимых со временем переходных процессов мышечного сокращения $\tau_{п.п.}$, жесткость мышцы будет близка к ее установившемуся значению $E_0 + E_{1\infty}$ на всем протяжении ее переходного процесса, за исключением начального участка. Если же время релаксации больше времени переходного процесса, то и сам переходный про-

цесс будет затянут вследствие продолжающегося изменения жесткости мышцы.

Если мышца обладает несколькими типами релаксаций, причем их постоянные времени сильно различаются, то при быстром исчезновении внутримышечного напряжения модуль упругости E не может скачком вернуться к начальному состоянию E_0 — для этого потребуется некоторое время, в течение которого мышца будет обладать повышенной по сравнению с исходным расслабленным состоянием жесткостью. Подобное запаздывание в уменьшении жесткости мышцы после снятия напряжения было обнаружено в экспериментах по исследованию движений мягкого нёба. В этих экспериментах отводились электрические потенциалы от правой и левой ветвей мышцы, поднимающей небную занавеску (ПНЗ), и язычно-небной мышцы (ЯН) [1.27] (подробнее об этих экспериментах см. в § 3.2). Мышца ЯН не проявила сколько-нибудь заметной электрической активности ни при произвольных движениях небной занавески во время вдоха, выдоха или дутья через рот, ни во время артикуляции отдельных звуков или звукосочетаний. Из этого следует, что небная занавеска поднимается за счет активного сокращения ПНЗ, тогда как опускается чисто пассивно — за счет упругости своих тканей. В результате можно было бы ожидать, что крутизна переходных процессов небной занавески при опускании должна быть не больше крутизны переходных процессов при опускании, поскольку напряженному состоянию ПНЗ соответствует и большая жесткость, повышающая согласно (1.46) собственные частоты системы «мышца — управляемый объект». Такое соотношение действительно имеет место, если время, в течение которого небная занавеска остается поднятой, сравнительно мало (порядка 0,1—0,2 с), как это обычно бывает при артикуляции (рис. 1.8). Если же это время порядка 1—2 с, то небная

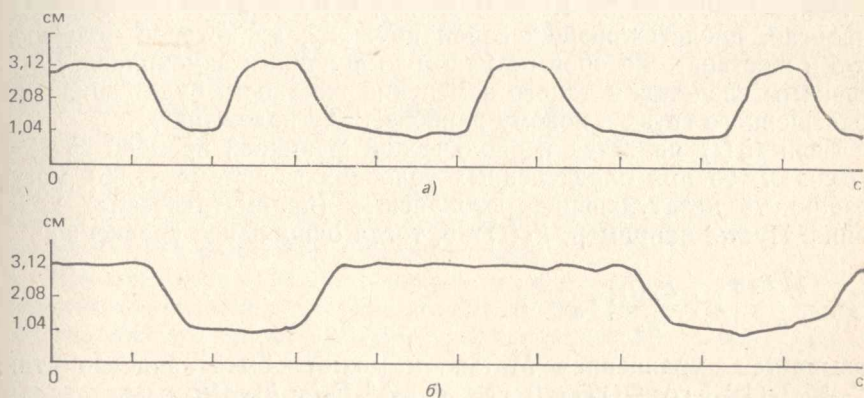


Рис. 1.8. Вертикальные смещения (усл. ед.) небной занавески в последовательности:

а — «ПИ — вдох — ПИ ...»; б — «И — вдох — И ...».

Отсчеты времени — через 0,43 с

занавеска иногда опускается даже быстрее (по крайней мере, на начальных этапах), чем поднимается. Это явление может быть объяснено гистерезисом жесткости, связанным с релаксационными процессами: переходные процессы при движении небной занавески вверх заканчиваются до того, как жесткость мышцы достигнет установившегося значения. Если мягкое нёбо находится в поднятом положении достаточно долго, то жесткость мышцы продолжает повышаться и сохраняется таковой еще в течение некоторого времени после прекращения импульсации ПНЗ, начиная затем медленно понижаться. При таких условиях пассивное опускание небной занавески происходит при более высоком значении коэффициента жесткости, чем для расслабленной мышцы ПНЗ перед началом подъема нёбной занавески, и, следовательно, начальные этапы опускания будут протекать быстрее. Эффект затянувшегося изменения жесткости мышцы получил название «медленной компоненты» мышечного сокращения, причем длительность переходных процессов этой компоненты равна 1—15 с [1.2, 1.29].

Модуль упругости E есть функция только времени, поэтому в (1.55) можно применить метод разделения переменных, что дает для временной компоненты систему

$$T_k'' + \frac{\lambda_k^2}{\rho} [E_0 + E_1(t)] T_k = F_k(t). \quad (1.56)$$

Если положить $F_k(t) = 0$, то получившееся однородное уравнение носит в математике имя Хилла [не путать с уравнением Хилла (1.1)]. Анализ этого уравнения было уделено много внимания в силу распространенности процессов, описываемых этим уравнением. Основное его свойство состоит в том, что даже при отсутствии внешнего возбуждения $F(t)$ переменный коэффициент $E(t)$ создает так называемое параметрическое возбуждение. Параметрическое возбуждение мышцы может быть пояснено с помощью аналогии, представляющей собой пружину, на которой подвешен груз и жесткость которой по каким-то причинам изменилась. Естественным следствием такого изменения жесткости будет движение подвешенного груза к новому равновесному положению.

Если $E_1(t)$ является периодической функцией времени $E_1(t) = a \cos \Omega t$, то при определенных условиях в системе (1.56) могут возникнуть незатухающие резонансные (параметрические) колебания. Пусть, например, $F_k(t) = 0$, тогда однородное уравнение

$$T_k'' + \frac{\lambda_k^2 E_0}{\rho} \left[1 + \frac{E_1(t)}{E_0} \right] T_k = 0$$

называется уравнением Матье и может быть записано так: $T_k'' + \omega_k^2 (1 + h \cos \Omega t) T_k = 0$, где $\omega_k^2 = \lambda_k^2 E_0 / \rho$; $h = a / E_0$.

В [1.1] было показано, что параметрические резонансы возникают при отношении частот системы ω и возбуждения Ω : $\omega / \Omega = 1/2, 1, 3/2, \dots$. Если в системе присутствует демпфирование, то демпфирующий коэффициент $\delta_k = 2\eta \lambda_k^2$ определяет глубину моду-

лации h , при которой возникает резонанс, но наличие демпфирующего члена, в отличие от линейных непараметрических систем, не приводит к затуханию колебаний. Например, для первого резонанса должны выполняться условия

$$1 - \sqrt{\frac{h^2}{4} - \frac{4\delta_1^2}{\Omega^2}} < \left(\frac{2\omega_1}{\Omega}\right)^2 < 1 + \sqrt{\frac{h^2}{4} - \frac{4\delta_1^2}{\Omega^2}}; \quad h > 4\frac{\delta_1}{\Omega}.$$

Для второго резонанса условием вхождения в резонанс является $h > 2\sqrt{2\delta_2/\Omega}$, т. е. при $\delta_2 > \Omega$ требуется большая глубина модуляции, чем для первого резонанса. Таким образом, наиболее вероятно возникновение резонансных колебаний при $\Omega = 2\omega$. Стационарные автоколебания при параметрическом возбуждении устанавливаются лишь при наличии нелинейности. В мышцах такая нелинейность присутствует при напряжениях, близких к максимальным.

Принципиальная возможность возникновения параметрических резонансных колебаний в мышце при периодическом возбуждении может оказаться весьма важным элементом анализа чрезвычайно сложных процессов колебаний голосовых складок.

В 1950 г. появилась работа Юссона [1.23], в которой сообщалось, что в экспериментах по стимуляции голосовых мышц собаки наблюдались синхронные с электрическим возбуждением колебания голосовых складок. Это и последующие наблюдения Юссона легли в основу так называемой нейрохронаксической теории, согласно которой колебательные процессы при голосообразовании у человека непосредственно управляются центральной нервной системой путем задания вынуждающих нейромоторных команд, следующих с той же частотой, что и частота основного тона. Теория Юссона подверглась серьезной критике в отношении как экспериментальной методики, так и выводов о полностью управляемом процессе колебаний голосовых связок. Например, в [1.17] указывается, что Юссон мог измерять не первичные импульсы электрической активности мышц голосовых складок у человека, а вторичную электризацию, или так называемый микрофонный эффект, вызванный механическими колебаниями голосовых складок. Иными словами, Юссон мог принять следствие за причину. Основной довод против теории Юссона состоит в том, что, как доказано физическими моделями и математическим анализом, для объяснения механизма колебаний голосовых складок вполне достаточно учесть эффект Бернулли, возникающий в потоках воздуха и приводящий к падению давления в голосовой щели, за счет чего и поддерживаются автоколебания голосовых связок.

Признавая критику теории Юссона в целом, необходимо отметить, что параметрические явления могут сосуществовать с эффектом Бернулли, не заменяя его. Известно, что при периодическом изменении натяжения тонкой струны возникают ее поперечные колебания, частоты которых в точности определяются параметрическими резонансами. Голосовые складки имеют конечные размеры во всех трех измерениях, и продольные колебания в них должны сопровождаться возникновением поперечных колебаний, частота которых согласно (1.26) и (1.27) значительно ниже частоты продольных колебаний. Это явление повышает вероятность параметрического возбуждения поперечных колебаний голосовых скла-

док, поскольку при этом снижаются требования на частоту импульсации, управляющей переменной компонентой натяжения голосовых складок.

Эксперименты на модели голосового источника (см. § 2.2) показали, что небольшие изменения силы натяжения складок, имитирующие физиологический тремор мышц, влияют на восприятие синтетической речи. При этом оказалось, что при случайных отклонениях, не превышающих 1,5—2% исходной величины натяжения, синтетическая речь воспринимается как «тусклая», «машиноподобная», при отклонениях, превышающих 5%, голос оценивается как «грубый», «хриплый», тогда как диапазон отклонений порядка 2—4% придает голосу большую живость и натуральность. Этот эффект вполне согласуется с исследованиями на синтетической речи роли небольших случайных изменений амплитуды и периода следования импульсов голосового возбуждения, но попытка найти явную зависимость между этими параметрами не увенчалась успехом. Эта неудача объясняется тем, что в действительности эти изменения происходят согласованно и сопровождаются также вариацией формы импульсов, связанной с изменением амплитудных соотношений гармоник упругих колебаний голосовых складок.

Параметрические явления в сокращении мышц, таким образом, проявляются в процессах артикуляции и голосообразования и могут играть важную роль в восприятии речи.

ГОЛОСОВОЙ ИСТОЧНИК

2.1. СТРОЕНИЕ И КИНЕМАТИКА ГОРТАНИ

Человеческая речь с давних пор вызывала интерес как уникальное явление, а древние приписывали голосу магические свойства.

Гортань человека и протекающие в ней механические и аэродинамические процессы являются, пожалуй, наиболее сложными в речеобразующем аппарате. Гортань расположена на верхнем конце трахеи и образована несколькими хрящами и мышцами, которые показаны на рис. 2.1—2.3 (по [2.28, 2.58, 2.98]). Перстневидный хрящ находится над последним кольцом трахеи и связан со щитовидным и черпаловидным хрящами соединительными тканями и мышцами: перстне-щитовидной, задней перстне-черпаловидной и боковыми перстне-черпаловидными. В число наружных мышц гортани входит также щитовидно-подъязычная, соединяющая щитовидный хрящ с подъязычной косточкой, а также щитовидно-грудная и подъязычно-грудная. Мышцы гортани относятся к числу тонических, которые способны выдерживать длительное напряжение без усталости и имеют весьма короткое время переходных процессов. Двигательные единицы этих мышц содержат по 100—300 мышечных волокон, что обеспечивает высокую точность управления их сокращением [2.73].

Голосовые складки образованы связками, двумя мышцами: голосовой и черпало-щитовидной — и соединительными тканями (рис. 2.4, 2.5 по [2.15, 2.17, 2.108]). Эти мышцы и связки своими передними концами неподвижно прикреплены к щитовидному хрящу, а их задние концы соединены с черпаловидными хрящами и могут двигаться вместе с ними. Над голосовыми складками находятся так называемые ложные голосовые складки (см. рис. 2.4). Они не содержат внутренних мышц, но расстояние между ними может изменяться за счет сокращения некоторых наружных мышц гортани, причем площадь поперечного сечения нижнего отдела речевого тракта (фаринкса) может изменяться более чем в 20 раз — от 0,2 до 6,8 см² [2.68]. Эти изменения зависят от фонологического окружения данного звука и скорости артикуляции [2.85, 2.86]. В шепотной речи ложные складки сближаются, препятствуя фонации за счет повышения надсвязочного давления и способствуя турбулизации воздушного потока, приводящей к шумообразованию, тогда как голосовая щель остается открытой [2.34]. В случае удаления истинных голосовых складок ложные складки иногда принимают на себя функции голосообразования

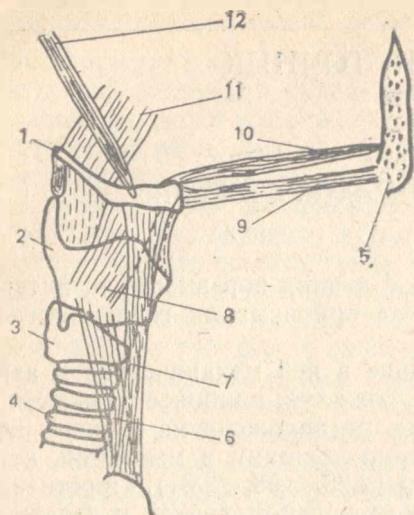


Рис. 2.1. Строение гортани и наружные мышцы — вид сбоку:

1 — подъязычная кость; 2 — щитовидный хрящ; 3 — перстневидный хрящ; 4 — трахея; 5 — нижняя челюсть; 6 — щитовидно-грудная мышца; 7 — подъязычно-грудная мышца; 8 — щитовидно-подъязычная мышца; 9 — подбородочно-подъязычная мышца; 10 — передняя двубрюшная мышца; 11 — язычно-подъязычная мышца; 12 — задняя двубрюшная мышца

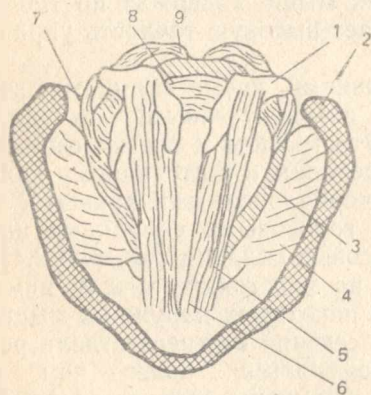


Рис. 2.3. Разрез гортани — вид сверху:

1 — черпаловидный хрящ; 2 — щитовидный хрящ; 3 — перстневидный хрящ; 4 — перстне-щитовидная мышца; 5 — черпало-щитовидная мышца; 6 — голосовая мышца; 7 — боковая перстне-черпаловидная мышца; 8 — внутренняя черпаловидная мышца; 9 — задняя перстне-черпаловидная мышца

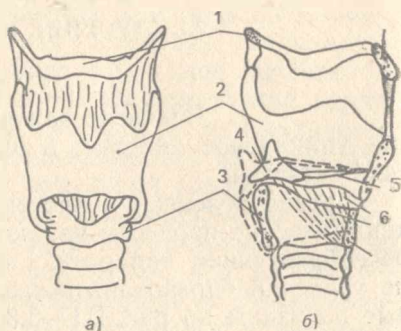


Рис. 2.2. Строение гортани: а — вид спереди; б — вид сбоку:

1 — подъязычная кость; 2 — щитовидный хрящ; 3 — перстневидный хрящ; 4 — черпаловидный хрящ; 5 — голосовые связки; 6 — перстне-щитовидная мышца

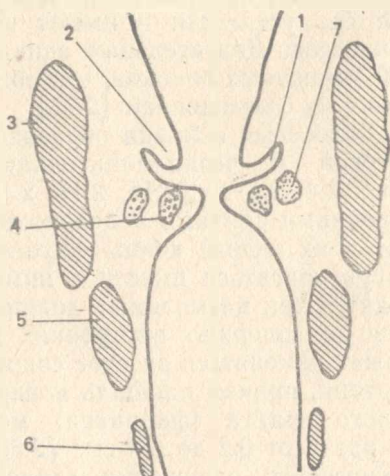


Рис. 2.4. Разрез гортани — вид спереди:

1 — морганиев желудочек; 2 — ложные голосовые складки; 3 — щитовидный хрящ; 4 — голосовая мышца; 5 — перстневидный хрящ; 6 — хрящ трахеи; 7 — черпало-щитовидная мышца

[2.23]. Между истинными и ложными голосовыми складками находится некоторое расширение, называемое морганиевым желушком.

Черпаловидные хрящи связаны с перстневидным задними и боковыми перстне-черпаловидными мышцами и внутренними черпаловидными мышцами. Черпаловидные хрящи могут поворачиваться в латеральной (горизонтальной) плоскости, образуя раз-

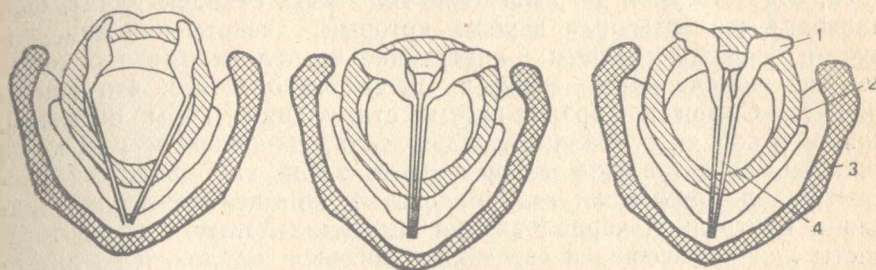


Рис. 2.5. Положения голосовых связок — вид сверху:

1 — черпаловидный хрящ; 2 — перстневидный хрящ; 3 — щитовидный хрящ; 4 — голосовые связки

личные формы голосовой щели (см. рис. 2.5). Сокращение перстне-щитовидной мышцы приводит к смещению черпаловидных хрящей в заднем направлении (см. рис. 2.2) и натяжению голосовых складок [2.94, 2.98]. К такому же результату приводит подъем перстневидного хряща за счет сокращения подъязычно-язычной мышцы [2.92] или же за счет подъема подъязычной косточки. Эти два последние вида движений вызывают увеличение натяжения голосовых складок и соответственно повышение частоты основного тона, связанное не с управлением мышцами гортани, а с артикуляцией различных звуков. Известно, в частности, что для так называемых высоких гласных, характеризующихся подъемом языка, частота основного тона F_0 выше, чем для низких гласных [2.66, 2.72, 2.80], причем это повышение F_0 происходит, несмотря на увеличение внутриротового давления за счет возрастания сопротивления воздушному потоку в речевом тракте [2.36]. Геометрические параметры внутренних вращающих моментов для перстневидного хряща приводятся в [2.98]. В частности, перстне-щитовидная мышца оказывается наиболее длинной (2,95 см для мужчин и 2,26 см для женщин) и толстой ($0,112 \text{ см}^2$).

Исследование процессов фонации с помощью волоконных световодов показывает, что при фонации гортань примерно на 1 см выше того положения, которое она занимает во время дыхания [2.35]. Иногда отмечается подъем гортани и при артикуляции глухих взрывных. Сокращение мышцы, опускающей подъязычную косточку, вызывает ослабление натяжения голосовых складок и соответственно падение частоты основного тона [2.30].

Между высотой гортани и частотой основного тона существует положительная корреляция, причем подъем гортани на 1 мм

приводит к повышению F_0 на 8—10 Гц [2.59, 2.77, 2.93]. На рис. 2.6 показана измеренная с помощью ультразвукового датчика зависимость частоты основного тона от высоты гортани [2.37]. Расстояние между голосовыми складками определяет величину голосовой щели. При дыхании связки полностью разведены, причем при выдохе размер щели несколько уменьшается. Задняя перстне-черпаловидная мышца является единственной мышцей в гортани, ответственной за разведение голосовых складок [2.42]. Она иннервируется язычным нервом, который одновременно иннервирует и сводящие мышцы — внутренние черпаловидные и боковые перстне-черпаловидные мышцы. Интересно отметить, что еще в 1890 г. Симон и Хорзли путем стимуляции коры обезьяны, кошки и собаки показали, что движения, сближающие голосовые складки, представлены в коре мозга, тогда как движения раскрытия голосовой щели — в подкорке. Задняя перстне-черпаловидная и внутренняя черпаловидная мышцы действуют как антагонисты для разведения и сведения голосовых складок, и их активность в норме попеременно во времени [2.88].

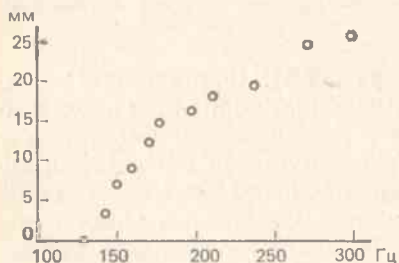


Рис. 2.6. Зависимость высоты гортани от частоты основного тона

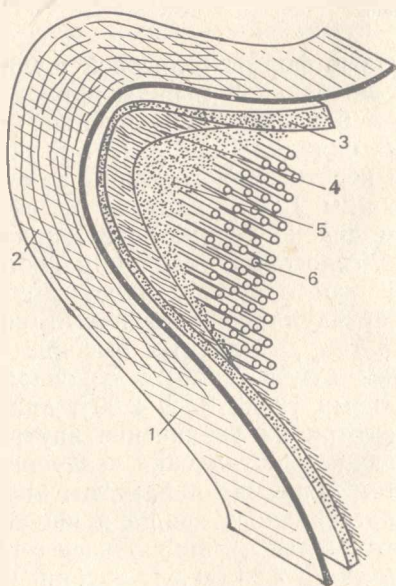


Рис. 2.7. Строение голосовых складок:
1 — реснитчатый эпителий; 2 — слоистый эпителий; 3 — поверхностный слой; 4 — промежуточный слой; 5 — глубокий слой; 6 — голосовая мышца

Амплитуда колебаний голосовых складок в латеральной плоскости достигает до 0,2 см в зависимости от регистра голоса, причем глухие взрывные согласные обычно артикулируются при постоянном раскрытии голосовой щели, близком к амплитуде колебаний складок при фонации [2.35, 2.98]. В некоторых языках, например в английском, артикуляция глухих взрывных может происходить в зависимости от фонетического окружения не только с открытой, но и с закрытой голосовой щелью во время

смычки, создавая аспиративные и неаспиративные взрывы, т. е. с придыханием и без него [2.89]. Степень раскрытия голосовой щели для глухих взрывных зависит также от положения согласного в слове — у двух из трех испытуемых в [2.87] ширина голосовой щели для начальных согласных была в 2—4 раза больше, чем для звуков, находящихся в середине и конце слова.

Амплитуда колебаний голосовых складок у звонких фрикативных такая же, как и у других звонких звуков, а ширина голосовой щели для глухих фрикативных примерно вдвое больше амплитуды колебаний складок при артикуляции звонких фрикативных [2.99]. По данным [2.71] размер голосовой щели для глухих согласных имеет две-три градации: самая большая площадь щели для глухих фрикативных, меньшая — для глухих аспиративных взрывных, наименьшая — для аспиративных. Время, за которое голосовые складки переходят из полностью разведенного состояния при свободном дыхании в исходное положение для фонации, составляет по [2.69] примерно 0,2 с, причем скорости сведения и разведения связок примерно одинаковы. В [2.95] приводится другая оценка — 0,08—0,15 с от начала смыкания голосовых складок до начала фонации.

Внутренние мышцы голосовых складок — голосовая и черпало-щитовидная — при своем сокращении изменяют жесткость, длину и форму складок. Изменение площади поперечного сечения складок при совместном воздействии внутренних мышц гортани составляет от —28 до +16%, а длина складок изменяется от —10 до 24% [1.22]. Уменьшение толщины голосовых складок при повышении частоты основного тона и удлинение их длины до 25% исходной величины наблюдались в [2.44—2.47].

В поперечном сечении голосовые складки представляют собой многослойную структуру, причем механические свойства каждого слоя сильно отличаются друг от друга [1.22, 2.40] и, по-видимому, их механические параметры по-разному реагируют на натяжение голосовых складок при сдвиге черпаловидных хрящей назад и при сокращении голосовой и черпало-щитовидной мышц. Поверхность складок покрыта эпителием толщиной около 0,05 мм, который играет роль капсулы, удерживающей форму складок (рис. 2.7 по [1.22]). Этот слой простирается в среднем на 1,8 мм в сторону и на 2,3 мм ниже верхнего края складок, почти перекрывая область контакта складок при схлопывании во время фонации. Под эпителием находятся три слоя, различающиеся содержанием волокон разного типа. Поверхностный слой толщиной около 0,3 мм довольно беден волокнами и подобен очень мягкому желатину, играющему роль амортизатора при схлопывании складок. Промежуточный слой состоит главным образом из упругих резиноподобных волокон. Глубокий слой содержит коллагеновые волокна, подобные хлопчатобумажным нитям. Общая толщина промежуточного и глубокого слоев составляет примерно 0,8 мм, причем они плавно переходят друг в друга. Коллагеновые волокна глубокого слоя

сплетаются с волокнами голосовой мышцы, которые идут почти параллельно краям голосовых складок. Можно условно разделить поперечное сечение голосовых складок на три слоя (по [1.22]): покрывающий, который включает в себя эпителий и поверхностный слой; переходный, состоящий из промежуточного и глубокого слоев, и тело складок, образованное мышечными волокнами голосовой и черпало-щитовидной мышц толщиной около 5 мм. Общая толщина голосовых складок 8—13 мм, а их высота 12—16 мм. Длина голосовых складок в ненапряженном состоянии составляет в среднем 1,5 см для мужчин и 1,1 см для женщин [2.11].

2.2. МЕХАНИЧЕСКИЕ СВОЙСТВА ТКАНЕЙ ГОЛОСОВЫХ СКЛАДОК

Геометрические размеры и механические параметры голосовых складок определяют частоту и форму их колебаний. Эти данные необходимы для построения количественных моделей, описывающих колебания голосовых складок. Имеется очень ограниченное число работ, посвященных измерениям параметров голосовых складок, причем значительная их часть выполнена на препарате удаленной гортани собаки, что создает проблему оценки степени их применимости к живым голосовым складкам человека. Тем не менее эти данные помогают разработать некоторую исходную систему, допускающую проверку и сопоставление с измерениями на человеческих тканях и, кроме того, по ряду параметров ткани собаки и человека оказываются весьма близкими. В одной из первых работ, посвященной измерениям параметров голосовых складок [2.54], их упругость при напряжении внутренних мышц оценивалась как $c \approx 8 \cdot 10^4$ дн/см, а декремент затухания — как величина порядка 0,1—0,2. Если принять, что связки в поперечном сечении имеют вид трапеции с основаниями 0,8 и 0,4 см, а высотой $h=0,6$ см, то по этим данным можно оценить и модуль упругости тканей складок: $E=ch/S \approx 10^4$ Па (здесь S — площадь поперечного сечения).

В [1.22, 2.40] приводятся результаты исследования голосовых складок собаки и человека с дифференциацией различных слоев. В этих исследованиях подвергаемый испытанию образец вырезался из голосовых складок путем параллельных сечений во фронтальной плоскости таким образом, что расстояние между сечениями составляло 3—5 мм, а расстояние между верхними и нижними краями складок было около 12—16 мм при ширине 8—13 мм. Эти образцы нагружались различными массами, и под микроскопом измерялось удлинение каждого слоя. Нагрузочные характеристики для этих слоев показаны на рис. 2.8 по [1.22], где по оси абсцисс отложено относительное удлинение $\Delta l/l_0$, а по оси ординат — нагрузка. Пользуясь этими характеристиками и учитывая геометрические размеры образцов, можно приблизительно рассчитать диапазон значений для модуля упругости различных слоев: $E=P/\epsilon S$,

где P — вес нагрузки; S — площадь слоя; ε — относительное удлинение.

Из-за особенностей формы образцов эти вычисления не очень точны, но тем не менее позволяют оценить E по порядку величины. При этом оказывается, что $E \approx (1 \div 6) \cdot 10^4$ Па, т. е. достаточно близко к значениям, полученным по данным [2.54]. Следует отметить, что в обоих случаях оценки упругости относятся к силе, действующей в плоскости, перпендикулярной к направлению мышечных волокон, и, как отмечалось в гл. 1, упругость тканей вдоль мышечных волокон может превышать даже в расслабленном состоянии упругость в поперечной плоскости на порядок.

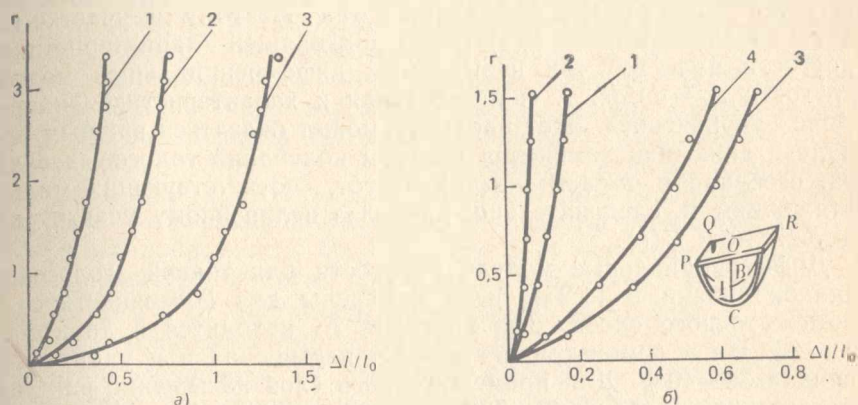


Рис. 2.8. Нагрузочные характеристики различных слоев голосовых складок: а — собаки ($PQ=3,5$ мм, $QR=12$ мм, $OB=5,33$ мм, $BC=3$ мм); б — человека ($PQ=5,2$ мм, $QR=16,3$ мм, $OB=10,74$ мм, $BI=1,67$ мм, $IC=0,69$ мм): 1 — общая; 2 — внутренний слой; 3 — поверхностный слой; 4 — переходный слой

Натяжение голосовых складок за счет сокращения перстне-щитовидной мышцы моделировалось в [1.22, 2.40] для препаратов голосовых складок собаки и человека. Было обнаружено повышение жесткости тканей: для человека примерно в 2 раза при удлинении складок на 3 мм, а для собаки примерно в 3—4 раза. Там же исследовалось сокращение голосовой мышцы собаки с помощью электрической стимуляции. При сравнительно невысоких уровнях стимуляции жесткость голосовых складок уменьшилась, очевидно, за счет расслабления поверхностных слоев. В дальнейшем, при увеличении интенсивности стимуляции, общая жесткость увеличивалась, причем в продольном направлении (вдоль мышечных волокон) относительное увеличение жесткости по сравнению с расслабленным состоянием было около 5 раз, а в поперечной плоскости жесткость увеличивалась незначительно.

Детальная оценка упругих характеристик в продольном направлении для различных слоев голосовых складок собаки была произведена в [1.26]. Измерялись модуль упругости, модуль сдви-

га и вязкость трех слоев — поверхностного (эпителия), промежуточного и глубокого (содержащего в основном мышечные волокна). В диапазоне малых относительных удлинений для эпителия модуль Юнга в среднем имеет величины порядка 10^5 — 10^6 Па, для промежуточного слоя — 10^4 — 10^5 Па, а упругость мышечных тканей находится в диапазоне $5 \cdot 10^3$ — $5 \cdot 10^4$ Па. Следует иметь в виду, что показатели модуля упругости для мышечных тканей относятся к расслабленному состоянию и должны быть увеличены в 5—10 раз для напряженной мышцы.

Однако модуль Юнга $E = \sigma/\epsilon$, определенный как отношение напряжения σ к относительному линейному удлинению ϵ , недостаточно полно характеризует упругие свойства сред, в которых не соблюдается закон Гука, т. е. E зависит от ϵ . Поэтому на участке нелинейной зависимости $E(\epsilon)$ более подходящей характеристикой упругости является так называемый дифференциальный модуль упругости $E_d = d\sigma/d\epsilon$, т. е. касательная к характеристике «напряжение—укорочение». Этот параметр может оказаться важным при анализе динамики изменения частоты колебаний голосовых складок, особенно в области низких частот, соответствующих малым натяжениям и, следовательно, наиболее нелинейному участку характеристики $E(\epsilon)$.

Дифференциальные модули упругости для тканей голосовых складок собаки в [1.26] были измерены для поверхностного и промежуточного слоев. Для эпителия E_d находится в диапазоне 10^5 — 10^7 Па и демонстрирует эффект насыщения при удлинениях порядка 30—40%. Для промежуточного слоя насыщение наступает позже — в районе 40—70% удлинения, а дифференциальный модуль упругости находится в диапазоне 10^4 — 10^7 Па. Дифференциальный модуль упругости промежуточного слоя в поперечной плоскости не обнаруживает насыщения вплоть до разрыва тканей (при удлинении в 200—300%) и примерно линейно зависит от удлинения, имея величины порядка $5 \cdot 10^3$ — $5 \cdot 10^4$ Па в диапазоне удлинений 5—70%, т. е. примерно на порядок меньше дифференциального модуля упругости этого слоя в продольном направлении. Промежуточный слой содержит эластичные и коллагеновые волокна, а также желатинообразную среду. Модуль упругости вдоль коллагеновых волокон очень велик — он оценивается от 10^8 Па (по [2.32]) до 10^9 Па (по [2.1]), тогда как упругость эластичных волокон значительно ниже — порядка $6 \cdot 10^5$ Па [2.32]. Сопоставляя эти величины с данными об упругих характеристиках промежуточного слоя, можно сделать вывод, что при умеренных удлинениях упругость определяется эластичными волокнами, а при больших — коллагеновыми.

Модуль сдвига тканей голосовых складок определяется в [1.26] с помощью кручения образца заданным моментом силы, изменяющейся периодически с частотой 0,1 Гц, и измерения угла поворота для расслабленного состояния и при удлинении складок примерно на 40%. Модуль сдвига G при этом изменялся незначительно и находился в диапазоне $(2-3) \cdot 10^4$ Па как для про-

межуточного слоя, так и для мышечной ткани. Вместе с тем коэффициент вязкости, определенный по затуханию колебаний, примерно в 2—3 раза больше для мышечных тканей, чем для промежуточного слоя, оставаясь по порядку величины близкой к 10 Пуаз, причем эти соотношения и абсолютные величины изменятся незначительно при удлинении образца на 40%.

Эти экспериментальные данные о механических параметрах тканей голосовых складок создают основу для конструирования количественных математических моделей, хотя и нуждаются в коррекции в связи с тем, что измерения производились на препаратах гортани и ряд данных относится к тканям собаки. Осуществить подобную коррекцию в известных пределах позволяют эксперименты на живых голосовых складках человека, проведенные в [2.60]. Эти эксперименты проливают также свет на некоторые особенности механики голосовых складок. В [2.60] применялся механический вибратор, возбуждающий колебания тканей голосовых складок человека в диапазоне частот 30—300 Гц. Одновременно ультразвуковой эхо-локатор измерял амплитуду этих колебаний в латеральной плоскости при различной степени напряженности складок. Наряду с амплитудой колебаний измерялась и их добротность $Q = f_0 / \Delta f$, где f_0 — резонансная частота; Δf — приращение частоты, при котором амплитуда колебаний падает до $1/\sqrt{2}$ от амплитуды на резонансной частоте f_0 . Соответственно коэффициент затухания $\delta = 1/2Q$.

При исследовании 17 взрослых мужчин и 19 женщин было установлено, что резонансные частоты колебаний ненапряженных голосовых складок для мужчин лежат в диапазоне 91—145 Гц со средним значением 128 Гц, а для женщин — в диапазоне 115—167 Гц со средним значением 136 Гц. Первая величина совпадает со статистической оценкой средней частоты основного тона мужчин — 129 Гц [2.6], тогда как резонансные характеристики голосовых складок для женщин заметно отличаются от средне-статистической частоты их основного тона, оцениваемой как 240 Гц. Добротность колебаний для мужчин находится в пределах 2,1—5,1 со средним значением 3,4, а для женщин — в пределах 2,6—7,1 со средним значением 4, что соответствует коэффициентам затухания от 0,08 до 0,25 и довольно близко к оценке 0,1—0,2, приведенной в [2.54].

Во время фонации механические характеристики голосовых складок изменяются за счет натяжения и сокращения голосовой и перстне-черпаловидной мышц. Измерения частотных характеристик складок непосредственно перед началом фонации с заданной частотой основного тока выявили неожиданные и важные явления, поскольку обнаружилась разница в характеристиках одного и того же человека при фонации с низкой или высокой частотой основного тона, а также между характеристиками мужчин и женщин. У мужчин при низкой частоте основного тона F_0 порядка 100 Гц резонансная частота складок совпадает с F_0 , а при вы-

соких частотах основного тона появляются два механических резонанса: один вблизи 100 Гц, а другой — на частоте основного тона. У женщин два резонанса обычно присутствуют независимо от частоты основного тона, причем один также находится вблизи 100 Гц, несколько повышаясь с повышением F_0 , а второй — вблизи F_0 . Затухание же не изменяется существенно ни для высокого, ни для низкого основного тона.

Эти данные свидетельствуют о том, что в голосовых складках существуют по крайней мере две механические структуры, причем свойства одной из них мало зависят от натяжения и напряжения голосовых складок, тогда как другая изменяет свои параметры в соответствии с активностью мышц гортани. Можно предположить, что последняя структура связана главным образом с голосовой и перстне-черпаловидной мышцами, а первая — с немышечными тканями голосовых складок, возможно с нижними, примыкающими к трахее участками голосовых складок, которые мало поддаются влиянию натяжения.

2.3. КОЛЕБАТЕЛЬНЫЕ ПРОЦЕССЫ В ГОЛОСОВЫХ СКЛАДКАХ

Частота колебаний голосовых складок у мужчин может достигать величин порядка 160 Гц, а у женщин — более 350 Гц, причем спектр звонких звуков (т. е. образованных с участием голосового источника) распространяется до 5 кГц. Поэтому для исследования колебаний голосовых складок должны применяться либо непрерывные методы измерений, либо скоростная киносъемка. С помощью таких методов было установлено, что в процессе фонации голосовые складки совершают колебательные движения в трех направлениях: в латеральной плоскости поперек голосовой щели и во фронтальной плоскости вдоль оси речевого тракта и поперек голосовой щели [2.13, 2.31, 1.22, 2.40, 2.84, 2.96]. Кроме того, после схлопывания складок по их верхней поверхности бегут поверхностные волны [2.41]. Скоростная киносъемка колебаний в латеральной плоскости со скоростью около 2000 кадр./с была впервые выполнена в [2.31]. Уже на такой сравнительно малой скорости можно увидеть, что голосовые складки изгибаются и что существует не только одна гармоника изгибных колебаний. Скоростная киносъемка и стробоскопические наблюдения за колебаниями складок в латеральной плоскости, а также кинорентгено съемка во фронтальной плоскости [2.84] или ультразвуковая локация [2.60] выявляют сдвиг по фазе колебаний верхнего и нижнего краев складок. Этот сдвиг по фазе, однако, существует не всегда, а зависит от регистра, в котором происходит фонация, т. е. от соотношения активности перстне-щитовидной, голосовой и щито-черпаловидной мышц.

Исследование движений голосовых складок во фронтальной плоскости было впервые выполнено в [2.13] на препаратах горта-

ни собаки с помощью кинорентгено съемки траекторий трех металлических частичек, внедренных в голосовые складки вблизи края, в средней и боковой частях (рис. 2.9 по [2.14]). Траектории этих частичек свидетельствуют о том, что вертикальное смещение складок не меньше их горизонтального смещения, а для некоторых точек в 2—4 раза больше. Аналогичные эксперименты были выполнены и в [2.84] для препаратов не только гортани собаки, но и гортани человека, а также на функционирующем голосовом аппарате человека. На препаратах использовались либо три частички, расположенные примерно так же, как и в [2.13], либо шесть ча-

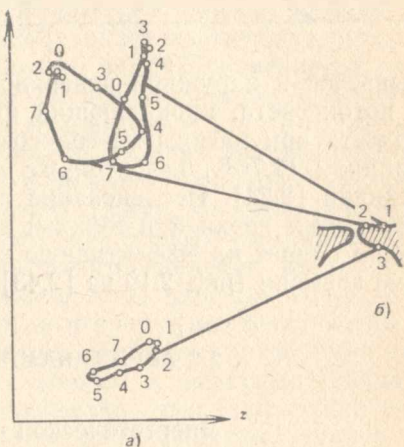


Рис. 2.9. Траектории (а) и расположение (б) металлических частичек в тканях голосовых складок. Цифры на траекториях — отчеты через равные промежутки времени ($F_0 = 100$ Гц)

стичек, распределенных в поверхностном слое в верхней части голосовых складок. В первом случае было установлено, что траектории частичек близки к эллипсоидальным с наименьшим эксцентриситетом для краев складок и с наибольшим (с преобладанием вертикальной оси над горизонтальной в 3—5 раз) для боковой и внутренних частичек. Во втором случае траектории частичек продемонстрировали существование поверхностной волны, распространяющейся от нижней кромки складок вверх и в стороны на верхней поверхности. Съемки колебаний складок во фронтальной плоскости на живом человеке, хотя и не обеспечивают четкого изображения, все же подтверждают заключение о волновом характере колебаний складок во фронтальной плоскости и показывают, что форма и размер морганиевого желудочка также изменяются в процессе фонации. Скорость распространения поверхностной волны, измеренная по смещению краев голосовых складок и кровеносных сосудов, изменяется от 2,5 до 1,5 м/с при удалении от краев складок [2.41] (измерения производились со скоростью 5000 кадр./с). При этом амплитуда горизонтального смещения тканей падала от 0,96 до 0,44 мм при удалении от краев складок к стенкам щитовидного хряща.

Наряду с измерениями смещений голосовых складок во время фонации представляют интерес измерения площади голосовой щели, поскольку характеристики голосового возбуждения во многом определяются именно изменениями площади голосовой щели во времени. Эта площадь может быть измерена с помощью не только скоростной киносъемки, но и путем так называемого метода трансиллюминации, при котором сильный источник света прикла-

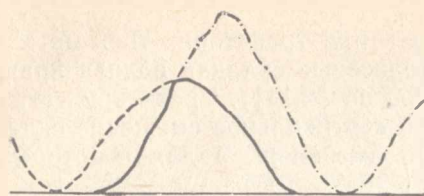


Рис. 2.10. Траектории верхней (---) и нижней (—) кромок голосовых складок. Сплошная линия — результирующая траектория, определяющая площадь голосовой щели

дывается к наружной поверхности шен под щитовидным хрящом и поток света, проникающий из подсвязочной области в речевую полость при раскрытии голосовой щели, регистрируется фотоприемником [2.70]. Для этой же цели применяется и телевизионная техника [2.24]. Исследования показали, что сдвиг фаз между колебаниями верхней и нижней кромок голосовых складок существенно влияет на эффективную площадь голосовой щели как функцию времени (рис. 2.10 из [2.43]).

2.4. АЭРОДИНАМИКА ГОЛОСОВОГО ИСТОЧНИКА

Энергетической основой колебаний голосовых складок служит избыточное давление ΔP , создаваемое в легких путем сокращения дыхательных мышц. Это давление изменяется от 30—40 мм вод. ст., когда фонация начинается, до 1300 мм вод. ст. при взрыве глухих согласных [В.3, 2.81]. Обычно ΔP изменяется в пределах 35—420 мм вод. ст. [2.19]. На голосовые складки действует сила, зависящая от перепада подсвязочного и надсвязочного давлений. Надсвязочное давление определяется сопротивлением воздушному потоку в речевом тракте, так что ΔP в конечном счете есть функция не только степени напряжения дыхатель-

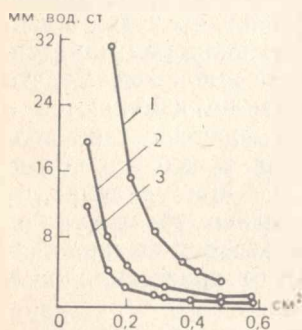


Рис. 2.11. Перепад давления в зависимости от площади голосовой щели и речевого тракта

Площадь сужения в тракте:
1 — $S_T = 0$; 2 — $S_T = 0,05$; 3 — $S_T = 0,1$ см²

ных мышц, но и площади голосовой щели и сопротивления речевого тракта. Эта зависимость была исследована на механическом аналоге речевого аппарата в [2.106] (рис. 2.11). Механизм изменения и влияния перепада давления посвящено много работ, в частности [2.26, 2.52, 2.74].

Как видно из рис. 2.11, при смычке в речевом тракте ΔP должно было бы стремиться к нулю. Однако во многих случаях при артикуляции звонких взрывных колебания складок не прекращаются во время смычки. Рентгенографические измерения показывают, что в этом случае расширяется фарингиальная область речевого тракта, за счет чего и поддерживается необходимый для фонации перепад давлений над голосовыми складками и

под ними [2.61, 2.79]. Существуют различные мнения о механике расширения фарингиальной области, которые можно свести к двум гипотезам: пассивное расширение фаринкса за счет податливости тканей при повышении давления в речевой полости и активное его расширение за счет сокращения некоторых мышц. Исследования мышечной активности фарингиальной области и данные об опускании подъязычной косточки [2.16, 2.61] свидетельствуют в пользу активного расширения фаринкса. Однако в [2.97] отмечается, что в слове *dad* по сравнению со слогом *tat* после начального *d* подъязычная косточка не опускается, а поднимается и к тому же не зарегистрировано различий в активности заднего участка подборочно-язычного мышцы, которая содействует расширению фаринкса. Существует еще одна гипотеза [2.81], согласно которой небольшой проход в носовую полость на интервале смычки может препятствовать повышению внутриротового давления выше критической величины, за которой прекращается фонация голосовых складок. Этот фактор, однако, вряд ли играет существенную роль, как можно заключить из сопоставления кинорентгенограмм в сагиттальной плоскости для глухих, звонких и назальных звуков. Рентгенографические измерения показывают, что расширение фарингиальной области происходит не только при артикуляции звука *Б* по сравнению со звуком *П*, но и при артикуляции звука *М* по сравнению со звуком *А*, хотя к моменту смычки проход в носовую полость полностью открыт.

Поскольку сопротивление воздушному потоку в речевом тракте во время артикуляции изменяется, для сохранения стабильности колебаний голосовых складок должны существовать какие-то механизмы, управляющие активностью дыхательных мышц и соответственно подсвязочным давлением. Имеются некоторые данные, что такой механизм действительно существует и что задержка в реакции дыхательных мышц на изменение сопротивления речевого тракта составляет лишь несколько десятков миллисекунд (33—80 мс по [2.27, 2.90]). Участие дыхательных мышц в формировании речи, однако, исследовано недостаточно полно. Так, известно, что эти мышцы играют основную роль в изменении интенсивности голоса [2.38, 2.39, 2.53], но данные об участии этих мышц в формировании слоговой структуры речи весьма противоречивы. Например, в [2.25, 2.100] сообщается о наблюдениях пиков подсвязочного давления, синхронных со слогами в слитной речи, тогда как в [2.65] это явление обнаружено не было.

Перепад подсвязочного и надсвязочного давлений ΔP влияет не только на интенсивность голоса, но и на частоту колебаний голосовых складок, т. е. на частоту основного тона F_0 . Этому явлению было посвящено много исследований, причем было установлено, что повышение ΔP , например при внезапном давлении на живот, приводит к повышению F_0 [2.20]. По [2.65] частота основного тона увеличивается на 1/2 октавы при увеличении подсвязочного давления на 60—70 мм вод. ст. При малых и больших значениях частоты основного тона связь между ΔP и F_0 несколько

различна [2.38, 2.39], причем считается, что на каждые 10 мм вод. ст. F_0 увеличивается на 2—4 Гц [2.51, 2.64, 2.78], тогда как значения в 3—18 Гц на 10 мм вод. ст., приводящиеся в [2.67], по-видимому, завышены и не были подтверждены другими исследованиями.

Избыточное подсвязочное давление ΔP создает силу, вызывающую движение потока воздуха из легких в ротовую полость через голосовую щель. Уже в [2.107] на механической модели гортани была установлена нелинейная зависимость скорости потока воздуха от избыточного давления ΔP . Дальнейшие исследования подтвердили важную роль сил, связанных с эффектом Бернулли в движущихся потоках воздуха, в процессах колебаний голосовых складок [2.21], причем эмпирически были определены компоненты аэродинамического сопротивления голосовой щели: $R = R_{тр} + R_B$, где сопротивление трения в пограничном слое принималось равным $R_{тр} = 12\mu h l_{г.щ}^2 / S_{г.щ}^3$, а динамическое сопротивление, возникающее вследствие эффекта Бернулли,

$$R_B = \kappa_B \rho_0 \omega / 2S_{г.щ}^2.$$

Здесь μ — коэффициент вязкости воздуха; h — длина голосовой щели вдоль направления воздушного потока и оси речевого тракта; $l_{г.щ}$ — длина голосовых складок; $S_{г.щ}$ — площадь голосовой щели; ρ_0 — плотность воздуха; ω — объемная скорость потока через голосовую щель. Коэффициент $\kappa_B = 0,875$ при динамическом сопротивлении [2.21].

Основываясь на этих результатах, в [В.3] и [В.4] были сконструированы нелинейные уравнения воздушного потока через голосовую щель:

$$R(t)\omega(t) + L(t)\frac{d\omega(t)}{dt} = \Delta P \quad \text{и} \quad R(t)\omega(t) + \frac{d}{dt}[L(t)\omega(t)] = \Delta P$$

соответственно. Здесь $L(t)$ отражает инерционность воздуха в голосовой щели. Отмечалось, что общее решение этих нелинейных уравнений при произвольной зависимости $S_{г.щ}$ от времени затруднительно, причем в [В.4] принималось, что инерционностью переходных процессов воздушного потока в голосовой щели можно пренебречь, считая ее малой (порядка 0,25 мс) по сравнению со средним значением длительности интервала открытой голосовой щели (около 2,5 мс).

Дальнейшие исследования свойств воздушного потока были выполнены на механической модели гортани [2.55]. Было показано, что внутри голосовой щели поток ламинарен, поскольку число Рейнольдса Re ниже порогового значения $Re_{кр} \approx 1600$, при котором для данной формы поперечного сечения возникает турбулентное (с завихрениями) течение воздушного потока. Действительно, поскольку число Рейнольдса определяется как $Re = vd/\nu$, где d — ширина голосовой щели; v — средняя линейная скорость воздушного потока; ν — кинематическая вязкость воздуха, то, принимая по экспериментальным данным максимальную линейную скорость

потока как $v=50$ м/с, $d=0,3$ см, а $\nu=0,166$ см²/с, получаем $Re \approx 920 < 1600$.

При выходе из голосовой щели воздушный поток, однако, попадает в морганиев желудочек, имеющий гораздо большую площадь поперечного сечения (примерно до 2—3 см²), в результате чего число Рейнольдса уже превышает пороговое значение и в потоке воздуха образуются завихрения, сопровождающиеся турбулентными шумами. Спектр этих шумов зависит от геометрии морганиева желудочка, играющего роль фильтра низких частот с передаточной функцией

$$W(j\omega) = \frac{-j \frac{c_1}{\omega}}{-j \frac{c_1}{\omega} + j\omega c_2 + c_3},$$

где c_1, c_2, c_3 — коэффициенты, зависящие от геометрических параметров морганиева желудочка [2.18].

В [2.55] теоретически и экспериментально исследовалась аэродинамика голосовой щели, где, в частности, было установлено, что коэффициент динамического сопротивления зависит от соотношения площадей голосовой щели $S_{г.щ}$ и морганиева желудочка $S_{м.ж}$ над верхней поверхностью голосовых складок:

$$\kappa_B = 1,375 - 2 \frac{S_{г.щ}(t)}{S_{м.ж}(t)} \left[1 - \frac{S_{г.щ}(t)}{S_{м.ж}(t)} \right].$$

Результаты этих исследований в общем были подтверждены в [2.91].

В аэродинамике зависимость скорости воздушного потока в канале, ось которого направлена вдоль оси x , от площади поперечного сечения канала $S(x)$ при отсутствии потерь описывается уравнением Гюгонио [2.4]

$$(M^2 - 1) \frac{dv}{v} = \frac{dS}{S},$$

где $M = v/c_0$ — число Маха; c_0 — скорость звука в воздухе. Принимая максимальную линейную скорость потока через голосовую щель порядка 50 м/с, имеем $M^2 \approx 0,02$ и, следовательно, фактором сжимаемости воздуха в процессах колебаний голосовых складок можно пренебречь. Тогда уравнение Гюгонио вырождается в обыкновенное дифференциальное уравнение первого порядка $-dv/v = dS/dS$, решение которого есть $vS = \text{const}$, т. е. условие неразрывности потока при дозвуковых течениях.

Рассмотрим теперь динамику воздушного потока в голосовой щели с изменяющейся во времени, но постоянной вдоль оси голо-

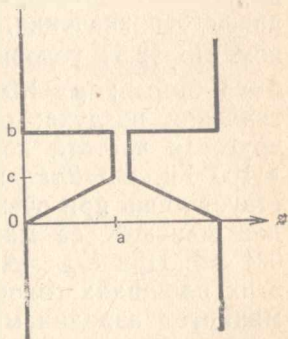


Рис. 2.12. Геометрическая модель голосовых складок

совой щели площадью поперечного сечения $S_{г.щ}$, считая поток ламинарным и несжимаемым. Пользуясь упрощенной моделью, показанной на рис. 2.12, обозначим длину голосовой щели (равную длине голосовых складок) $l_{г.щ}$, ширину голосовой щели в латеральной плоскости (расстояние между складками) $b_{г.щ}$, а глубину голосовой щели в направлении трахея — речевой тракт $h_{г.щ}$. Составим уравнение баланса сил, действующих на воздух, находящийся в голосовой щели. Тогда сила инерции этого слоя

$$F_{\text{и}} = \rho_0 h_{г.щ} \frac{d(v S_{г.щ})}{dt},$$

сила, возникающая за счет эффекта Бернулли,

$$F_{\text{Б}} = \kappa_{\text{Б}} \rho_0 S_{г.щ} \frac{v^2}{2},$$

а сила, создаваемая перепадом подсвязочного и надсвязочного давлений,

$$F_{\text{д}} = \Delta P S_{г.щ}.$$

Эта система сил должна быть дополнена еще и силами трения, преобладающими при стремлении площади голосовой щели к нулю: $F_{\text{тр}} = \kappa_{\text{тр}} v S_{г.щ}$. Обычно принимается, что сила трения соответствует силе, возникающей в так называемых капиллярных каналах, считая форму голосовой щели в латеральной плоскости прямоугольной:

$$\kappa_{\text{тр}} = 12 \mu h_{г.щ} / b_{г.щ}^2. \quad (2.1)$$

Коэффициент трения, определяемый по (2.1), справедлив лишь для величин $b_{г.щ}$, сравнимых с толщиной пограничного слоя потока Δ , т. е. такого слоя вблизи стенок щели, в котором скорость воздушного потока изменяется от нуля (на стенках) до установившегося значения, на которое уже не влияет присутствие стенок. По [2.5] условие справедливости (2.1) есть $b_{г.щ} \leq \Delta \sqrt{2}$, где $\Delta = \sqrt{2\mu/\rho_0\omega}$; $\mu = 1,86 \cdot 10^{-4}$ г/см·с — коэффициент вязкого сопротивления воздуха; $\rho_0 = 1,14 \cdot 10^{-3}$ г/см³ — плотность воздуха; ω — круговая частота колебаний. При частоте основного тона $F_0 = 100$ Гц, $\omega_0 = 2\pi F_0$ и $\Delta \approx 0,02$ см, т. е. капиллярное трение действует лишь при очень малых площадях голосовой щели, на порядок меньших ее максимального значения, примерно равного 0,2 см². При $b_{г.щ} > \Delta$ действует уже другой закон трения, но при этих площадях голосовой щели, как будет показано в гл. 4, доминирует аэродинамическое сопротивление, так что вязким трением вообще можно пренебречь.

Итак, баланс сил для воздушного потока есть $F_{\text{и}} + F_{\text{тр}} + F_{\text{Б}} = F_{\text{д}}$, и, таким образом, уравнение динамики потока представляется в виде нелинейного дифференциального уравнения, относящегося к типу уравнения Риккати:

$$\rho_0 h_{г.щ} \frac{d(v S_{г.щ})}{dt} + \kappa_{\text{тр}} v S_{г.щ} + \kappa_{\text{Б}} \rho_0 S_{г.щ} \frac{v^2}{2} = \Delta P S_{г.щ}. \quad (2.2)$$

Если приближенно принять форму голосовой щели близкой к прямоугольной, то коэффициент трения можно представить как

Перепишем затем (2.2) в виде

(2.1).

$$v' + \left(-\frac{\kappa_{\text{ТР}}}{\rho_0 h_{\text{г.щ}}} + \frac{S'_{\text{г.щ}}}{S_{\text{г.щ}}} \right) v + \frac{\kappa_{\text{Б}}}{2h_{\text{г.щ}}} v^2 + \frac{\Delta P}{\rho_0 h_{\text{г.щ}}} = 0,$$

где штрих означает производную по времени t , и произведем подстановку

$$\bar{v} = v e^{\int \frac{\kappa_{\text{Б}} v}{2 h_{\text{г.щ}}} dt} \quad (2.3)$$

В результате этой подстановки получим линейное уравнение второго порядка относительно новой переменной

$$\bar{v}' - \left(\frac{\kappa'_{\text{Б}}}{\kappa_{\text{Б}}} + \frac{S'_{\text{г.щ}}}{S_{\text{г.щ}}} + \frac{\kappa_{\text{ТР}}}{\rho_0 h_{\text{г.щ}}} \right) \bar{v} - \frac{\Delta P \kappa_{\text{Б}}}{2 \rho_0 h_{\text{г.щ}}} \bar{v} = 0.$$

Другой подстановкой

$$\tilde{v} = \bar{v} e^{\int \frac{g(t)}{2} dt},$$

где

$$g(t) = \frac{\kappa'_{\text{Б}}}{\kappa_{\text{Б}}} + \frac{S'_{\text{г.щ}}}{S_{\text{г.щ}}} + \frac{\kappa_{\text{ТР}}}{\rho_0 h_{\text{г.щ}}},$$

получим так называемую каноническую форму для уравнения второго порядка [2.2]

$$\tilde{v}'' - \tilde{g}(t) \tilde{v} = 0, \quad (2.4)$$

где

$$\tilde{g}(t) = \frac{\Delta P \kappa_{\text{Б}}}{2 \rho_0 h_{\text{г.щ}}} + \frac{g^2(t)}{4} + \frac{g'(t)}{2}.$$

Из (2.3) имеем $v = \frac{2h_{\text{г.щ}}}{\kappa_{\text{Б}}} \frac{\bar{v}}{\bar{v}}$, так что

$$v = \frac{2h_{\text{г.щ}}}{\kappa_{\text{Б}}} \left(\frac{\tilde{v}}{\tilde{v}} - \frac{\kappa_{\text{ТР}}}{2 \rho_0 h_{\text{г.щ}}} \right). \quad (2.5)$$

В частном случае постоянной во времени площади голосовой щели, когда $\tilde{g}$ не зависит от времени, уравнение (2.4) решается в квадратурах:

$$\tilde{v} = C_1 \sqrt{\tilde{g}t} + C_2 \text{sh} \sqrt{\tilde{g}t}. \quad (2.6)$$

Обозначим начальную скорость воздушного потока в момент $t=0$ через v_0 . Тогда с учетом (2.5) и (2.6) имеем

$$v(t) = \frac{2h_{г.щ}}{\kappa_B} \left[\sqrt{\tilde{g}} \operatorname{th}(\sqrt{\tilde{g}} t + \theta) - \frac{\kappa_{тр}}{2\rho_0 h_{г.щ}} \right], \quad (2.7)$$

где

$$\tilde{g} = \frac{\Delta P \kappa_B}{2\rho_0 h_{г.щ}^2} + \left(\frac{\kappa_{тр}}{2\rho_0 h_{г.щ}} \right)^2;$$

$$\theta = \operatorname{Arth} \left[\frac{1}{2 h_{г.щ} \sqrt{\tilde{g}}} \left(v_0 \kappa_B + \frac{\kappa_{тр}}{\rho_0} \right) \right].$$

При отсутствии потерь на трение, т. е. при $\kappa_{тр}=0$, решение (2.7) совпадает с решением, полученным в [2.3]. Как видно из (2.7), при малых площадях голосовой щели $S_{г.щ} \rightarrow 0$ коэффициент трения $\kappa_{тр}$ стремится к бесконечности и скорость потока $v \rightarrow 0$. Обозначим установившееся значение скорости потока при $t \rightarrow \infty$ через v_∞ . Тогда

$$v_\infty = \frac{2h_{г.щ}}{\kappa_B} \left(\sqrt{\tilde{g}} - \frac{\kappa_{тр}}{2\rho_0 h_{г.щ}} \right).$$

и (2.7) можно переписать как

$$v(t) = \left(v_\infty + \frac{\kappa_{тр}}{\rho_0 \kappa_B} \right) \operatorname{th}(\sqrt{\tilde{g}} t + \theta) - \frac{\kappa_{тр}}{\rho_0 \kappa_B},$$

или

$$v(t) = v_\infty \operatorname{th}(\sqrt{\tilde{g}} t + \theta) + \frac{2\kappa_{тр}}{\rho_0 \kappa_B} [\operatorname{th}(\sqrt{\tilde{g}} t + \theta) - 1]. \quad (2.8)$$

При скачкообразном раскрытии голосовой щели до максимального значения, равного $b_{г.щ}=0,2$ см, и скорости потока $v_\infty=50$ м/с текущая величина скорости потока v отличается от установившегося значения v_∞ не более чем на 5% примерно через 0,4 мс. Эта длительность переходного процесса для скорости потока в голосовой щели близка к величине 0,25 мс, полученной в [В.4], хотя и превышает ее почти в 1,5 раза.

При произвольной зависимости $S_{г.щ}(t)$ для (2.2) не существует аналитического решения. В этом случае необходимо применить приближенные или численные методы решения дифференциальных уравнений. Для решения (2.2) удобно использовать модифицированный метод Эйлера, который приводит к следующей системе:

$$\left. \begin{aligned} \tilde{w} \left(t + \frac{\Delta t}{2} \right) &= w(t) + \frac{\Delta t}{2} f \left[w(t), t + \frac{\Delta t}{4} \right]; \\ w(t + \Delta t) &= w(t) + \Delta t f \left[\tilde{w} \left(t + \frac{\Delta t}{2} \right), t + \frac{\Delta t}{2} \right], \end{aligned} \right\} \quad (2.9)$$

где $w = v S_{г.щ}$ — объемная скорость воздушного потока через голосовую щель, а

$$f[w(\tau), t] = \frac{1}{\rho_0 h_{г.щ}(t)} \left\{ \Delta P S_{г.щ}(t) - w(\tau) \left[\kappa_{тр}(t) + \frac{\rho_0 \kappa_B(t) w(\tau)}{2 S_{г.щ}(t)} \right] \right\}.$$

При сопоставлении решения (2.8) с конечно-разностным решением системы (2.9) выяснилось, что для $S'_{г.щ} \ll S_{г.щ}$ они довольно близки. Например, если принять $S_{г.щ}(t) = a \sin 2\pi F_0 t$, где $a \leq 0,1$, $50 \text{ Гц} \leq F_0 \leq 200 \text{ Гц}$, то разница по абсолютной величине между численным и аналитическим решениями не превышает 10%. При $a > 0,1$ влияние скорости изменения площади поперечного сечения $S_{г.щ}$ начинает возрастать и функция скорости потока $v(t)$ становится несимметричной относительно середины периода — скорость нарастания v на начальных участках раскрытия голосовой щели становится меньше, чем дает (2.8), тогда как на конечных участках при закрывании голосовой щели скорость v больше сохраняет большие значения, а затем круче, чем в (2.8), падает до нуля. Большая крутизна v на участках схлопывания голосовых складок приводит к большему значению импульса, возбуждающего акустические колебания в речевом тракте, поэтому квазистационарные решения типа (2.8) имеют ограниченную область применения (при малых максимальных значениях площади голосовой щели). Пренебрежение динамикой воздушного потока может привести к существенному ухудшению перцептивных характеристик голосового источника. Например, в [2.75] не удалось улучшить звучание синтетической речи лишь за счет использования в качестве возбуждения функции площади голосовой щели $S_{г.щ}(t)$, измеренной с помощью скоростной киносъемки со скоростью до 4000 кадр./с. Отсутствие улучшения натуральности в этом случае, по крайней мере частично, можно объяснить недостаточно точным описанием нелинейной зависимости между площадью голосовой щели и скоростью воздушного потока.

Численные методы решения уравнений, однако, при всей простоте вычислительной схемы порождают свои проблемы, если они реализуются на цифровых вычислительных машинах. Ограниченная длина кодового машинного слова, представляющего числа, может привести к неустойчивости решения. Подобная неустойчивость возникает при использовании модифицированного метода Эйлера для решения уравнения голосовой щели при малых значениях площади ее поперечного сечения. Для устранения этой неустойчивости приходится применять специальные приемы, поиск которых иногда весьма трудоемок. В случае уравнения голосовой щели неустойчивость решения при малых площадях поперечного сечения можно устранить за счет перехода к асимптотическому аналитическому решению, воспользовавшись именно этим фактором.

Предположение о постоянстве площади голосовой щели в поперечном направлении приводит к занижению инерционности воз-

духа в ней, в результате чего скорость нарастания потока при раскрытии щели оказывается близкой к скорости его торможения при ее закрытии. Это приводит к тому, что в модели речевого тракта возбуждаются акустические колебания примерно одинаковой амплитуды как передним, так и задним фронтами импульса объемной скорости, что противоречит наблюдаемым свойствам речевого сигнала. Вследствие явления сдвига фаз к моменту раскрытия голосовой щели нижние кромки уже расходятся на определенное расстояние, так что масса воздуха в щели на начальных этапах ее раскрытия оказывается значительно большей, в результате чего скорость нарастания потока уменьшается и возбуждение акустических колебаний в речевом тракте вызывается главным образом задним фронтом импульса объемной скорости.

Перепад давления ΔP есть разность подсвязочного P_n и надсвязочного P_n давлений. В подсвязочной области при раскрытии голосовой щели воздушный поток начинает ускоряться еще в трахее на некотором расстоянии от голосовой щели. В этой области для среднего давления справедлив закон Бернулли

$$P_n + \frac{\rho_0 v^2}{2} = P_0,$$

где P_0 — давление в той области легких, в которой скорость потока равна нулю.

Скорость потока в трахее v можно рассчитать исходя из уравнения неразрывности, если известна скорость потока в голосовой щели. Эти расчеты показывают, что падение давления в трахее в результате непосредственного влияния эффекта Бернулли незначительно. Например, при объемной скорости, равной $700 \text{ см}^3/\text{с}$, минимальной площади голосовой щели, равной $0,2 \text{ см}^2$, и площади трахеи 3 см^2 падение давления равно 3 Па при давлении в легких 10^3 Па . Непосредственные же измерения подсвязочного давления с помощью миниатюрного микрофона, спускаемого в трахею через голосовую щель, показывают, что уже на расстоянии 4 см ниже голосовой щели в трахее возникают колебания давления, причем максимум подсвязочного давления строго соответствует минимуму надсвязочного давления при колебании складок, а изменения среднего надсвязочного давления не влияют на форму колебаний давления под голосовыми складками [2.63]. Минимальное подсвязочное давление при раскрытии голосовой щели может быть в 5 — 6 раз меньше максимального [2.62] и по данным [2.43] может опускаться до 20 — 30 мм вод. ст. Таким образом, падение давления значительно превышает расчетное, если принимать во внимание только эффект Бернулли в трахее. Причина этого явления заключается в том, что энергетический источник речи — легкие — не является источником тока, причем сопротивление легких зависит от скорости воздушного потока в бронхиолах, т. е. в конечном счете от эффекта Бернулли, но действующего в других отделах подсвязочной области. Таким образом, ΔP в (2.2) оказывается функцией скорости потока v . Про-

цессы изменения подсвязочного и надсвязочного давлений будут подробно описаны в гл. 4.

Влияние сопротивления речевого тракта воздушному потоку уже обсуждалось выше, однако колебания акустического давления в речевом тракте также влияют на надсвязочное давление, хотя этот эффект маскируется другими, более сильными влияниями [2.36]. Вихревые движения воздушного потока, возникающие в морганиевом желудочке вследствие резкого расширения канала, также могут повлиять на мгновенное значение давления над верхней поверхностью голосовых складок. Из-за шумовой природы турбулентных вихрей изменение давления носит случайный характер, и вполне возможно, что это явление вызывает небольшие (порядка нескольких процентов) случайные изменения периода и амплитуды импульсов голосового источника, которые наблюдаются в естественной речи и играют важную роль в восприятии натуральности речевого сигнала [2.50, 2.76, 2.83].

2.5. МЕХАНИКА КОЛЕБАНИЙ ГОЛОСОВЫХ СКЛАДОК

Приступая к конструированию количественной модели голосового источника, прежде всего необходимо определить цели моделирования, поскольку ясно, что модель, стремящаяся охватить все свойства реального объекта, будет вычислительно нереализуемой. Цель создания такой модели — обеспечение звучания синтетической речи, максимально приближающейся к звучанию естественной. Но и при этом необходимо решить вопрос о том, какие же физические свойства и процессы определяют натуральность звучания речи. Практика показывает, что время, потраченное на выявление главных элементов моделируемого явления, окупается при экспериментальной работе с моделью. При этом часто удается получить приближенные аналитические решения, снижающие остроту проблемы устойчивости и точности вычислений, порождаемую дискретизацией параметров и конечной точностью представления чисел в ЭВМ. Не менее важным, однако, является и то, что подобные решения позволяют глубже проникнуть в существо проблемы, указывают на области, требующие дополнительных экспериментальных исследований, а иногда позволяют обнаружить такие свойства и явления, которые вообще были неизвестны.

Основой количественных моделей служат экспериментальные исследования. Модели голосового источника развивались по мере накопления экспериментальных данных о процессах голосообразования. Прежде всего на основе экспериментальных исследований [2.21, 2.22] был описан эффект Бернулли в так называемой одномассовой модели [2.33], рассматривающей голосовые складки как систему с сосредоточенными параметрами. Затем появились двухмассовые модели [2.29, 2.56], учитывающие помимо эффекта Бернулли тот факт, что верхняя и нижняя кромки голосовых

складок во время фонации колеблются со сдвигом фаз. В дальнейшем были разработаны непрерывные модели голосовых складок [2.7, 2.8, 2.101—2.105], описывающие волны упругих деформаций, которые иногда отчетливо видны при регистрации колебаний складок. Необходимость представления голосовых складок как систем с распределенными параметрами может быть продемонстрирована путем сопоставления линейных размеров складок с длиной распространяющихся в них волн [2.105]. Воспользуемся уравнением из гл. 1 для скорости поперечных колебаний $c_n = \sqrt{E/\rho}$, подставим значение модуля упругости $E \approx 10^5$ Па и вычислим длину волны первой моды колебаний складок $\lambda_0 = c_n/F_0$ для частоты основного тона $F_0 = 100$ Гц. Принимая плотность тканей $\rho = 1,06$ г/см³, получаем длину волны λ_0 порядка 3 см, что сравнимо с длиной голосовых складок. Таким образом, ясно, что адекватная модель колебаний голосовых складок должна тем или иным образом описывать упругие деформации тканей складок и их зависимость от давления и аэродинамической силы.

Опираясь на результаты экспериментальных исследований колебаний голосовых складок, описанных в предыдущих параграфах, рассмотрим качественно механику этих колебаний. Примем, что геометрия и механические параметры обеих складок идентичны, т. е. что анализ сил и деформаций для одной из складок справедлив и для другой. Введем декартову систему координат, направив ось x вдоль складок, ось y вертикально вдоль оси речевого тракта, ось z горизонтально во фронтальной плоскости. Выберем некоторое поперечное сечение складок во фронтальной плоскости при $z = z_0$ и рассмотрим силы и моменты, возникающие в этом сечении в процессе фонации.

В число этих сил входит сила $F_{\text{пд}}$, создаваемая перепадом статического давления над и под складками ΔP , аэродинамическая сила F_a , возникающая за счет эффекта Бернулли при обтекании потоком воздуха криволинейного профиля голосовых складок, сила упругих деформаций тканей голосовых складок F_y , сила инерции F_n и вес складок. Кроме того, действуют силы натяжения складок F_n и реакции боковых стенок F_p . Эти силы вызывают, с одной стороны, деформацию профиля складок, а с другой — стремятся сместить складки вверх или вниз. Оценим качественно изменение активных сил $F_{\text{пд}}$ и F_a в течение одного цикла, представленного на рис. 2.13, составленном по данным [1.22, 2.13, 2.14, 2.43, 2.84].

В начальный момент, когда складки находятся в сомкнутом состоянии, действуют только силы статического давления на нижний скошенный участок складок. Точка приложения и направление вектора равнодействующей этих сил $F_{\text{пд}}$ условно показаны на фазе 1. Эта сила вызывает деформацию контура складок, которые начинают расходиться снизу в стороны и подниматься вверх в силу свойства несжимаемости тканей. Одновременно вертикальная составляющая этой силы создает изгибающий момент, кото-

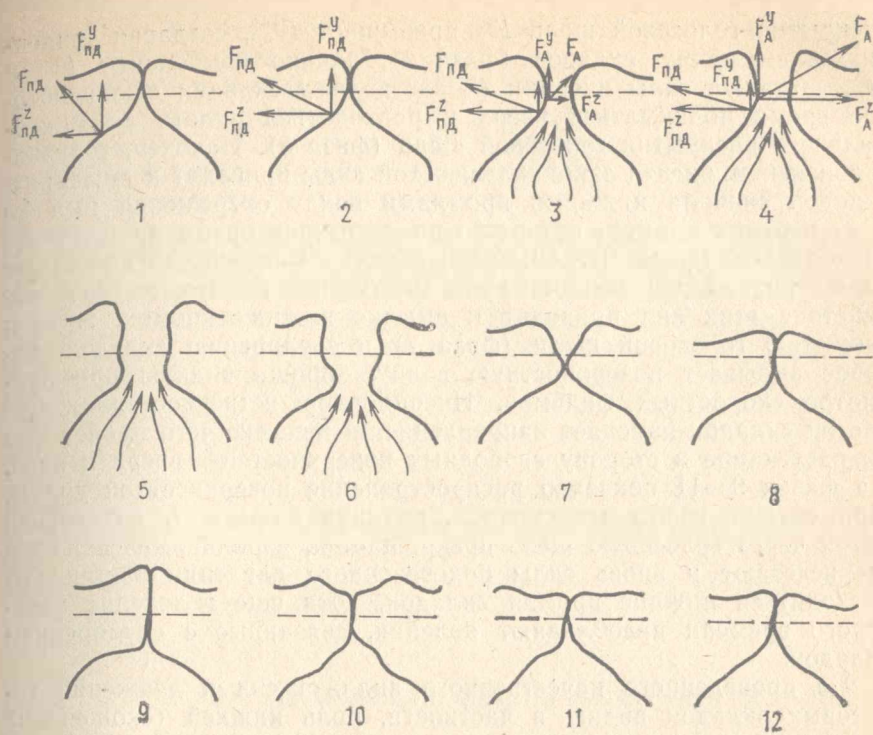


Рис. 2.13. Фазы колебаний голосовых складок и действующие силы

рый заставляет внутренние края складок прогибаться вверх. На фазе 2 нижние кромки складок разомкнулись и площадь их соприкосновения уменьшилась, в результате чего подсвязочному давлению противостоят уже более тонкие верхние кромки, что вызывает еще больший прогиб их вверх. По мере размыкания нижних кромок сила $F_{пд}$ возрастает за счет увеличения площади, на которую воздействует подсвязочное давление, точка ее приложения смещается вверх, а угол между вектором и горизонтальной плоскостью уменьшается, приводя к уменьшению вертикальной составляющей $F_{пд}^y$ при одновременном увеличении горизонтальной составляющей $F_{пд}^z$. Этот процесс завершается размыканием и верхних кромок, т. е. раскрытием голосовой щели. На фазе 3 представлен момент начальной стадии раскрытия щели, когда поток воздуха из легких устремляется в речевой тракт. С раскрытием щели возникает аэродинамическая сила F_A , горизонтальная компонента которой F_A^z направлена противоположно горизонтальной компоненте сил статического давления $F_{пд}^z$, тогда как вертикальная компонента F_A^y направлена в ту же сторону и способствует вертикальному смещению складок вверх. Быстрое возрастание аэродинамических сил, пропорциональных квадрату скорости воздушного потока, приводит к тому, что вскоре после

раскрытия голосовой щели F_{z_a} превышает $F_{z_{пд}}$, создавая тенденцию к смыканию складок (фаза 4), однако компоненты F_{v_a} и $F_{v_{пд}}$, а также силы инерции F_i заставляют складки еще некоторое время подниматься вверх и расходиться вплоть до максимального раскрытия голосовой щели (фаза 5). Наконец, горизонтальная компонента аэродинамической силы приводит к смыканию складок сначала нижними кромками ввиду опережения ими по фазе верхних кромок, которые еще некоторое время продолжают расходиться (фаза 6). Смыканию складок способствуют также силы упругих деформаций F_y и натяжения F_n . Совместное воздействие этих сил приводит к смычке нижних кромок, т. е. к закрытию голосовой щели (фаза 7) и соударению складок, которое вызывает поверхностную волну, хорошо видную при просмотре скоростных фильмов. Инерционное встречное движение частиц складок изменяет направление, вследствие чего происходит их растекание в сторону свободных поверхностей — вверх и вниз. На фазах 9—11 показано распространение поверхностной волны. При смычке голосовых складок, начиная с фазы 7, воздушный поток через голосовую щель прекращается, аэродинамические силы исчезают и вновь силы подсвязочного давления заставляют расходиться нижние кромки складок, хотя еще в течение некоторого времени преобладают явления, связанные с соударением складок.

Из проведенного качественного анализа сил и движений голосовых складок видна, в частности, роль нижней (скошенной) поверхности в формировании вертикальных движений и образовании сдвига фаз между верхней и нижней их частями. Более того, без учета этой поверхности невозможно описать процесс расхождения полностью сомкнутых складок, поэтому в моделях с сосредоточенными параметрами процессы фонации обычно начинаются со стадии раскрытой голосовой щели, а на стадии закрытой щели для формирования раздвигающих усилий применяются искусственные приемы. Роль асимметрии голосовых складок в направлении вдоль оси y может быть продемонстрирована попыткой генерирования звуков речи при обратном направлении воздушного потока — на вдохе. Тембр образуемых при этом звуков отличается неестественностью, поскольку нарушаются соотношения между горизонтальными и вертикальными движениями складок, а также изменяется аэродинамика голосовой щели.

Голосовые складки движутся вверх с наибольшей скоростью после их соударения из-за совместного действия подсвязочного давления и растекания частиц массы складок в стороны (в том числе и вверх). При напряжении голосовой и перстне-черпаловидной мышц натяжение наружных слоев уменьшается, их жесткость падает и под влиянием подсвязочного давления амплитуда вертикальных колебаний складок увеличивается. Это явление обнаруживается на препаратах гортани и при исследованиях трехмерной модели [2.105]. При расхождении складок вследствие нежимаемости их тканей снова происходит растекание частиц вверх

и вниз, но на этот раз вертикальная компонента равнодействующей силы подвязочного давления и аэродинамических сил уменьшается и в момент расхождения складок на максимальное расстояние в латеральной плоскости начинается их опускание. Это означает, что в вертикальных движениях складок основную роль играют силы давления и аэродинамические силы, а свойство несжимаемости тканей оказывает меньшее влияние.

Кроме рассмотренных сил необходимо учитывать и адгезионные силы, вызывающие слипание сомкнутых складок и задерживающие их расхождение. Это явление хорошо видно при просмотре скоростных кинофильмов. Влажная поверхность голосовых складок способствует их молекулярному сцеплению, и этот эффект может играть важную роль в процессах фонации. К тому же известно, что сухие голосовые складки теряют способность к фонации. Это может происходить потому, что при смыкании складок за счет неодинаковости механических и геометрических параметров возникают касательные движения складок относительно друг друга и сухая поверхность складок создает силы трения, гасящие автоколебания. Кроме того, жесткость сухой оболочки голосовых складок повышается, что также сужает область устойчивости автоколебаний.

2.6. ОДНОМЕРНАЯ МОДЕЛЬ ГОЛОСОВЫХ СКЛАДОК

Из теории упругости известно, что если один из размеров упругого тела в 3—5 раз превышает остальные, то такое тело можно рассматривать как брус и исследовать лишь одномерные упругие деформации. Такой подход существенно упрощает анализ колебаний и соответствующий математический аппарат. Как упоминалось ранее, средняя длина голосовых складок для мужчин в среднем равна 1,5 см (с возможным удлинением до 30% при сдвиге черпаловидных хрящей), толщина голосовых складок в латеральной плоскости примерно равна 0,6 см, а их высота в вертикальном направлении в зоне контакта при схлопывании — около 0,4 см. Такие геометрические соотношения уже позволяют рассматривать складки как брус при анализе колебаний в латеральной плоскости. Если же принять во внимание слоистую структуру голосовых складок и тот факт, что толщины этих слоев находятся в отношении примерно 1:5:7 [1.22], то вполне допустимо представление голосовой складки в виде двуслойного бруса, состоящего из наружного слоя и пучка мышечных волокон, лежащего на упругом основании.

Для вывода уравнений одномерных колебаний можно воспользоваться результатами гл. 1, однако проще рассмотреть деформации некоторого элемента длиной dx под влиянием приложенных усилий и моментов. Исходя из условия равновесия этого элемента приравняем нулю действующие моменты M , а также

проекции перерезывающей силы Q и распределенной нагрузки q на ось Y :

$$Q - q dx - Q - dQ = 0; \quad M + Q dx + q dx \frac{dx}{2} - M - dM = 0.$$

Пренебрегая величиной второго порядка малости, получаем:

$$Q = \frac{dM}{dx}; \quad q = -\frac{dQ}{dx} = -\frac{d^2 M}{dx^2}. \quad (2.10)$$

Под влиянием изгибающего момента кривизна средней линии элемента $K = -M/EJ_y$, где E — модуль упругости; J_y — момент инерции поперечного сечения относительно оси y . За среднюю линию принимается геометрическое место точек бруса, в котором отсутствуют как растяжение, так и сжатие (при изгибе бруса одна сторона его поверхности подвергается растяжению, а другая — сжатию). Выражая кривизну K в дифференциальном виде, получаем

$$K = \frac{d^2 u}{dx^2} \frac{1}{\left[1 + \left(\frac{du}{dx}\right)^2\right]^{3/2}},$$

где $u(x)$ — смещение средней линии в боковом направлении. Ограничиваясь малыми перемещениями, для которых справедливо $du/dx \approx 0$, получаем дифференциальное уравнение изгиба голосовых складок $EJ_y \frac{d^2 u}{dx^2} = -M$ или с учетом (2.10) $\frac{d^2 u}{dx^2} \left(EJ_y \frac{d^2 u}{dx^2}\right) = q$.

Сдвиг черпаловидных хрящей в заднем направлении создает некоторое удельное натяжение N , которое для малых деформаций не зависит от координаты x . Действительно, обозначая через α угол наклона касательной к поверхности складок, имеем

$$\sin \alpha = \frac{\operatorname{tg} \alpha}{\sqrt{1 + \operatorname{tg}^2 \alpha}} \approx \operatorname{tg} \alpha = \frac{du}{dx}.$$

Отсюда следует, что $\cos \alpha(x) \approx \cos \alpha(x+dx) \approx 1$ и из условия равновесия элемента складок $N(x) \cos \alpha(x) - N(x+dx) \cos \alpha(x+dx) = 0$ имеем $N(x) = N(x+dx)$. Поэтому поперечная сила, создаваемая натяжением,

$$N \sin \alpha(x) - N \sin \alpha(x+dx) = N \left(\frac{du}{dx} \Big|_x - \frac{du}{dx} \Big|_{x+dx} \right) \approx N \frac{d^2 u}{dx^2}.$$

В [2.9] было показано, что при больших значениях силы натяжения изменяется эквивалентная жесткость упругого тела, поэтому уравнение сил записывается так:

$$\frac{d^2}{dx^2} \left[J_y (E + N) \frac{d^2 u}{dx^2} \right] - NS \frac{d^2 u}{dx^2} = q.$$

Распределенная нагрузка q создается силами инерции, вязкого трения и упругостью подстилающего слоя:

$$-q = \rho_r \frac{\partial^2 u}{\partial t^2} + r \frac{\partial u}{\partial t} + c_n u,$$

где $\rho_T = \rho_T S$ — погонная плотность голосовых складок; $\rho_0 = 1,06$ г/см³; S — площадь поперечного сечения складок; r — погонный коэффициент трения; $c_n = E_n S_n / l_{г.с}$ — погонная упругость подстилающего слоя; E_n — модуль упругости подстилающего слоя; S_n — площадь поперечного сечения подстилающего слоя; $l_{г.с}$ — длина голосовых складок. Пренебрегая объемными силами в виде веса складок, получаем в результате дифференциальное уравнение одномерной модели колебаний голосовых складок

$$\frac{\partial^2}{\partial x^2} \left[J_y (E + N) \frac{\partial^2 u}{\partial x^2} \right] - N S \frac{\partial^2 u}{\partial x^2} + c_n u + r \frac{\partial u}{\partial t} + \rho_T \frac{\partial^2 u}{\partial t^2} = F(x, t).$$

Здесь $F(x, t)$ — распределенная нагрузка на голосовые складки в боковом направлении, создаваемая комбинацией сил подвязочно-го давления и аэродинамических сил воздушного потока через голосовую щель.

Вообще говоря, форма поперечного сечения складок зависит от координаты x , так что параметры J_y , ρ_T , r и c_n также зависят от x . Однако практически эта зависимость очень слабая, и параметры можно рассматривать как константы, за исключением, быть может, упругости подстилающего слоя, толщина которого изменяется примерно по параболическому закону от заднего к переднему концу складок. Модуль упругости тканей также можно считать не зависящим от координаты x . С учетом сказанного уравнение одномерных упругих колебаний голосовых складок представится как

$$J_y (E + N) \frac{\partial^4 u}{\partial x^4} - N S \frac{\partial^2 u}{\partial x^2} + c_n(x) u + r \frac{\partial u}{\partial t} + \rho_T \frac{\partial^2 u}{\partial t^2} = F(x, t). \quad (2.11)$$

При тех геометрических соотношениях, которыми обладают голосовые складки, необходимо учитывать также и силы инерции, возникающие при вращении поперечного сечения, равные $J_y \rho_T \frac{\partial^4 u}{\partial x^2 \partial t^2}$, и распределенные силы, вызываемые сдвигом и равные

$$\frac{E J_y \rho_T}{\eta G} \frac{\partial^4 u}{\partial x^2 \partial t^2} + \frac{J_y \rho_T^2}{\eta G} \frac{\partial^4 u}{\partial t^4},$$

где G — модуль сдвига; η — коэффициент формы поперечного сечения. Эти члены существенно усложняют анализ дифференциального уравнения, но, к счастью, существует возможность учесть их влияние в виде поправок на собственные значения [2.10].

Как и всякое дифференциальное уравнение в частных производных, (2.11) требует дополнения в виде начальных и граничных условий. Закрепление голосовых складок на щитовидном хряще с достаточной степенью точности принимается жестким [2.25], что соответствует граничным условиям при $x=0$

$$u(0, t) = 0; \quad \left. \frac{\partial u(x, t)}{\partial x} \right|_{x=0} = 0. \quad (2.12)$$

Задний конец складок, прикрепленный к черпаловидным хрящам, по-видимому, может смещаться с амплитудой, зависящей от частоты колебаний складок и напряжения перстне-черпаловидных мышц. Такие смещения приведут к параметрическим эффектам, но в упрощенной модели можно принять с некоторым приближением, что закрепление передних концов складок также жесткое:

$$u(l_{г.с}, t) = 0; \quad \left. \frac{\partial u(x, t)}{\partial x} \right|_{x=l_{г.с}} = 0. \quad (2.13)$$

В гл. 1 обсуждалась возможная роль переменного натяжения в возбуждении автоколебаний голосовых складок. Такое параметрическое возбуждение обычно требует двойной частоты изменения силы натяжения по сравнению с частотой возбуждаемых колебаний и приводит к уравнениям типа уравнения Хилла или Матье. Медленные по сравнению с частотой основного тона F_0 изменения натяжения, формирующие интонацию, на некотором интервале времени (десятки миллисекунд) хорошо аппроксимируются линейным законом изменения натяжения во времени: $N(t) = N_0 + at$, где N_0 — начальное значение силы натяжения при $t=0$, a — скорость изменения натяжения. Как будет показано ниже, при линейном законе $N(t)$ имеем дело с разновидностью уравнения Бесселя. Для простоты допустим, что натяжение постоянно во времени, т. е. $\dot{N} = N_0$, а упругость подстилающего слоя не зависит от координаты x . Применим к однородному уравнению, полученному из (2.11) при $F(x, t) = 0$, метод разделения переменных, рассматривающийся в гл. 1. Пусть $u(x, t) = U(x)T(t)$, тогда получим систему:

$$U^{IV} - \frac{NS}{(E+N)J_y} U'' - \alpha^2 U = 0; \quad (2.14)$$

$$T'' + 2gT' + \omega^2 T = 0, \quad (2.15)$$

где

$$2g = \frac{r}{\rho_T}; \quad \omega^2 = \frac{c_n + \alpha^2(E+N)J_y}{\rho_T}.$$

Уравнение (2.14) имеет следующее решение:

$$U_n(x) = a_{1n}(\operatorname{ch} p_{1n} x - \cos p_{2n} x) + a_{2n}(\operatorname{sh} p_{1n} x - \frac{p_{2n}}{p_{1n}} \sin p_{2n} x),$$

где корни p_{1n} , p_{2n} и коэффициент a_{2n} определяются из граничных условий (2.12) и (2.13), а коэффициент a_{1n} — из условия ортонормированности собственных функций:

$$\int_0^{l_{г.с}} U_n^2(x) dx = l_{г.с}/2. \text{ Собственные}$$

числа a_n находятся как $a_n = p_{1n} p_{2n}$.

Из экспериментов с препаратами гортани известно, что удельное натяжение превышает модуль упругости на порядок, и, следовательно, в (2.14) можно принять $NS/(E+N)J_y \approx S/J_y = \text{const}$,

т. е. это соотношение почти не изменяется при переменном натяжении $N(t)$ голосовых складок. Это позволяет ограничить анализ влияния переменного натяжения одним лишь уравнением (2.15), которое в случае линейного изменения $N(t)$ представляется в следующем виде:

$$T_n'' + 2g T_n' + (b_{1n} + b_{2n} t) T_n = 0, \quad (2.16)$$

где

$$b_{1n} = \frac{c_n + \alpha_n^2 (N_0 + E) J_y}{\rho_T}; \quad b_{2n} = \frac{\alpha_n^2 a J_y}{\rho_T}.$$

Подстановкой $T_n = \bar{T}_n e^{-gt}$ и заменой переменной $\tilde{t} = b_{1n} + b_{2n} t$ приведем (2.16) к каноническому виду уравнения Бесселя

$$\bar{T}_n'' + \frac{\tilde{t}}{b_{2n}^2} \bar{T}_n = 0,$$

которое имеет решение [2.2]

$$\bar{T}_n = \sqrt{\tilde{t}} Z_{\frac{1}{3}}(\tilde{t}), \quad (2.17)$$

где

$$\tilde{t} = \frac{2\sqrt{\tilde{t}^3}}{3b_{2n}}; \quad Z_{\frac{1}{3}}(\tilde{t}) = C_{1n} J_{\frac{1}{3}}(\tilde{t}) + C_{2n} N_{\frac{1}{3}}(\tilde{t}).$$

Здесь $J_{\frac{1}{3}}(\tilde{t})$ — цилиндрическая функция первого рода:

$$J_{\frac{1}{3}}(\tilde{t}) = \sum_{k=0}^{\infty} \frac{(-1)^k \left(\frac{\tilde{t}}{2}\right)^{\frac{1}{3} + 2k}}{k! \Gamma\left(\frac{4}{3} + k\right)};$$

Γ — гамма-функция: $\Gamma\left(\frac{4}{3} + k\right) = \frac{4}{3} \cdot \frac{7}{3} \cdots \frac{4+k}{3}$;

$N_{\frac{1}{3}}(\tilde{t})$ — цилиндрическая функция второго рода:

$$N_{\frac{1}{3}}(\tilde{t}) = \frac{J_{\frac{1}{3}}(\tilde{t}) \cos \frac{\pi}{3} - J_{-\frac{1}{3}}(\tilde{t})}{\sin \frac{\pi}{3}}.$$

Постоянные C_{1n} и C_{2n} определяются из начальных условий. При больших значениях аргумента $\tilde{t} \gg 1$ справедливо асимптотическое представление функций Бесселя [2.12]:

$$J_v(\tilde{t}) = \sqrt{\frac{2}{\pi \tilde{t}}} \cos \left(\tilde{t} - \frac{v\pi}{2} - \frac{\pi}{4} \right);$$

$$N_v(\tilde{t}) = \sqrt{\frac{2}{\pi \tilde{t}}} \sin \left(\tilde{t} - \frac{v\pi}{2} - \frac{\pi}{4} \right).$$

Если натяжение голосовых складок не меняется во времени ($a=0$), то (2.16) имеет решение

$$T_n(t) = (C_{1n} \cos \omega_n t + C_{2n} \sin \omega_n t) e^{-gt}. \quad (2.18)$$

Выражения (2.17) и (2.18) представляют собой общие решения однородного уравнения, полученного из (2.14) при $F(x, t)=0$. Это свободные колебания системы, совершаемые при ненулевых начальных условиях (об определении коэффициентов C_{1n} и C_{2n} см. § 1.3). Сила $F(x, t)$, действующая на каждую голосовую складку, подобна аэродинамической подъемной силе, возникающей при обтекании крыла самолета набегающим потоком воздуха. Эта сила, как мы видели, может быть описана компонентами статического давления и сил, вызываемых эффектом Бернулли в голосовой щели:

$$F(x, t) = [\Delta P - \kappa_B \frac{\rho_0 v^2(x, t)}{2}] h, \quad (2.19)$$

где ΔP — перепад подсвязочного и надсвязочного давлений; κ_B — коэффициент подъемной силы; v — линейная скорость воздушного потока через голосовую щель; h — толщина профиля складок вдоль оси y . Как было показано в § 2.4, скорость воздушного потока v есть функция площади голосовой щели $S_{г.щ}$ и, таким образом, в целом неоднородное уравнение (2.11) является нелинейным, поскольку вынуждающая сила зависит от переменной u . Анализируя вынужденные движения в одномерной модели, прежде всего отмечаем, что различиями в скорости воздушного потока для разных координат x согласно [2.55] можно пренебречь и, таким образом, $F(x, t) = F(t)$, т. е. в первом приближении можно принять, что аэродинамические силы распределены равномерно вдоль голосовых складок. Далее площадь голосовой щели определяется отклонениями каждой складки от положения равновесия. Принимая складки идентичными, можно записать

$$S_{г.щ}(t) = 2 \int_0^{l_{г.щ}} u(x, t) dx,$$

где $u(x, t)$ — решение уравнения (2.11). Заметим, что площадь голосовой щели зависит только от времени и не зависит от координаты x . Поэтому метод Фурье применим и к неоднородному уравнению (2.11), а парциальные возбуждающие воздействия для каждой гармоники колебаний голосовых складок описываются следующим образом:

$$F_n(T_n, t) = \frac{2 b_{1n}}{\rho_T l_{г.с}} \int_0^{l_{г.с}} F(u, x, t) u_n(x) dx.$$

Следовательно, для вынужденных колебаний временные моды описываются нелинейными уравнениями

$$T_n'' + 2g T_n' + \omega_n^2 T_n = F_n(T_n, t). \quad (2.20)$$

Если учитывать силы поверхностного сцепления, то в (2.19) должен быть добавлен член (со знаком минус) $F_{пс}$, пропорциональный площади соприкосновения складок. Эта сила замедляет расхождение складок, и она быстро убывает по мере раскрытия голосовой щели.

Нелинейные уравнения (2.20) можно решать различными способами, в том числе и методом последовательных приближений: сначала скорость воздушного потока принимается постоянной $v = v_0$ и определяется площадь голосовой щели $S_{г.щ0}(t)$, затем вычисляется переменная скорость воздушного потока $v_1(t)$ и т. д. На рис. 2.14 показаны две итерации процесса решения (2.11). Видно, что уже вторая итерация дает решение, обладающее всеми характерными особенностями нелинейного процесса.

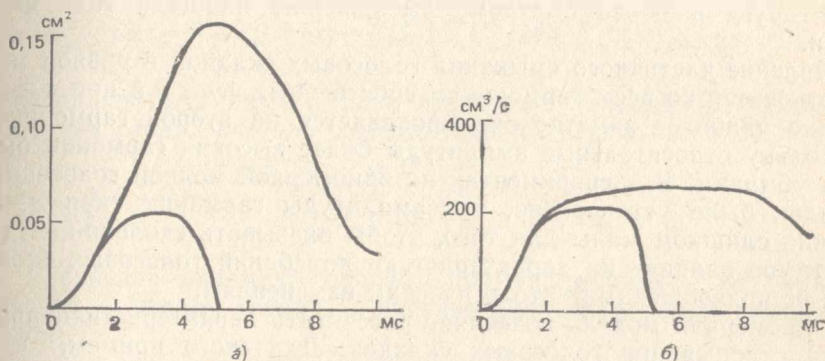


Рис. 2.14. Площадь голосовой щели (а) и объемная скорость (б) (две итерации)

Учитывая возможное изменение во времени силы натяжения голосовых складок, приводящее к решениям через бесселевы функции, частное решение неоднородного уравнения (2.20) на каждом i -м шаге итерации будем искать как

$$T_n^{(i)}(t) = N \frac{1}{3}(\tilde{t}) \int \frac{J_{\frac{1}{3}}(\tilde{t}) F_n(T_n^{(i-1)}, t)}{W} dt - \\ - J_{\frac{1}{3}}(\tilde{t}) \int \frac{N \frac{1}{3}(\tilde{t}) F_n(T_n^{(i-1)}, t)}{W} dt + C_{1n} J_{\frac{1}{3}}(\tilde{t}) + C_{2n} N \frac{1}{3}(\tilde{t}).$$

где детерминант Вронского

$$W = J_{\frac{1}{3}}(\tilde{t}) N'_{\frac{1}{3}}(\tilde{t}) - N_{\frac{1}{3}}(\tilde{t}) J'_{\frac{1}{3}}(\tilde{t})$$

по [2.2];

$$W = -2 \sin \frac{\pi}{3} / \pi \tilde{t}.$$

Четные гармоники собственных функций U_n имеют суммарную площадь отклонения от равновесного положения, равную нулю, и, казалось бы, не влияют на площадь голосовой щели. Однако при малых ее раскрытиях те отклонения гармоник, которые направлены навстречу друг другу, не могут проникнуть далее линии равновесного положения, поскольку складки встречаются. В результате складки оказываются сомкнутыми у одного конца и разомкнутыми — у другого. Такие конфигурации голосовой щели, наблюдаемые при просмотре скоростных кинофильмов, создают нелинейный эффект, влияние которого на конечных стадиях схлопывания складок приводит к ускорению их движения навстречу друг другу и более быстрому уменьшению площади голосовой щели.

Явление частичного смыкания голосовых складок в равной мере относится ко всем гармоникам собственных функций при $n \geq 2$, однако наиболее заметно оно проявляется на второй гармонике, поскольку относительные амплитуды более высоких гармоник быстро убывают. В экспериментах на одномерной модели голосовых складок было установлено, что амплитуды гармоник при $n > 3$ обычно слишком малы для того, чтобы оказывать сколько-нибудь ощутимое влияние на характеристики колебаний голосовых складок, не подверженных патологическим изменениям.

Одномерная модель позволяет рассчитать характеристики процесса схлопывания голосовых складок. Для этого примем, что поперечное сечение складок в плоскости $zoу$ имеет форму трапеции (см. рис. 2.12), причем для простоты допустим, что геометрические параметры находятся в следующих соотношениях: $ob = 2cb = 2h$ (обычно принимают $oa = d = 0,6$ см; $h = 0,4$ см). При схлопывании складки соприкасаются вдоль меньшего основания трапеции и дальнейшее движение наружной поверхности в направлении оси Z невозможно. Однако центры масс поперечных сечений еще в течение некоторого времени продолжают двигаться навстречу друг другу. Смещение центров масс от равновесного положения сопровождается деформацией профиля поперечного сечения, изгибом средней линии и потерями энергии, в результате чего складки расходятся с меньшей скоростью, чем скорость столкновения.

Примем, что в нейтральном положении складки соприкасаются, и обозначим через H расстояние, на которое сближаются центры масс их поперечных сечений в плоскости $zoу$ при некоторой

координате $x=x_0$ после удара складок. В момент встречи складок $H=0$; скорость $H'=2v_0$, т. е. удвоенной скорости движения каждой складки v_0 ; кинетическая энергия соударения $E^{(0)}_{кин}=\rho_T v_0^2$, а кинетическая энергия сжатия профиля $E^{(сж)}_{кин}=\rho_T (H'/2)^2$, потенциальная энергия сжатия

$$E_{пот}^{(сж)} = \frac{(E+N)h}{d} H^2 \left(1 + \frac{h}{3d} \right) \approx \frac{(E+N)h}{d} H^2,$$

поскольку $h \ll 3d$, потенциальная энергия изгиба

$$E_n^{(и)} = J_y (E+N) \int_0^{l_{г.с}} [f''(x)]^2 dx + NS_x \int_0^{l_{г.с}} [f'(x)]^2 dx,$$

где S_x — площадь поперечного сечения складки в плоскости $zoу$; $f(x)$ — геометрическое место точек центров масс каждого поперечного сечения.

Допустим, что $f(x)$ пропорциональна первой собственной функции $U_1(x)$, амплитуда прогиба равна $H/2$, и рассмотрим деформации поперечного сечения при $x_0=l_{г.с}/2$. Умножим (2.14) на $U_1(x)$, возьмем интеграл от 0 до $l_{г.с}$ и примем во внимание, что равенство

$$\int_0^{l_{г.с}} U_n^2(x) dx = \frac{l_{г.с}}{4} [U_n^2(x) - 2U'_n(x)U''_n(x) + U_n'^2(x)]_{x=l_{г.с}}$$

справедливо для любого типа граничных условий [2.10]. Тогда получим выражение для потенциальной энергии изгиба:

$$E_{пот}^{(и)} \Big|_{x=\frac{l_{г.с}}{2}} = \frac{J_y (E+N)}{8} \left[\frac{\alpha_1 H}{U_1 \left(\frac{l_{г.с}}{2} \right)} \right]^2.$$

Уравнение динамики прогиба при ударе найдем из баланса энергий с учетом трения и противодействия подсызочного давления;

$$\frac{d}{dt} (E_{кин} + E_{пот}) + E_T = 0,$$

откуда имеем

$$H'' + \frac{2r_{эк}}{\rho_T} H' + \omega_c^2 H = -\frac{\Delta P h}{\rho_T},$$

где $r_{эк}$ — эквивалентный коэффициент вязкого трения, и

$$\omega_c = \frac{4(E+N)}{\rho_T} \left\{ \frac{h}{d} + \frac{J_y}{8} \left[\frac{\alpha_1}{U_1 \left(\frac{l_{г.с}}{2} \right)} \right]^2 \right\}.$$

Отсюда можно найти и период T_c , во время которого складки находятся в сомкнутом состоянии: $T_c = \pi/\omega_c$. Величина T_c оказывает важное воздействие на восприятие характеристик голосового источника. На этот параметр наибольшее влияние оказывает по-

тенциальная энергия сжатия. Эквивалентный коэффициент трения [2.56] $r_{\text{эк}} = 2\xi \sqrt{mc}$, где m — масса; c — упругость складок; ξ — относительный декремент затухания, оценки которого приводились в § 2.2.

Одномерная распределенная модель, несмотря на свою относительную простоту, описывает многие явления, сопутствующие колебаниям реальных голосовых складок, в том числе и влияние физиологического тремора мышц гортани, изменяющее натяжение и эквивалентную жесткость складок. В экспериментах с синтетической речью была выявлена важная роль небольших случайных изменений периода и амплитуды от импульса к импульсу голосового возбуждения в субъективной оценке натуральности речевого сигнала [2.48, 2.83]. Попытки анализа взаимоотношений между изменениями периода и амплитуды импульсов в этих работах не привели, однако, к установлению ясной зависимости. В основе этого явления, по-видимому, лежат два механизма: случайное изменение надсвязочного давления за счет вихревых потоков воздуха в морганиевом желудочке, как это обсуждалось в § 2.4, и физиологической тремор мышц гортани. В [2.14] были обнаружены колебания натяжения перстне-черпаловидной мышцы, причиной которых является сравнительно малое число двигательных единиц в этой мышце, в результате чего происходит неполное пространственно-временное сглаживание импульсов их сокращения. Введение физиологического мышечного тремора в одномерную модель в виде случайных изменений натяжения голосовых складок позволило продемонстрировать следующие эффекты. Если относительная амплитуда случайных отклонений от исходной величины натяжения не более 1—1,5%, то перцептивное влияние его неощутимо. Если отклонения превышают 5%, то синтетическая речь приобретает резкий, грубоватый оттенок. В диапазоне отклонений 2—4% наблюдается улучшение натуральности звучания. Это совпадает с заключением других исследователей о диапазоне наибольшего влияния случайных отклонений, однако в одномерной модели изменения амплитуды и периода колебаний, в том числе и отношения длительности интервала, в течение которого голосовая щель открыта, к интервалу, в течение которого голосовая щель закрыта, происходят согласованно. Кроме того, эти изменения сопровождаются изменениями формы импульсов голосового источника.

2.7. ДВУХМЕРНАЯ МОДЕЛЬ

Как описывалось в § 2.3, помимо изгибных колебаний вдоль голосовых складок наблюдаются и деформации в поперечной — фронтальной — плоскости. Для описания этих деформаций необходимо рассмотреть двухмерную модель колебаний голосовых складок. В этом случае, принимая во внимание слоистую

структуру и геометрические размеры складок, их можно представить в виде одной или двух плит, лежащих в плоскости $хоу$. Элементарная теория ненапряженной плиты дает следующее уравнение упругих колебаний [2.9]:

$$\frac{E d^3}{2\rho(1-\nu^2)} \nabla^4 \xi + \frac{\partial^2 \xi}{\partial t^2} = F(x, y, z), \quad (2.21)$$

где ξ — функция смещения, аналогичная функция Эри (см. § 1.2); E — модуль упругости; d — толщина пластины; ρ — плотность материала; ν — коэффициент Пуассона; $F(x, y, t)$ — распределенная сила, действующая перпендикулярно к плоскости $хоу$;

$$\nabla^2 = \frac{\partial^2}{\partial x^2} + \frac{\partial^2}{\partial y^2}.$$

Поскольку длина голосовых складок значительно превышает остальные размеры, мы имеем дело с плоской задачей теории упругости.

Сложность решения уравнений типа (2.21) зависит от конфигурации пластины и вида граничных условий. Если бы форма поперечного сечения складок была прямоугольной, то решение этого дифференциального уравнения не составляло бы труда для некоторых видов граничных условий. Но трапецевидная (а фактически более сложная) форма сечения существенно усложняет процесс решения и заставляет искать упрощенные подходы. Один из таких подходов описан в [2.8] и состоит в исключении из уравнения члена со смешанной производной $\partial^4 \xi / \partial x^2 \partial y^2$, что позволяет воспользоваться преимуществами метода разделения переменных. Это, конечно, вносит в решения уравнения некоторую погрешность, но она сопоставима с погрешностью допущения прямоугольной формы поперечного сечения и в то же время значительно облегчает поиск решений.

Допустим сначала, что поверхностный и внутренний слои голосовых складок составляют одно целое, и рассмотрим колебания одной пластины в плоскости $зоу$ в направлении оси z . Учитывая влияние натяжения N , описанное ранее, составим уравнение колебаний пластины:

$$(E + N) \left(J_y \frac{\partial^4 \xi}{\partial x^4} + J_x \frac{\partial^4 \xi}{\partial y^4} \right) - N \left(S_x \frac{\partial^2 \xi}{\partial x^2} + S_y \frac{\partial^2 \xi}{\partial y^2} \right) + \rho_T \frac{\partial^2 \xi}{\partial t^2} + r \frac{\partial \xi}{\partial t} + c_n \xi = F(x, y, t),$$

где J_x, J_y — моменты инерции сечений относительно осей x и y ; S_x, S_y — площади сечений в плоскостях $зоу$ и $хоу$; ρ_T — плотность тканей; r — коэффициент вязкого трения; c_n — коэффициент упругой деформации подстилающего слоя; $F(x, y, t)$ — внешняя сила. В этом уравнении изгибные деформации представлены первым членом, сила натяжения — вторым, третий член учитывает инерцию, четвертый — вязкое трение в тканях, пятый — упругие деформации. Используя метод Фурье, запишем $\xi(x, y, t) =$

$=X(x)Y(y)T(t)$ и разделим временные и пространственные переменные, получив при этом систему:

$$\rho_T T'' + r T' + (c_{\Pi} + \lambda^2) T = 0; \quad (2.22)$$

$$J_y X^{IV} - \frac{NS_x}{E+N} X'' - \beta^2 X = 0; \quad (2.23)$$

$$J_x Y^{IV} - \frac{NS_y}{E+N} Y'' - \alpha^2 Y = 0, \quad (2.24)$$

где $\lambda^2 = \alpha^2 + \beta^2$.

Граничные условия для (2.23) снова (так же, как и для одномерной модели) принимаем жесткими:

$$X(0) = X'(0) = 0; \quad X(l_{r.c.}) = X'(l_{r.c.}) = 0. \quad (2.25)$$

Тогда собственные функции (2.24) есть

$$\Phi_n(x) = a_{1n} (\operatorname{ch} p_{1n} x - \cos p_{2n} x) + a_{2n} \left(\operatorname{sh} p_{1n} x - \frac{p_{2n}}{p_{1n}} \sin p_{2n} x \right),$$

где корни p_{1n} , p_{2n} и коэффициент a_{2n} определяются из граничных условий (2.25), а коэффициент a_{1n} находим из условия ортонормированности функций $\Phi_n(x)$. Собственные числа $\beta_n = p_{1n} p_{2n}$.

Граничные условия для (2.24) определить сложнее, однако в соответствии с принципом Сен-Венана погрешность в граничных условиях для точки o мало скажется на форме колебаний наиболее важного для процесса фонации участка cb . Исходя из опыта моделирования поведения тканей примем для точки o нулевые смещения и моменты сопротивления повороту. В точке b перерезывающие усилия, по-видимому, можно принять равными нулю, а соотношение между степенью податливости и моментом сопротивления вращению запишем через некоторый коэффициент γ . Тогда граничные условия запишутся так:

$$Y(0) = Y''(0) = 0; \quad \gamma Y(b) + (1 - \gamma) Y''(b) = 0; \quad Y'''(b) = 0. \quad (2.26)$$

Отсюда найдем собственные функции для (2.24)

$$\Psi_m(y) = a_{3m} \operatorname{sh} q_{1m} y + a_{4m} \sin q_{2m} y.$$

В результате функцию деформации поверхности складок можно записать так:

$$\xi(x, y, t) = \sum_{n,m} \Phi_n(x) \Psi_m(y) T_{n,m}(t),$$

где

$$T_{n,m}(t) = (C_{1n,m} \cos \omega_{n,m} t + C_{2n,m} \sin \omega_{n,m} t) e^{-gt};$$

$$\omega_{n,m} = \sqrt{\frac{\lambda_{n,m}^2 + c_{\Pi}}{\rho_T}} - g^2; \quad g = \frac{r}{2\rho_T}.$$

Коэффициенты $C_{1n,m}$ и $C_{2n,m}$, как обычно, определяются из начальных условий. Вынуждающая сила для каждой гармоники (n, m)

$$F_{n,m}(t) = \frac{4}{\rho_T l_T l_{r,c}} \int_0^{l_T} \int_0^{l_{r,c}} F(x, y, t) \varphi_n(x) \cdot \psi_m(y) dx dy,$$

где $l_T = ob$.

На интервале, в течение которого голосовая щель открыта, перепад давлений ΔP действует на всю площадь складок, а сила от динамического эффекта Бернулли $F_B(t)$ — главным образом на участок cb , т. е.

$$F_{n,m}^{(от)}(t) = \frac{4 I_n^{(l_{r,c})}}{\rho_T l_{r,c}} [\Delta P I_m^{(l_T)} - \kappa F_B(t) I_m^{(cb)}],$$

где

$$I_n^{(l_{r,c})} = \int_0^{l_{r,c}} \varphi_n(x) dx; I_m^{(l_T)} = \int_0^{l_T} \psi_m(y) dy; I_m^{(cb)} = \int_c^b \psi_m(y) dy;$$

κ — коэффициент области влияния эффекта Бернулли.

Уравнение вынужденных движений для (2.22) остается нелинейным, как и для одномерной модели, поскольку сила $F_B(t)$ зависит от квадрата площади поперечного сечения голосовой щели $S_{г.щ}$, однако последняя зависит не только от изгибных колебаний вдоль голосовых складок, но и от деформаций их профиля:

$$S_{г.щ}(t) = \min_y S_{г.щ}(y, t) = 2 \sum_{n,m} T_{n,m}(t) \varphi_m(y_0) \int_0^{l_{r,c}} \varphi_n(x) dx,$$

где y_0 — координата, при которой проекция голосовой щели на плоскость $хоz$ имеет минимальную площадь.

Эксперименты с двухмерной моделью показывают, что для получения реального диапазона частот основного тона необходимо повышение модуля упругости связок на порядок за счет их натяжения. Согласно гл. 1 подобное увеличение жесткости тканей является обычным в процессах сокращения мышц, следовательно, двухмерная модель не выходит за рамки физических ограничений, наложенных на входные параметры моделируемого объекта. Процессы схлопывания складок в двухмерной модели в первом приближении те же, что и в одномерной.

Исследование поведения модели голосового источника производилось на ЭВМ при различных геометрических параметрах складок, значениях перепада воздушного давления и натяжения складок. Использовались две собственные функции $\varphi_1(x)$ и $\varphi_3(x)$ и две собственные функции $\psi_1(y)$ и $\psi_2(y)$ (рис. 2.15). При $oa = 0,6$ см; $bc = oc = 0,4$ см; $l_{r,c} = 1,6$ см; $2 \cdot 10^2$ Па $\leq \Delta P \leq 1,6 \cdot 10^3$ Па и $3 \leq N/E \leq 24$ частота основного тона попадает в диапазон 55 Гц $\leq F_0 \leq 160$ Гц, а максимальная площадь раскрытия голосовой щели — в диапазон 0,03 см² $\leq S_{г.щ} \leq 0,14$ см². Если $\Delta P \leq 2 \cdot 10^2$ Па и $N/E \geq 12$, то автоколебания не возникают. Наибольшее влияние на частоту основного тона ΔP оказывает в диапазоне $(4-6) \cdot 10^2$ Па при наибольшем натяжении $N/E = 24$; оно состоит в 4% F_0 на приращение перепада давления в $2 \cdot 10^2$ Па. При уста-

новившихся процессах отношение длительности интервала открытой щели $T_{от}$ к длительности интервала закрытой щели T_c лежит в диапазоне $0,8 \leq T_{от}/T_c \leq 0,9$. На рис. 2.16, 2.17 показаны зависимости установившихся значений E_0 , $T_{от}$ и $S_{г.щ \max}$ от N/E .

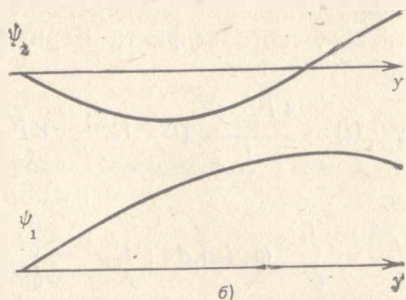
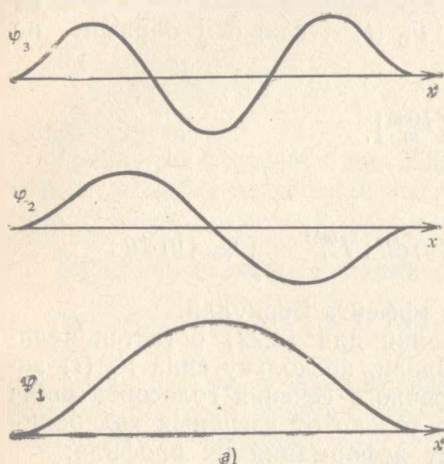


Рис. 2.15. Собственные функции голосовых складок:
а — продольные; б — поперечные

Эти статические величины, однако, не исчерпывают свойств модели.

Форма голосовой щели изменяется при движении складок в плоскостях как xoz , так и zoy (рис. 2.18). Полученные формы голосовой щели аналогичны зарегистрированным путем кино- и рент-

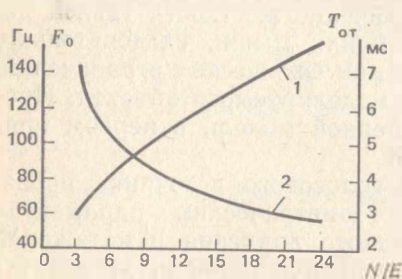


Рис. 2.16. Частота основного тона F_0 (1) и длительность интервала открытой голосовой щели (2) в зависимости от натяжения складок, $\Delta P = 8 \cdot 10^2$ Па

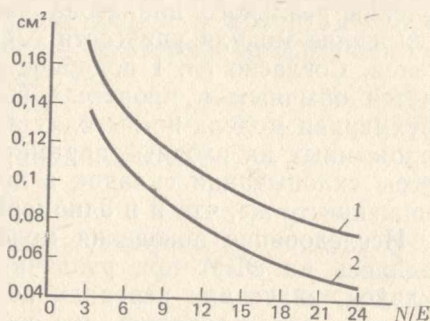


Рис. 2.17. Зависимость максимальной площади голосовой щели $S_{г.щ}$ от натяжения складок: 1 — $\Delta P = 8 \cdot 10^2$ Па; 2 — $\Delta P = 4 \cdot 10^2$ Па

геносъемки. Моделирование, однако, показало, что для образования правдоподобных форм колебаний складок в плоскости zoy необходимо допустить, что упругим деформациям в этой плоскости подвергается, главным образом, наружный слой, толщина которо-

го около $0,1 \text{ см}^1$. При этом коэффициент γ в граничных условиях (2.26) должен быть примерно $0,1$. Выяснилось также, после начала автоколебаний по всем выходным параметрам модели

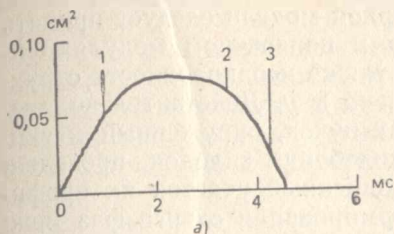
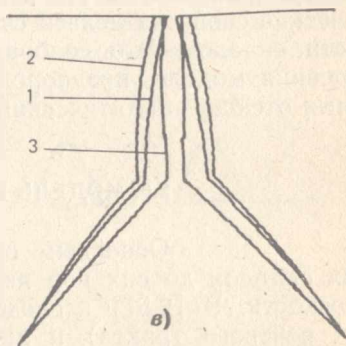
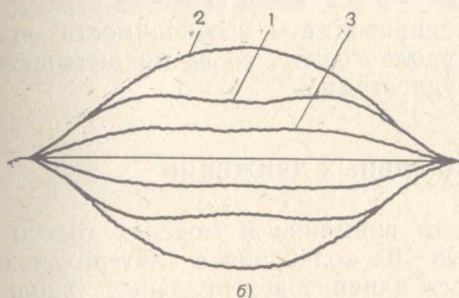


Рис. 2.18. Площадь голосовой щели (а) и формы голосовых складок в латеральной (б) и фронтальной (в) плоскостях: 1, 2, 3 — различные моменты времени



F_0 , $S_{г.щ \text{ max}}$, $T_{от}$ и амплитудам гармоник (n , m) — наблюдаются переходные процессы на протяжении трех—семи импульсов. На рис. 2.19 изображены переходные процессы по $S_{г.щ \text{ max}}$ и F_0 для первых пяти импульсов при $\Delta P = 8 \cdot 10^2 \text{ Па}$ и $N/E = 15$. Поскольку время, на котором складки сомкнуты, зависит только от N/E , при переходных процессах меняется не только частота основного тона F_0 , но и отношение $T_{от}/T_c$, причем в большей степени, чем F_0 . Соотношения между амплитудами гармоник (n , m) изменяются как при различных параметрах ΔP и N/E , так и в течение переходных процессов, вследствие чего изменяется и форма импульсов объемной скорости. Анализируя поведение двух-

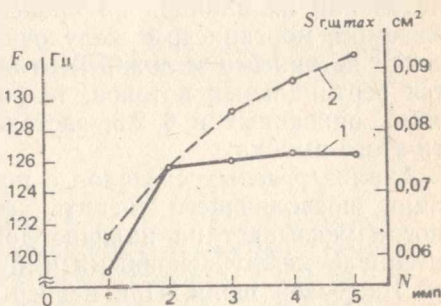


Рис. 2.19. Переходные процессы для частоты основного тона (1) и площади щели $S_{г.щ \text{ max}}$ (2)

¹ Предполагается, что основная масса складки колеблется в плоскости $хоz$, а тонкий поверхностный слой накладывает на эти колебания деформации в плоскости $гоу$.

мерной модели в плоскости zoy , можно заметить, что амплитуда колебаний все же слишком мала по сравнению с экспериментально наблюдаемыми. Аналогичное явление существует и в трехмерной модели [2.105].

При дальнейшем развитии двухмерной модели следует принять во внимание различие в продольном и поперечном модулях упругости тканей голосовых складок, а также неоднородность структуры складок вдоль оси не только z , но и y . Действительно, мышечные волокна образуют в складках пучок, окруженный мягкими тканями. Из анализа механики колебаний складок, проведенного в § 2,5, следует, что нижний скошенный участок их профиля принимает активное участие в формировании сдвига фаз между движениями верхней и нижней кромок. Но мышечные волокна уже не проникают в эту часть складок и, следовательно, ее механические свойства нельзя считать теми же, что и свойства верхней части складок. Это соображение приводит к необходимости построения модели, неоднородной вдоль оси y , со всеми вытекающими отсюда аналитическими трудностями.

2.8. МОДЕЛЬ ВЕРТИКАЛЬНЫХ ДВИЖЕНИЙ

Основным объектом внимания в моделях голосовых складок до сих пор являлись их колебания в латеральной плоскости. В [2.57] исследовалось влияние вертикальных (вдоль оси речевого тракта) и латеральных (в плоскости, перпендикулярной этой оси) потоков воздуха на акустические характеристики речевого сигнала, в результате чего было сделано заключение о пренебрежимо малом влиянии этих потоков. Можно было бы согласиться с мнением о малой роли латеральных потоков, хотя и с оговоркой о необходимости проведения дополнительных исследований их влияния на процессы вихреобразования в так называемом морганиевом желудочке (расширении речевого тракта между истинными и ложными голосовыми складками). Что касается вертикальных потоков, то ряд экспериментальных исследований, описанных в § 2.3, заставляет усомниться в справедливости этого вывода.

Анализ речевых сигналов с помощью метода обратной фильтрации, позволяющего оценить форму импульса возбуждения, иногда указывает на наличие дополнительного возбуждения в интервале между основными импульсами [2.49], т. е. при закрытой голосовой щели. При непосредственных измерениях шумовой компоненты речевого сигнала в [2.82] был обнаружен шум при закрытой голосовой щели относительно небольшого (порядка —50 дБ), но тем не менее ощутимого уровня. Появление этого шума можно объяснить только вертикальными колебаниями голосовых складок, существующими независимо от того, открыта или закрыта голосовая щель. Двухмассовая модель, на которой был получен отрицательный результат относительно влияния вер-

тикальных колебаний складок, во многих отношениях не согласована с физикой голосообразования, а перенос результатов моделирования на ней на вертикальные колебания, по-видимому, вообще неправилен. Исследования на трехмерной модели [2.105], значительно более адекватной физике голосообразования, показали, что в количественном отношении нельзя пренебрегать вкладом вертикальных движений в характеристики источника голосового возбуждения. Исходя из этих данных можно заключить, что роль вертикальных движений голосовых складок в восприятии речи может оказаться значительно большей, чем это предполагалось ранее.

Как уже говорилось, в течение периода колебаний голосовых складок каждая точка их верхней поверхности совершает сложное движение во фронтальной плоскости, напоминающее эллипс, иногда с небольшой петлей в вершине. Чем ближе рассматриваемая точка к краям складок, тем больше эксцентриситет эллипса, и отношение вертикального смещения к горизонтальному может

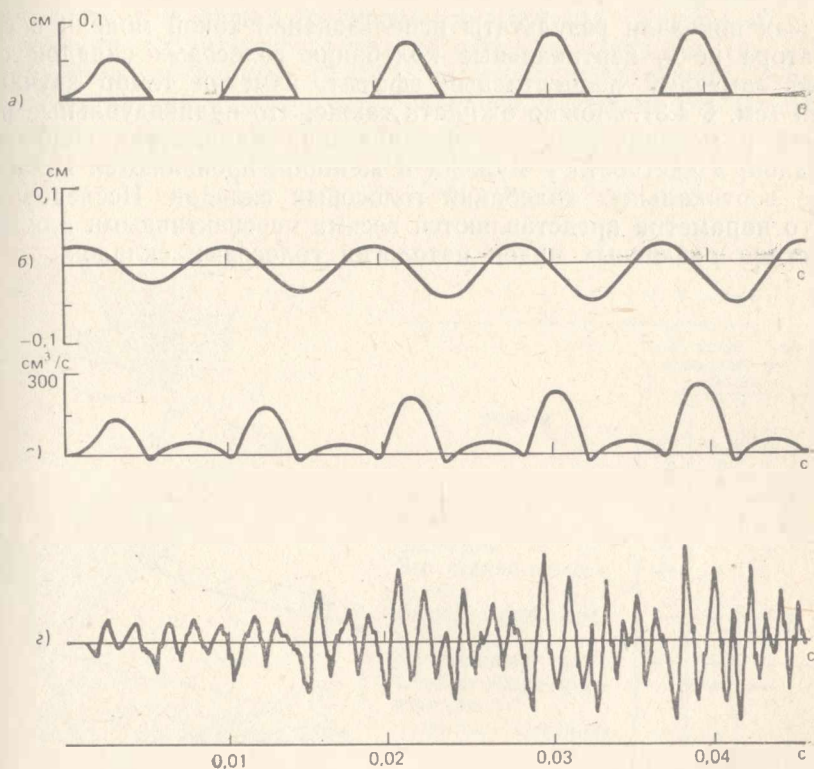


Рис. 2.20. Формы колебаний голосовых складок и импульсов голосового источника:

a — горизонтальные колебания; *б* — вертикальные колебания; *в* — объемная скорость воздушного потока на выходе из голосовой щели; *г* — синтезированный речевой сигнал (гласный А)

достигать 2—4. Анализ этих траекторий приводит к заключению, что свойство несжимаемости тканей играет меньшую роль в вертикальных движениях складок по сравнению с вертикальными компонентами сил давления и аэродинамических сил. На величину же и направление вектора аэродинамических сил, как известно, существенно влияет форма обтекаемой поверхности, и, поскольку форма профиля голосовых складок в процессе фонации непрерывно изменяется, детальное описание взаимодействий воздушного потока и складок дать чрезвычайно трудно. Поэтому моделирование трехмерных упругих деформаций складок подобно тому, как это сделано в [2.105], весьма трудоемко, и, по крайней мере на первом этапе, предпочтительно использовать более простую модель, с помощью которой можно вычислить вертикальные движения голосовых складок путем эллиптического преобразования их горизонтальных движений, генерируемых двухмерной моделью, описанной в предыдущем параграфе.

Форма импульсов, порождаемых описанной моделью, показана на рис. 2.20.

Как показали результаты использования такой модели в синтезаторе речи, вертикальные колебания голосовых складок создают заметный перцептивный эффект, изменяя тембр звучания речи (см. § 4.3). Можно ожидать также, что индивидуальные различия в механических и геометрических параметрах голосовых складок, в частности у мужчин и женщин, проявляются в характере вертикальных колебаний голосовых складок. Исследования этого параметра представляются весьма перспективными для диагностики различных видов патологии голосовых складок.

3.1. НИЖНЯЯ ЧЕЛЮСТЬ

МЫШЕЧНЫЙ АППАРАТ И КИНЕМАТИКА

Нижняя челюсть имеет сложную трехмерную форму и способна поворачиваться относительно двух осей, а также сдвигаться вперед, назад и вбок. По сравнению с большинством других артикуляторных органов динамика нижней челюсти может быть описана относительно простой моделью, поскольку сама нижняя челюсть и связанные с ней пассивные ткани представляют собой систему с сосредоточенными диаметрами. Движение нижней челюсти управляются главным образом мышцами (рис. 3.1, составленный по данным [3.25]): височной, жевательной, наружной и внутренней крыловидной, передней и задней двубрюшной, подъязычно-язычной, подбородочно-подъязычной и челюстно-подъязычной. В табл. 3.1 плюсами отмечено участие той или иной

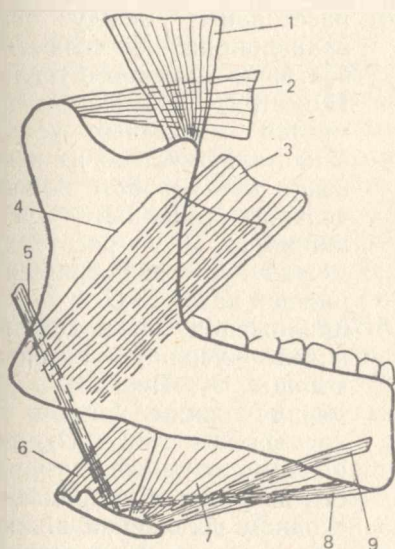


Таблица 3.1

Мышцы	Участие мышц в движении нижней челюсти			
	вверх	вниз	вперед	назад
Жевательная	+	—	+	—
Височная	+	—	—	+
Внутренняя крыловидная	+	—	+	—
Внешняя крыловидная	—	+	+	—
Двубрюшная	—	+	—	+
Подбородочно-подъязычная	—	+	—	+
Челюстно-подъязычная	—	+	—	+

Рис. 3.1. Мышцы, управляющие движениями нижней челюсти:

1 — височная; 2 — внешняя крыловидная; 3 — внутренняя крыловидная; 4 — жевательная; 5 — задняя двубрюшная; 6 — подъязычно-язычная; 7 — челюстно-подъязычная; 8 — передняя двубрюшная; 9 — подбородочно-подъязычная

мышцы в каком-либо движении, а минусами — отсутствие заметной активности [3.18, 3.25]. Помимо участия в опускании нижней челюсти двубрюшная и подбородочно-подъязычная мышцы также активизируются и на конечных стадиях ее подъема [3.19], выполняя, очевидно, тормозящую роль. Из всех мышц, управляющих движениями нижней челюсти, лишь двубрюшная и подбородочно-подъязычная создают усилия, которые можно рассматривать как сосредоточенные. Остальные же мышцы могут изменять точку и направление равнодействующей создаваемых усилий за счет активизации различных мышечных волокон. Как можно было бы заключить из рис. 3.1, мышцы в известной степени дублируют друг друга в отношении какого-либо направления движений нижней челюсти. Однако на самом деле степень активизации той или иной мышцы зависит от характера деятельности нижней челюсти: кусания, пережевывания пищи, артикуляции и т. д. — и от силовых и динамических характеристик самих мышц. Как отмечается в [2.1], например, при попытке оторвать кусок пищи использование лишь одной жевательной мышцы может привести к вывиху нижней челюсти, и в данном случае основное усилие создается височной мышцей. При разгрызании требуются большие усилия, и их развивает жевательная мышца.

Движения нижней челюсти приводят к смещениям подъязычной косточки, т. е. влияют на положение гортани и натяжение голосовых складок, как это отмечалось в гл. 2. Роль какой-либо мышцы в том или ином движении нижней челюсти определяется создаваемым ею моментом — произведением развиваемого усилия на плечо, образованное кратчайшим расстоянием между осью

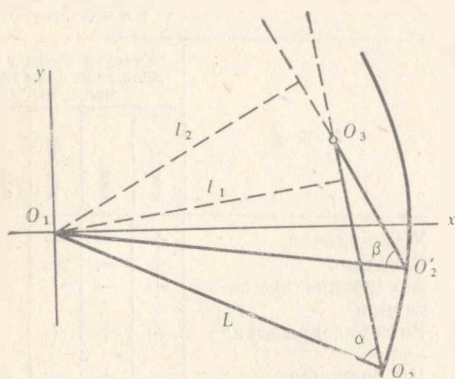


Рис. 3.2. Зависимость эффективности мышцы от поворота нижней челюсти

вращения нижней челюсти и вектором мышечного усилия. Геометрический анализ моментов мышечных усилий свидетельствует об их изменении при повороте нижней челюсти (рис. 3.2). Пусть, например, L — расстояние между точкой вращения нижней челюсти O_1 и точкой приложения равнодействующей усилий височной мышцы O_2 . Поскольку эта мышца одним концом закреплена в точке O_3 , при подъеме нижней челюсти изменяется направление

создаваемого усилия. Плечо силы в исходном положении нижней челюсти $l_1 = L \sin \alpha$, а плечо силы при повороте ее на некоторый угол $l_2 = L \sin \beta$. Отсюда видно, что при подъеме нижней челюсти угол β меньше угла α и, следовательно, момент височной мышцы уменьшается. Таким образом, система, включающая в себя ниж-

ную челюсть и прикрепленные к ней мышцы, является нелинейной. Более того, момент каждой мышцы нелинейно зависит от угла поворота нижней челюсти, так как угол между вектором усилия и линией, соединяющей точку вращения O_1 с точкой приложения усилия к нижней челюсти, весьма далек от нуля. В частности, по оценке [3.11] эффективность депрессоров (двухбрюшная и челюстно-подъязычная мышцы) падает в 2 раза при повороте нижней челюсти на 15° . В то же время эффективность жевательной мышцы и внутренней крыловидной мышцы мало меняется при повороте нижней челюсти, и относительно этих мышц система может считаться линейной.

Плечо силы височной мышцы по [3.19] составляет около 4 см, а плечо силы жевательной мышцы — около 4,8 см, так что создаваемые ими моменты при равных усилиях весьма близки. Однако жевательная мышца рассчитана на развитие больших усилий и, как будет показано ниже, весьма инерционна. Обычно считается, что жевательная мышца редко принимает участие в артикуляционных процессах [3.49], а иногда даже высказывается мнение о ее полной пассивности в речеобразовании, однако при артикуляции осмысленной речи в быстром темпе иногда наблюдается ее активность (см. § 5.7).

Задержка между началом стимуляции жевательной мышцы и подъемом нижней челюсти составляет в среднем около 35 мс (по результатам экспериментов с внешней стимуляцией, описываемых ниже). Начало активности двухбрюшной мышцы опережает опускание нижней челюсти примерно на 60—80 мс при нормальном темпе артикуляции и на 70—90 мс — при ускоренном. Длительность команд, поступающих на двухбрюшную мышцу, 195—230 мс для нормального темпа артикуляции и 130—150 мс — для ускоренного. При этом наблюдаются небольшие отклонения как в опережении, так и в длительности команд в зависимости от места слога в звукосочетании и его ударности. Наибольшая активность двухбрюшной мышцы обычно регистрируется на ударном слоге, хотя иногда наблюдается ее увеличение на последнем слоге. Эти данные получены с помощью экспериментов, проводившихся в Коннектикутском университете на носителе русского языка совместно с Т. Геом (речевой материал — бессмысленные трехсложные звукосочетания типа БАБАБА с ударением на втором слоге).

Положение, которое нижняя челюсть занимает в процессе артикуляции, зависит от контекста. В табл. 3.2 приведены отклонения нижней челюсти (в сантиметрах) от исходного состояния при сомкнутых челюстях. Первая строка этой таблицы — результаты измерений на кинорентгенограммах, причем речевой материал представлял собой изолированные звуки — губные *Б, М, П, В, Ф*, переднеязычные *Д, Ж, З, Н, С, Т, Ш*, заднеязычные *Г, К, Х* и гласные, разбитые на две группы, *А, О, Э*, которым соответствует первая цифра в графе «гласные», и *Ы, У, И*, которым соответствует вторая цифра в этой графе. Нижняя челюсть опускается ниже всего при артикуляции открытых гласных, тогда как закрытые

Таблица 3.2

Метод измерений		Отклонение нижней челюсти, см, от нейтрального положения для звуков			
		согласных			гласных
		губных	переднеязычных	заднеязычных	
Кинорентгено съемка		0,34	0,24	0,39	0,88 0,39
Оптический	Норма	0,22	0,25	0,22	0,24
	Быстро	0,19	0,21	0,19	0,24

гласные характеризуются примерно одинаковой степенью опускания с согласными. Для переднеязычных согласных нижняя челюсть занимает более высокое положение.

Кинорентгенографические измерения показывают, однако, что те же звуки в слогах имеют другие показатели положения нижней челюсти. Например, в слоге АДА положение нижней челюсти для первого А есть 0,8 см, а для Д — 0,3 см; в слоге АГА для первого А — 0,88 см, для Г — 0,88 см; в слоге АБА для первого А — 1 см, а для Б — 0,62 см. Как видно, положение нижней челюсти для первого А в этих слогах осталось примерно тем же, что и для изолированного произнесения этого гласного, но показатели для согласных Б, Д, Г изменились в сторону большего опускания нижней челюсти — примерно в 2—3 раза по сравнению с их изолированным произнесением. Это наблюдение согласуется с мнением о том, что коартикуляция приводит к зависимости опускания нижней челюсти для согласного от степени открытости последующего гласного [3.49].

Другая серия измерений положения нижней челюсти и скорости ее движения была выполнена с помощью системы фотодатчиков, разработанной В. Эваном (Коннектикутский университет, США). Речевой материал содержал бессмысленные трехсложные сочетания с ударением на втором слоге типа СГСГСГ, где Г — гласная — была всегда одна и та же — А, а в качестве согласных использовались губные Б, П, М, В, Ф, переднеязычные Т, Д, З, Ш и заднеязычный звук Г. Каждое звуко сочетание произносилось в нормальном и ускоренном темпах артикуляции тем же испытуемым, что и в первом случае. По сравнению с изолированным произнесением вариации положения нижней челюсти по разным группам звуков гораздо меньше, причем и для гласного звука А степень опускания нижней челюсти значительно уменьшается. Ускорение темпа артикуляции уменьшает амплитуду движений нижней челюсти.

В артикуляционных моделях, учитывающих положение нижней челюсти, обычно принимается во внимание лишь ее поворот относительно горизонтальной оси. Наряду с этим кинорентгено-

графические измерения указывают на существование горизонтальных сдвигов — вперед или назад, так что в сагиттальной плоскости любая точка нижней челюсти описывает эллипс. На рис. 3.3 показаны контуры кинорентгенограмм для двух положений — артикуляции согласных *С* и *Ш*. Как видно, существует сдвиг нижней челюсти в горизонтальном направлении, причем величина сдвига, примерно равная 0,2—0,3 см, сравнима со степенью ее вертикального смещения для этих звуков. Вообще нижняя челюсть может играть в процессах артикуляции весьма важную роль, особенно для переднеязычных звуков и при ускоренном темпе артикуляции, как это будет видно ниже.

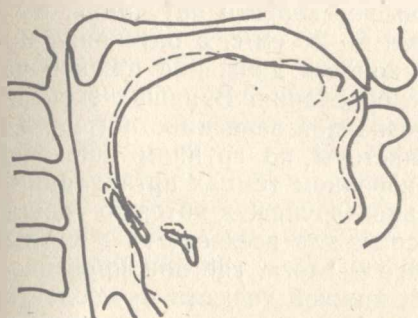


Рис. 3.3. Контурсы рентгенограмм для звуков *С* (—) и *Ш* (---)

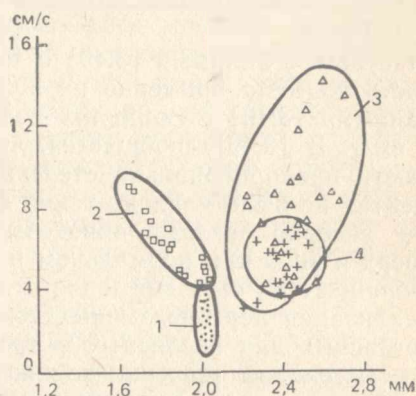


Рис. 3.4. Зависимость скорости движения нижней челюсти от амплитуды смещений

1 — медленный подъем; 2 — быстрый подъем; 3 — быстрое опускание; 4 — медленное опускание

Разброс в положении нижней челюсти, оцениваемый как относительная дисперсия $\delta = (\sigma/h) \cdot 100\%$ (σ — дисперсия; h — положение), вычисленная по всем слогам и реализациям звукосочетаний, остается в пределах 2—4% независимо от звука, порядка слога и направления движения. При увеличении скорости артикуляции относительная ошибка по положению увеличивается незначительно — менее чем на 1%. Относительная ошибка по скорости, как это видно из табл. 3.3, значительно выше — 15—25%, причем нестабильность скорости при движении вниз несколько больше, чем при движении вверх, независимо от скорости артикуляции. Правда, при этом и скорость движения вниз при любом темпе превышает скорость подъема в среднем на 40% по всем слогам. Такое различие в скорости может объясняться различиями в динамических характеристиках мышц, управляющих подъемом и опусканием нижней челюсти.

Превышение скорости опускания нижней челюсти над скоростью ее подъема, обнаруженное в наших экспериментах, не со-

Таблица 3.3

Порядок слога в звукосочета- нии	Максимальная скорость движения нижней челюсти V_{max} , см/с, и ее относительная ошибка ϵ_V , %, при							
	подъеме				опускании			
	норма		быстро		норма		быстро	
	V_{max}	ϵ_V	V_{max}	ϵ_V	V_{max}	ϵ_V	V_{max}	ϵ_V
1	3,9	15	5	25	6,3	20	8,7	25,5
2	4,6	20	6,9	18,5	5,7	21,5	7	21,5
3	2,8	15	2,4	17,5	3,7	25	4,6	25

гласуется с данными [3.49] о том, что для губных взрывных скорость подъема примерно на 20% выше скорости опускания, и с данными [3.10] о скоростях подъема 6—20 см/с, а опускания 3—5 см/с. В [3.49] сообщается, что скорость движения нижней челюсти пропорциональна степени ее опускания. В наших экспериментах не было найдено такой однозначной зависимости (рис. 3.4, где условно овалами обведены кластеры по группам движений вверх и вниз при нормальном и ускоренном темпах артикуляции). Обращает на себя внимание отдельная группа, в которую вошли реализации переднеязычных согласных для всех слогов и губных согласных для последнего — третьего — слога, где при практически постоянном значении положения нижней челюсти наблюдается разброс скорости движения в 4 раза. Подобные противоречия между различными авторами, касающиеся характеристик движений артикуляторных органов, по-видимому, отражают различия в тактике и параметрах системы управления артикуляцией у разных дикторов для различных звукосочетаний и языков.

ДИНАМИЧЕСКИЕ ХАРАКТЕРИСТИКИ

Определение динамических характеристик артикуляторных органов обычно наталкивается на серьезные трудности, связанные с тем, что большинство этих органов представляет собой системы с распределенными параметрами, причем форма возбуждающего сигнала и распределение активности управляющих мышц неизвестны, тогда как от состояния мышечной активности зависят и упругость, и вязкость. В этом отношении идентификация параметров нижней челюсти облегчается тем, что с достаточной степенью точности ее вместе с присоединенными тканями можно рассматривать как систему с сосредоточенными параметрами, описываемую обыкновенным дифференциальным уравнением либо относительно угла поворота $I\ddot{\varphi} + B\dot{\varphi} + K\varphi = M(t)$, либо относительно линейного смещения положения ее центра тяжести $m\ddot{y} + r\dot{y} + cy = F(t)$, (3.1)

где J , B и K — соответственно момент инерции, вязкость и жесткость системы; $M(t)$ — момент внешних сил; m , r и c — масса, вязкость и жесткость системы (в другой системе единиц); $F(t)$ — внешняя сила. В дальнейшем мы будем пользоваться последним уравнением.

При исследовании динамических характеристик нижней челюсти входной управляющий сигнал можно сформировать с помощью внешних механических средств, вынуждающих ее движения [3.18, 3.20, 3.21]. Другой метод создания вынужденных движений нижней челюсти — внешняя электрическая стимуляция [3.6, 3.48]. Последняя вызывает меньшую рефлекторную активность мышц-антагонистов и, следовательно, меньшее искажение нормальных значений присоединенной упругости и вязкости, поэтому этот метод обладает преимуществом перед механическим возбуждением. Известны примеры, когда за счет рефлекторного повышения упругости мышц при внешнем механическом возбуждении оценка резонансной частоты нижней челюсти оказывалась завышенной во много раз. Возможно, что этот эффект проявился и в экспериментах [3.18], когда расчетный момент инерции челюсти $J_p = 9,8 \cdot 10^{-3} \text{ кгм}^2$ оказался значительно выше измеренного на препарате нижней челюсти $J_n = 2,3 \cdot 10^{-3} \text{ кгм}^2$. Момент инерции J_p вычислялся путем подгонки вычисленного переходного процесса к измеренному. Резонансная частота в такой системе $\omega_0 = \sqrt{K/J}$, поэтому в случае повышения жесткости K для сохранения частотных характеристик необходимо увеличить и момент инерции J . Правда, при этом упадет коэффициент усиления системы, но, к сожалению, не сообщается о том, потребовалась ли его подгонка при подгонке переходных процессов.

Жевательная мышца — вполне удобный объект для электрической стимуляции. Она доступна и не изменяет своей эффективности при повороте нижней челюсти. Следовательно, применение математического аппарата для анализа линейных систем жевательной мышцы более оправданно, чем для других мышц.

Наши эксперименты заключались в записи движения нижней челюсти в вертикальном направлении при нагрузке ее пружиной и (или) массами и электрической стимуляции жевательной мышцы от внешнего генератора. Стимуляция производилась биполярно через проволочные электроды, прикрепленные к мышце. Частота следования импульсов была 100—200 с^{-1} , длительность импульсов 2—5 мс, амплитуда 4—14 В. Последовательность импульсов модулировалась другим генератором, «вырезавшим» пакеты с частотой 0,5—10 Гц, что позволяло измерять как переходные процессы нижней челюсти, так и ее амплитудно-частотную характеристику (АЧХ).

Положение нижней челюсти определялось по тени, отбрасываемой на управляемую ЭВМ систему фотодатчиков штифтом, укрепленным на нижних зубах. Другая система фотодатчиков определяла положение головы для компенсации ее движений во вре-

мя эксперимента. К штифту прикреплялись грузики массой 50 и 100 г, а также пружина с упругостью $c_n=500$ г/см, другой конец которой фиксировался на верхних зубах. Смещение штифта в вертикальной плоскости достигало 14—18 мм. Стимуляция начиналась при фиксированных начальных условиях для отклонения нижней челюсти на 18 мм. При одних и тех же параметрах возбуждения формы переходных процессов были похожи и амплитуды движений были, как правило, весьма близки. В некоторых случаях, однако, наблюдались одно—три замедления движения при переходных процессах, что было интерпретировано как рефлекторное.

На рис. 3.5 показана амплитудно-частотная характеристика $A(f)$, полученная усреднением амплитуд по всем циклам для фиксированной частоты при различных условиях нагрузки нижней челюсти. Из рисунка видно, что на уровне $A=0,1$ $A(0)$ полоса пропускания при отсутствии нагрузки — около 5 Гц. Для проверки

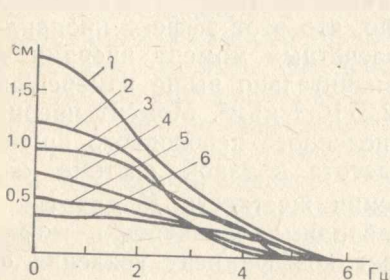


Рис. 3.5. Измеренные АЧХ нижней челюсти при различных нагрузках:

1 — без нагрузки; 2 — нагрузка 50 г; 3 — нагрузка 100 г; 4 — пружина; 5 — пружина и 50 г; 6 — пружина и 100 г

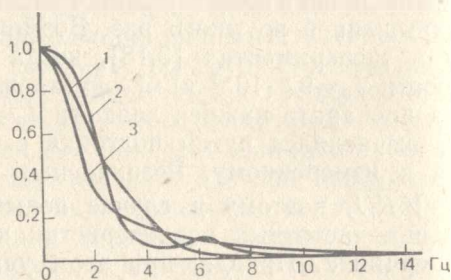


Рис. 3.6. Нормированные АЧХ системы нижней челюсти:

1 — стимуляция жевательной мышцы без нагрузки; 2 — медленные произвольные движения; 3 — быстрые произвольные движения

специфичности этой величины для данного эксперимента были измерены частоты среза для максимально быстрых произвольных движений и артикуляции в ускоренном темпе. В обоих случаях величины оказались близкими к 5 Гц, и это дало основание полагать, что внешняя стимуляция мышцы существенно не изменяет свойства периферического отдела системы управления движением. Результаты экспериментов по стимуляции мышц с присоединением упругой или инерционной нагрузки дают важную информацию о передаточных функциях исследуемых элементов системы управления движениями, однако нужно принимать во внимание условия эксперимента для того, чтобы избежать ошибочных заключений. Например, в [3.12, 3.13] для мышц кошки в режиме изометрического сокращения было найдено, что последовательное присоединение упругого элемента увеличивает коэффициент усиления систем пропорционально жесткости пружин, а

инерционная нагрузка почти не влияет на форму АЧХ мышцы в области низких частот. В дальнейшем было получено более непосредственное подтверждение того, что электрическая стимуляция жевательной мышцы почти не искажает формы амплитудно-частотного спектра передаточной функции нижней челюсти. С этой целью измерялась естественная электрическая активность жевательной мышцы при произвольных движениях нижней челюсти. С помощью быстрого преобразования Фурье были рассчитаны АЧХ входного (ЭМГ) и выходного (движения нижней челюсти) сигналов. Амплитудно-частотная характеристика передаточной функции нижней челюсти была найдена как отношение этих спектров — прием, вполне законный, поскольку мы удостоверились в том, что система «жевательная мышца — нижняя челюсть» — линейная относительно угла поворота челюсти. Подробнее о методике расчета АЧХ артикуляторных органов в естественных условиях будет рассказано в следующем параграфе. Сопоставляя полученные спектры со спектром, измеренным в экспериментах с внешней стимуляцией, обнаруживаем их большое сходство как по форме, так и в граничной частоте (рис. 3.6).

В наших экспериментах использовалось понятие передаточной функции, определяемое как отношение преобразования Лапласа выходного сигнала к преобразованию Лапласа входного сигнала. Применяя преобразование Лапласа к обеим частям уравнения (3.1) и используя отношение $X_{\text{вых}}(s)$ к $X_{\text{вх}}(s)$, где $X_{\text{вых}}$ и $X_{\text{вх}}$ — преобразования Лапласа от $y(t)$ и $F(t)$; s — комплексная частота, получаем передаточную функцию нижней челюсти W в общеизвестной форме для линейной системы второго порядка:

$$W(s) = \frac{\kappa_{\text{н.ч.}}}{T_{\text{н.ч.}}^2 s^2 + 2\xi_{\text{н.ч.}} T_{\text{н.ч.}} s + 1},$$

где $\kappa_{\text{н.ч.}} = 1/c_{\text{н.ч.}}$ — коэффициент усиления; $T_{\text{н.ч.}} = \sqrt{m/c_{\text{н.ч.}}}$ — постоянная времени; $\xi_{\text{н.ч.}}$ — относительный показатель затухания; $2\xi_{\text{н.ч.}} T_{\text{н.ч.}} = r/c_{\text{н.ч.}}$. Собственная (резонансная) частота нижней челюсти $\omega_{\text{н.ч.}} = 1/T_{\text{н.ч.}}$. Если к нижней челюсти на расстоянии l_n от точки вращения нижней челюсти присоединена некоторая нагрузка в виде пружины с упругостью c_n и массой m_n , то собственная частота

$$\omega_{\text{н.ч.}} = \sqrt{\frac{(c_{\text{н.ч.}} + c_n \alpha)}{m_{\text{н.ч.}} + m_n \alpha}}, \quad (3.2)$$

где $c_{\text{н.ч.}}$ — упругость мышц нижней челюсти; $m_{\text{н.ч.}}$ — масса нижней челюсти; $\alpha = l_n/l_{\text{ц.т.}}$; $l_{\text{ц.т.}}$ — расстояние центра тяжести нижней челюсти от точки ее вращения. Из (3.2) видно, что нагрузка в виде упругости повышает собственную частоту, а нагрузки массой ее понижают. Это свойство было использовано в экспериментах по анализу параметров системы управления с помощью электрической стимуляции жевательной мышцы и нагрузки нижней челюсти пружиной и массами. Общий вид передаточной функции си-

системы «нижняя челюсть — мышцы» можно искать в классе дробно-рациональных систем:

$$W(s) = K(X_{\text{вх}}) \frac{\prod_{n=1}^N (1 + T_n s)}{\prod_{m=1}^M (1 + T_m s)}, \quad (3.3)$$

где $K(X_{\text{вх}})$ — нелинейный коэффициент усиления всей системы, зависящий от амплитуды входного сигнала $F(t)$.

Фактор временной задержки сигналов в системе управления может оказаться существенным при некоторых обстоятельствах [3.3], однако в обработке результатов описываемых экспериментов он не учитывался.

Для анализа систем управления и идентификации их параметров по экспериментальным данным в теории автоматического управления пользуются амплитудно-частотными характеристиками, заменяя аргумент s на $j\omega$ (j — мнимая единица). Для (3.3) АЧХ есть

$$A(\omega) = K(X_{\text{вх}}) \frac{\prod_{n=1}^N |1 + j\omega T_n|}{\prod_{m=1}^M |1 + j\omega T_m|}.$$

Кроме того, применяются логарифмические амплитудно-частотные характеристики (ЛАЧХ) $L(\omega) = 20 \lg A(\omega)$, причем известно, что каждый член $1 + j\omega T$ в знаменателе передаточной функции (3.3) придает $L\omega$ наклон — 6 дБ/окт., а в числителе — +6 дБ/окт. Этим свойством ЛАЧХ пользуются для идентификации параметров передаточных функций типа (3.3), аппроксимируя ЛАЧХ ломаными прямыми с наклоном, кратным ± 6 дБ/окт., причем $T_i = 1/\omega_i$ определяются в точках излома аппроксимирующей функции [3.7].

Рассмотрим АЧХ $A^{(i)}(\omega)$, где $i=0$ соответствует отсутствию нагрузки, $i=1$ — нагрузке массой в 50 г, $i=2$ — нагрузке массой 100 г, $i=3$ — нагрузке пружиной, $i=4$ — нагрузке пружиной и массой 50 г, $i=5$ — нагрузке пружиной и массой 100 г. Известно, что при частоте $\omega=0$ $A^{(i)}(0) = B_{\text{вх}} \kappa^{(i)}$, где $B_{\text{вх}}$ — амплитуда скачкообразного возбуждения; $\kappa^{(i)}$ — общий коэффициент усиления: $\kappa^{(i)} = \kappa^{(i)}_{\text{с.у}} \kappa^{(i)}_{\text{н.ч}}$; $\kappa^{(i)}_{\text{н.ч}}$ — коэффициент усиления нижней челюсти; $\kappa^{(i)}_{\text{с.у}}$ — коэффициент усиления системы управления. Исходя из схемы эксперимента можно записать следующую систему уравнений:

$$\begin{aligned} \kappa^{(i)}_{\text{н.ч}} &= 1/c_{\text{н.ч}} \quad \text{при } i=0, 1, 2; \\ \kappa^{(i)}_{\text{н.ч}} &= 1/(c_{\text{н.ч}} + c_{\text{п}}) \quad \text{при } i=3, 4, 5; \\ \kappa^{(i)}_{\text{с.у}} &= \kappa^{(0)}_{\text{с.у}} (1 - \delta \kappa^{(i)}_{\text{с.у}}) \quad \text{при } i=1, 2, 3; \\ \kappa^{(i)}_{\text{с.у}} &= \kappa^{(0)}_{\text{с.у}} (1 - \delta \kappa^{(i)}_{\text{с.у}} - \delta \kappa^{(3)}_{\text{с.у}}) \quad \text{при } i=4, 5, \end{aligned} \quad (3.4)$$

где c_n — упругость нагрузочной пружины; $\delta\kappa^{(i)}_{c.y.}$ — изменение коэффициента усиления системы управления при нагрузке.

Вычисляя относительное изменение коэффициента усиления системы для разных условий, находим: $\delta\kappa^{(1)}_{c.y.}=0,33$; $\delta\kappa^{(2)}_{c.y.}=0,48$; $\delta\kappa^{(4)}_{c.y.}=0,32$; $\delta\kappa^{(5)}_{c.y.}=0,55$. Отсюда можно заключить, что нагрузка массой 50 г создает одинаковое изменение коэффициента усиления системы управления как при нагрузке пружиной, так и без нее, т. е. $\delta\kappa^{(1)}_{c.y.}\approx\delta\kappa^{(4)}_{c.y.}$. Это позволяет решить систему (3.4) относительно упругости, присоединенной к нижней челюсти $c_{н.ч.}$, причем это решение оказывается инвариантным к амплитуде входного воздействия $B_{вх.}$:

$$c_{н.ч.} = \frac{A^{(3)}(0) - A^{(4)}(0)}{A^{(0)}(0) - A^{(1)}(0) + A^{(4)}(0) - A^{(3)}(0)} c_n.$$

Для данного испытуемого присоединения упругость $c_{н.ч.}=315$ г/см. Определение абсолютной величины $c_{н.ч.}$ позволяет приблизительно оценить собственную частоту $f_{н.ч.}$ нижней челюсти испытуемого. Примем массу нижней челюсти как среднестатистическую величину, составляющую по [2.11] около 1,7% массы костей скелета, т. е. примерно 170 г, и добавим к ней массу присоединенных тканей, так что $m_{н.ч.}=200$ г. Допустим далее, что вся упругость присоединенных тканей сосредоточена в центре тяжести нижней челюсти, расположенном посредине между точкой вращения и наиболее удаленной от нее точкой нижней челюсти. Тогда собственная частота нижней челюсти $f_{н.ч.}\approx 6$ Гц оказывается лежащей вблизи границы экспериментально измеренной полосы пропускания системы и не может наблюдаться. Этот вывод подтверждается анализом поведения характерных частот f_i , полученных как координаты изломов кусочно-линейной аппроксимации логарифмических амплитудно-частотных характеристик отрезками прямых с наклоном, кратным -6 дБ/окт. (рис. 3.7). Из табл. 3.4 видно, что ни одна из характерных частот не увеличивается последовательно с добавлением пружины и не падает при нагрузке массой (частоты f_i даны в герцах, а постоянные времени T_i — в секундах). Следовательно, хотя форма АЧХ нижней челюсти и влияет на форму АЧХ всей системы, однако найденные экспериментальные частоты f_i характеризуют только свойства мышцы и, может быть, рецепторов обратной связи.

Результаты эксперимента позволяют вычислить еще одну величину — относительный коэффи-

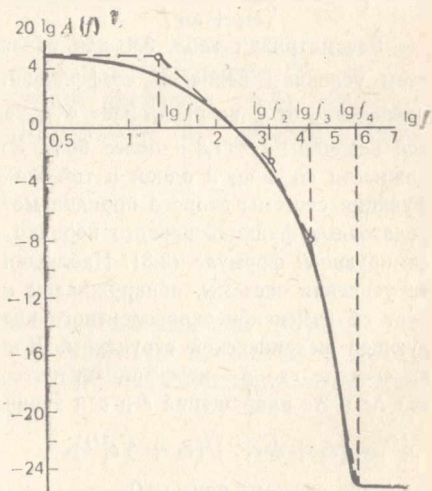


Рис. 3.7. Кусочно-линейная аппроксимация логарифмической АЧХ без нагрузки

Таблица 3.4

Параметры нижней челюсти	Зависимость идентифицированных параметров нижней челюсти от вида и величины нагрузки					
	без нагрузки	нагрузка, г		пружина	пружина +50 г	пружина +100 г
		50	100			
f_1	1,2	1,45	1,7	1,4	1,8	2,3
f_2	3	2,7	3,2	2,7	3,8	3,4
f_3	4,2	4,3	3,8	3,3	4,4	4,1
f_4	5,8	5,9	5	4,5	5,9	5,1
T_1	0,133	0,11	0,094	0,114	0,088	0,69
T_2	0,53	0,059	0,005	0,0059	0,0042	0,033
T_3	0,038	0,037	0,042	0,048	0,036	0,031
T_4	0,027	0,027	0,032	0,035	0,027	0,031
ξ	1,107	1,049	1,05	1,055	1,069	1,069
δc	—	0,5	0,9	—0,055	—	—
δr	—	0,24	0,35	—0,185	—	—

циент потерь в мышце ξ . Предположим, что первые две характерные частоты ω_1 и ω_2 есть сопряженные частоты некоторой механической структуры, описываемой обыкновенным дифференциальным уравнением второго порядка ($\omega = 2\pi f$). Тогда, как известно, между ними существуют следующие соотношения:

$$\omega_1 = \omega_0 (\xi - \sqrt{\xi^2 - 1}); \quad \omega_2 = \omega_0 (\xi + \sqrt{\xi^2 - 1}),$$

где ω_0 — собственная частота механической структуры. Отсюда

$$\xi = \sqrt{\frac{1}{1 - \left(\frac{\omega_2 - \omega_1}{\omega_2 + \omega_1} \right)^2}}.$$

Рассматривая табл. 3.4 для различных нагрузок, видим, что средняя по всем условиям величина коэффициента потерь $\bar{\xi} = 1,066$, причем наибольшее отклонение от $\bar{\xi}$ не превышает 4%, тогда как отклонение, например, f_1 от средней величины $\bar{f}_1 \approx 1,3$ — более 60%. Это оправдывает предположение о принадлежности ω_1 и ω_2 к одной и той же структуре. Поскольку $\xi > 1$, передаточная функция системы второго порядка может быть описана произведением двух передаточных функций первого порядка, что соответствует знаменателю в аппроксимирующей формуле (3.3). Наблюдая за изменением ω_1 и общего коэффициента усиления системы, обнаруживаем между ними корреляцию, что свидетельствует об изменении эквивалентного коэффициента потерь и упругости соответствующей механической структуры. Для апериодической системы первого порядка $\omega = r/c$, где r — коэффициент потерь; c — упругость системы. Обозначим через Δr и Δc приращения r и c и запишем

$$\omega_1^{(i)} = (r_1 + \Delta r_1^{(i)}) / (c_1 + \Delta c_1^{(i)}),$$

а отклик системы при $\omega = 0$

$$A^{(i)}(0) = \frac{B_{\text{вх}}^{(i)}}{(c_1^{(0)} + \Delta c_1^{(i)})(c_{\text{н.ч}} + c_{\text{н}})},$$

причем при $i=0$, $\Delta r^{(i)} = \Delta c_1^{(i)} = c_n = 0$. Тогда

$$\frac{A^{(0)}(0)}{A^{(i)}(0)} = \frac{c_1^{(0)} + \Delta c_1^{(i)}}{c_1^{(0)}} = 1 + \delta c_1^{(i)}, \quad \text{где } \delta c_1^{(i)} = \frac{\Delta c_1^{(i)}}{c_1^{(0)}},$$

$$\text{и } \frac{\omega_1^{(i)}}{\omega_1^{(0)}} = \frac{c_1^{(0)} + \Delta c_1^{(i)}}{c_1^{(0)}} \frac{r_1^{(0)}}{r_1^{(0)} + \Delta r_1^{(i)}} = \frac{1 + \delta c_1^{(i)}}{1 + \delta r_1^{(i)}},$$

$$\text{где } \delta r_1^{(i)} = \frac{\Delta r_1^{(i)}}{r_1^{(0)}}.$$

Учитывая, что

$$\frac{A^{(0)}(0)}{A^{(3)}(0)} = (1 + \delta c_1^{(3)}) \left(1 + \frac{c_n}{c_{н.ч}} \right),$$

получаем:

$$\delta c_1^{(i)} = \frac{A^{(0)}(0)}{A^{(i)}(0)} \frac{c_{н.ч}}{c_{н.ч} + c_n} - 1;$$

$$\delta r_1^{(i)} = \frac{A^{(0)}(0)}{A^{(i)}(0)} \frac{\omega_1^{(0)}}{\omega_1^{(i)}} - 1,$$

причем при $i < 3$, $c_n = 0$.

Из табл. 3.4 также видно, что инерционная нагрузка увеличивает как потери, так и упругость некоторой механической структуры в мышце, тогда как упругая нагрузка уменьшает их, как бы разгружая эту структуру. В отличие от относительного коэффициента затухания ξ , абсолютный коэффициент потерь r оказывается подверженным значительным изменениям.

Аппроксимация АЧХ четырьмя ломаными позволяет идентифицировать параметры четырех постоянных времени $T_m = 1/\omega_m$, $m=1, \dots, 4$, в знаменателе передаточной функции (3.3). Однако реконструкция АЧХ и переходных процессов по передаточной функции вида

$$W(j\omega) = \frac{\kappa}{\prod_{m=1}^4 (1 + j\omega T_m)}$$

и сравнение их с экспериментально определенными характеристиками показывают необходимость включения, по крайней мере, двух членов в числитель (3.3), т. е. требуют идентификации параметров передаточной функции вида:

$$W(j\omega) = \kappa \frac{\prod_{n=1}^2 (1 + j\omega T_{1n})}{\prod_{m=1}^4 (1 + j\omega T_m)}.$$

Процедура идентификации состояла в подборе таких параметров T_{11} , T_{12} и $\xi_{н.ч.}$, которые обеспечили бы наибольшую близость вычисленных и экспериментально измеренных АЧХ и переходных процессов, причем больший вес придавался сходству переходных процессов. Переходные процессы вычислялись как

$$G(t) = \frac{2}{\pi} \int_0^{\infty} \frac{P(\omega)}{\omega} \sin \omega t d\omega,$$

где $P(\omega)$ — действительная часть передаточной функции всей системы $W(j\omega) = P(\omega) + jQ(\omega)$. Идентифицированные параметры T_{11} , T_{12} , $\xi_{н.ч.}$ для различных нагрузок приведены в табл. 3.4, а соответствующие АЧХ и переходные процессы для случая отсутствия нагрузки показаны на рис. 3.8—3.11. Значение коэффициента затухания $\xi_{н.ч.} = 0,35 \div 0,45$ близко к величине 0,5, полученной в [3.18].

Методика описанного эксперимента позволила однозначно определить упругость тканей, присоединенных к нижней челюсти,

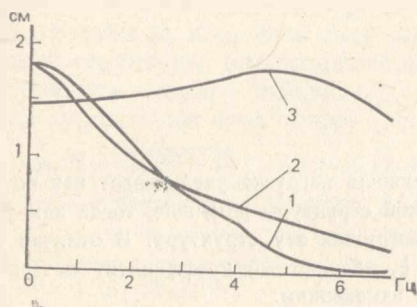


Рис. 3.8. Амплитудно-частотные характеристики системы нижней челюсти без нагрузки (движение вверх): 1 — измеренная; 2 — вычисленная; 3 — расчетная АЧХ нижней челюсти (усл. ед.)

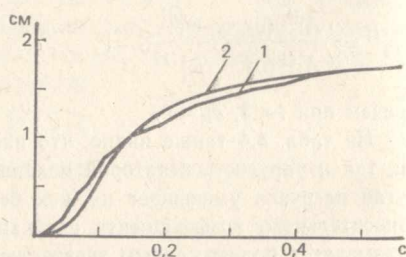


Рис. 3.9. Переходные процессы (движение вверх): 1 — измеренный; 2 — вычисленный

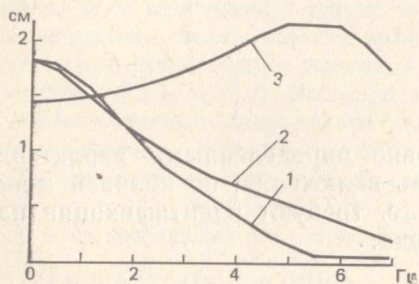


Рис. 3.10. Амплитудно-частотные характеристики системы нижней челюсти без нагрузки (движение вниз): 1 — измеренная; 2 — вычисленная; 3 — расчетная АЧХ нижней челюсти (усл. ед.)

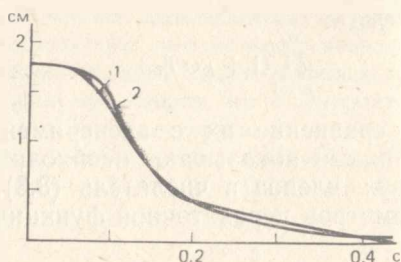


Рис. 3.11. Переходные процессы (движение вниз): 1 — измеренный; 2 — вычисленный

которая для данного испытуемого оказалась равной 315 г/см. При вычислении этого параметра предполагалось, что он не зависит от нагрузки, т. е. ткани принимались «пассивными». Однако в число этих тканей входят и другие мышцы, управляющие движением нижней челюсти, и даже левая жевательная мышца, поскольку электрической стимуляции подвергалась только правая. И хотя испытуемый старался максимально расслабить мышцы с тем, чтобы уменьшить произвольное сопротивление стимулированному движению, некоторое рефлекторное сокращение этих мышц могло привести к увеличению их жесткости. Следовательно, упругость абсолютно расслабленных мышц меньше найденной величины $c_{н.ч.}$.

Собственная частота нижней челюсти $f_{н.ч.}$, близкая к 6 Гц, может быть преувеличена наряду с преувеличением «истинного» значения $c_{н.ч.}$, если масса нижней челюсти с присоединенными тканями была занижена. Однако даже увеличение общей массы в 1,5 раза (до 300 г) соответствует снижению $f_{н.ч.}$ всего на 1 Гц, т. е. $f_{н.ч.}$ по-прежнему остается на границе измеренной полосы пропускания системы.

Последний параметр, связанный с нижней челюстью, — относительный коэффициент потерь $\xi_{н.ч.}$ может быть получен из анализа табл. 3.5. Как видно, значения $\xi_{н.ч.}$ различаются при движении нижней челюсти вверх (под действием стимулированного сокращения жевательной мышцы) и свободном движении вниз. Очевидно, что за величину, близкую к истинному $\xi_{н.ч.}$, следует взять постоянное значение при движении вниз. Увеличенное значение $\xi_{н.ч.}$ и зависимость его от нагрузки при движении нижней челюсти вверх свидетельствует об изменении вязких свойств мышц, не подвергающихся стимуляции.

Таблица 3.5

Условия эксперимента	Постоянные времени и коэффициент потерь при движении нижней челюсти					
	вверх			вниз		
	$\xi_{н.ч.}$	T_{11}	T_{12}	$\xi_{н.ч.}$	T_{11}	T_{12}
Без нагрузки	0,45	0,06	0,04	0,35	0,1	0,0
Нагрузка, г:						
50	0,6	0,07	0,04	0,35	0,1	0,015
100	0,8	0,08	0,07	0,35	0,08	0,025

В результате экспериментов было идентифицировано шесть параметров передаточной функции $W(j\omega)$ нейромышечной системы — четыре полюса и два нуля. Постоянные времени T_{11} и T_{12} , соответствующие нулям $W(j\omega)$, зависят от нагрузок и направления движения нижней челюсти. По-видимому, аналогичное изменение в зависимости от направления движения присуще и полюсам $W(j\omega)$, однако при данной экспериментальной методике это-

го обнаружить нельзя. Сравнение вычисленных и измеренных АЧХ и переходных процессов показывает возможность достижения хорошей аппроксимации передаточной функции нейромышечной системы для жевательной мышцы в классе дробно-рациональных передаточных функций с постоянными параметрами. Как известно, некоторые полюсы передаточной функции $W(j\omega)$ могут быть скомпенсированы близкими к ним нулями, однако обнаружить такие пары весьма затруднительно, а их результирующее влияние на поведение системы управления обычно невелико.

Итак, из описанных экспериментов следует, что, когда нижняя челюсть двигается при участии жевательной мышцы, полоса пропускания частот ее движений определяется полосой пропускания нейромышечной системы управления, а не механическими свойствами челюсти. Измерения спектра движений нижней челюсти при внешнем механическом ступенчатом или импульсном воздействии, выполненные в [3.18], свидетельствуют о том, что ширина этого спектра на уровне -20 дБ примерно равна 5 Гц. Та же ширина получается и при измерении спектра движений нижней челюсти при естественной артикуляции в ускоренном темпе. Наибольшая ширина спектра (7 Гц) соответствует максимально быстрому повторению слога *ТА*. Таким образом, разнообразные экспериментальные методики дают близкие оценки ширины амплитудно-частотной характеристики нижней челюсти — около 5 Гц. Это дает основание полагать, что и в процессах артикуляции передаточная функция системы «мышцы — нижняя челюсть» определяется в значительной степени механическими свойствами мышечного отдела этой системы, тогда как спектр передаточной функции собственно нижней челюсти значительно шире и не накладывает существенных ограничений на скорость артикуляторных движений.

3.2. НЕБНАЯ ЗАНАВЕСКА

МЫШЕЧНЫЙ АППАРАТ

Небная занавеска является следующим после нижней челюсти объектом по простоте конструкции и схеме действующих на нее сил. Хотя небная занавеска и представляет собой упругое тело, испытывающее деформации под воздействием мышечных усилий, тем не менее она в большинстве случаев может быть описана в терминах систем с сосредоточенными параметрами, поскольку непрерывность ее деформаций не играет существенной роли в процессах артикуляции.

Небная занавеска (рис. 3.12) образована мягкими тканями и проходящими через нее волокнами поднимающей (ПН), напрягающей (НН) и небно-язычной (НЯ) мышц. Небная занавеска служит для открывания или закрывания прохода из ротовой по-

лости в носовую. Открытый проход в носовую полость создает эффект назализации, являющийся отличительным признаком для согласных *М, Н* в русском языке. Признак назализации может создавать фонетическое отличие и для гласных звуков, как это, например, имеет место во французском языке. Нормальное положение небной занавески при свободном дыхании — опущенное, и лишь при артикуляции (в русском языке) она поднимается и закрывает проход в носовую полость для всех звуков, за исключением *М, Н*, так что в процессе речеобразования нормой является поднятое ее положение. Нарушение этого условия приводит к различным видам гнусавости.

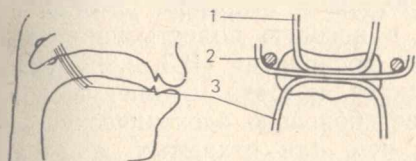


Рис. 3.12. Мышцы небной занавески: 1 — поднимающая; 2 — напрягающая; 3 — небно-язычная

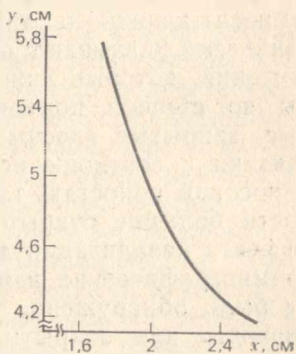


Рис. 3.13. Зависимость вертикальной координаты y небной занавески от ее горизонтальной координаты x

Один конец небной занавески свободен, а другой связан с тканями мягкого неба, поэтому при ее подъеме она описывает кривую, близкую к дуге окружности на участке до соприкосновения с задней стенкой прохода в носовую полость (рис. 3.13 по [3.32]). Как видно из этого рисунка, между высотой подъема небной занавески и степенью ее близости к задней поверхности существует нелинейное соотношение, увеличивающее погрешность при большой высоте небной занавески и затрудняющее применение правила [3.17], согласно которому площадь прохода в носовую полость может быть вычислена как минимальное расстояние в сагиттальной плоскости между задней поверхностью прохода и верхней поверхностью небной занавески, умноженное на коэффициент, равный 10. Другое ограничение на этот способ вычисления площади прохода накладывается тем фактом, что даже после соприкосновения средней линии небной занавески с задней стенкой проход в носовую полость остается еще не полностью перекрытым. В результате этого подъем небной занавески продолжается и после того, как, казалось бы, судя по кинорентгенограммам, проход уже перекрыт.

Высота небной занавески различна для разных звуков, и проход в носовую полость может быть приоткрыт даже для неназализованных звуков. Это связано с наличием порогового эффекта в восприятии назализации — если минимальная площадь прохо-

да в носовую полость менее $0,2 \text{ см}^2$, то звук не воспринимается как назализованный [3.30, 3.50]. Как было показано в [3.51], это определяется резким падением давления в ротовой полости при раскрытии прохода более чем на $0,2 \text{ см}^2$. Именно поэтому даже для глухих согласных, для которых высота небной занавески обычно максимальна, иногда наблюдается неполное перекрытие прохода с площадью отверстия $0,02\text{—}0,03 \text{ см}^2$, которая не влияет на фонетическое качество звуков. И все же эффект назализации не носит двоичного характера, поскольку измерения высоты небной занавески указывают на существование непрерывного спектра положений, которые лишь условно можно разбить на четыре группы по степени подъема: назальные согласные, открытые гласные, закрытые гласные и взрывные согласные [3.15, 3.37]. Это связано с соотношением акустического сопротивления в ротовой и носовой полостях, так что при большом открытии ротовой полости большее открытие прохода в носовую полость не создает эффекта назализации для слухового восприятия [В.3, 3.29]. Электромиографические измерения подкрепляют это мнение, поскольку было обнаружено, что ПН имеет большую электрическую активность для закрытых гласных, чем для открытых, а сдвиги языка в направлении «вперед — назад» не изменяют ЭМГ потенциалы этой мышцы [3.14, 3.22, 3.42].

Обсуждая вопрос о причинах различной высоты небной занавески для разных звуков, необходимо принять во внимание кинематику мягкого неба — его связь с языком и подъязычными структурами через мышцу НЯ. Как и всякая другая мышца, НЯ имеет ограниченную способность к растяжению, поэтому при опускании середины или корня языка возникает усилие, передающееся через эту мышцу как через пассивную структуру, которое может привести к опусканию небной занавески. Это опускание может быть частично предотвращено за счет соответствующего сокращения ПН, однако очевидно, что возможности подобной компенсации ограничены.

Непосредственные измерения воздушного потока через ноздри свидетельствуют о том, что иногда проход в носовую полость оказывается открытым при артикуляции звуков, удаленных от назального согласного на несколько фонетических единиц, причем на слух эти звуки все же не воспринимаются как назализованные [3.43, 3.51]. Таким образом, можно предположить, что слуховое восприятие эффекта назализации приспособилось к объективно существующему различию в высоте небной занавески, отличая каким-то образом ее произвольные и произвольные движения в соответствии с контекстом и фонетическим строем данного языка и противопоставляя по этому признаку лишь ограниченное число звуков.

Наряду со статическим описанием положения небной занавески для различных звуков большое значение имеет оценка ее динамических свойств, поскольку лишь знание динамики позволяет отделить те особенности поведения небной занавески в процессах артикуляции, которые вызваны управляющими командами из центральной нервной системы, от характеристик, связанных с ее механическими свойствами.

Исследования динамики обычно проводятся с использованием рентгеновской киносъемки. Установки типа микролучевого рентгеноскопа обладают преимуществом не только относительно малой дозой облучения, но и возможностью непосредственного ввода данных в ЭВМ и дальнейшей оперативной обработки [3.33]. Этот рентгеноскоп был использован для измерения речевых и неречевых движений небной занавески носителя русского языка [1.27]. В этих экспериментах отслеживались вертикальные и горизонтальные координаты небольших металлических горошин, приклеенных к небной занавеске, нижней челюсти и кончику языка, в то время как пара меток, укрепленных на переносице и верхних зубах, служила для учета движений головы. Скорость отсчетов составляла 232 кадр./с. Одновременно с измерением положения артикуляторных органов регистрировалась электрическая активность мышц, в том числе правой и левой ветвей ПН и НЯ, с помощью биполярных проволочных электродов. Сигналы ЭМГ записывались на магнитофон с частотной модуляцией, а затем вводились в ЭВМ, где они подвергались выпрямлению и интегрированию в окне 4,3 мс. Все эти сигналы синхронизировались с записью речевого сигнала.

В экспериментах были использованы следующие серии движений: 1: звук *И* — вдох через нос — звук *Ш*; 2: звук *И* — вдох — звук *И*; 3: звук *И* — вдох — звук *А*; 4: вдох — дутье (через рот); 5: последовательность слогов *ПА*; 6: последовательность слогов *ПИ*; 7: последовательность звуко сочетаний *ПИНПИ*; 8: последовательность звуко сочетаний *ПИМПИ*.

В обоих отведениях ЭМГ сигналов от НЯ не наблюдалось никакой активности, связанной с движениями небной занавески, хотя правильность введения электродов была подтверждена регистрацией активности при движениях языка. Относительно роли НЯ в процессах опускания небной занавески среди исследователей нет единого мнения. С одной стороны, наблюдается ее активность, возникающая на короткий промежуток времени (порядка 100 мс) при подавлении активности ПН перед началом опускания небной занавески [3.16, 3.37]. С другой стороны, данные об активности НЯ при артикуляции гласных и согласных противоречивы, а ряд исследователей вообще не обнаруживает какой-либо активности этой мышцы, в то время как другие считают, что ее использование зависит от индивидуальной тактики системы управления артикуляцией у данного диктора [3.14].

Мышца НН вследствие создаваемого ею направления усилий могла бы принимать участие в процессах опускания небной занавески, однако специальные исследования на этот счет не обнаружили ее активности в речевых движениях [3.30].

Таким образом, отсутствие активности НЯ, выявленное в описываемых экспериментах, служит подкреплением того мнения, что небная занавеска опускается, по крайней мере во многих случаях, без участия сокращения этой мышцы — за счет лишь упругих сил тканей, сжатых и изогнутых при ее подъеме. Поднимающая мышца является единственной мышцей, ответственной за подъем небной занавески, и в установившемся режиме существует линейная зависимость между уровнем ее ЭМГ активности и высотой подъема небной занавески. Это означает, что в статическом режиме система «ПН — небная занавеска» является линейной относительно ее вертикального смещения. В динамике же, конечно, уровень электрической активности этой мышцы зависит не только от целевого требуемого положения небной занавески, но и от ее исходного положения. Задержка между началом активности ПН и началом подъема небной занавески в наших экспериментах составляла в среднем 60 мс для 1—6-й серий движений (отдельные звуки и слоги, 42 реализации) и около 35 мс для 7-й и 8-й серий (звукосочетания, 10 реализаций).

Для изолированных звуков, а также при продувании воздуха через рот ЭМГ ПН демонстрирует отчетливый пик в самом начале (рис. 3.14). Этот пик становится еще выше при артикуляции звукосочетаний (рис. 3.15). На рисунке также хорошо видно, что при артикуляции назального согласного *М* небная занавеска не опускается до положения полного открытия прохода в носовую полость примерно наполовину. Приведенный пример служит иллюстрацией порогового эффекта в восприятии назализа-

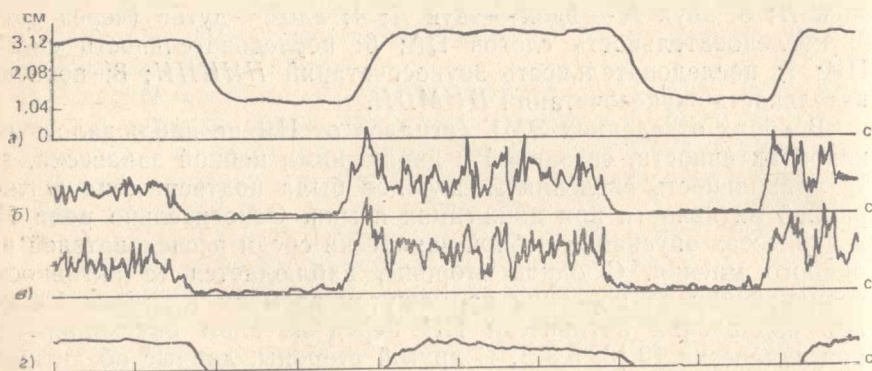


Рис. 3.14. Активность поднимающей мышцы и движения небной занавески. Последовательность И — вдох — Ш (интервалы между метками 0,43 с):

а — высота небной занавески; б — ЭМГ потенциалы левой ветви поднимающей мышцы; в — ЭМГ потенциалы правой ветви поднимающей мышцы; г — акустический сигнал

ции — увеличение площади прохода в носовую полость свыше некоторой величины не изменяет назального качества звука.

Для оценки динамических характеристик подъема и опускания небной занавески амплитудно-частотные характеристики ЭМГ ПН и вертикального смещения небной занавески вычислялись с помощью быстрого преобразования

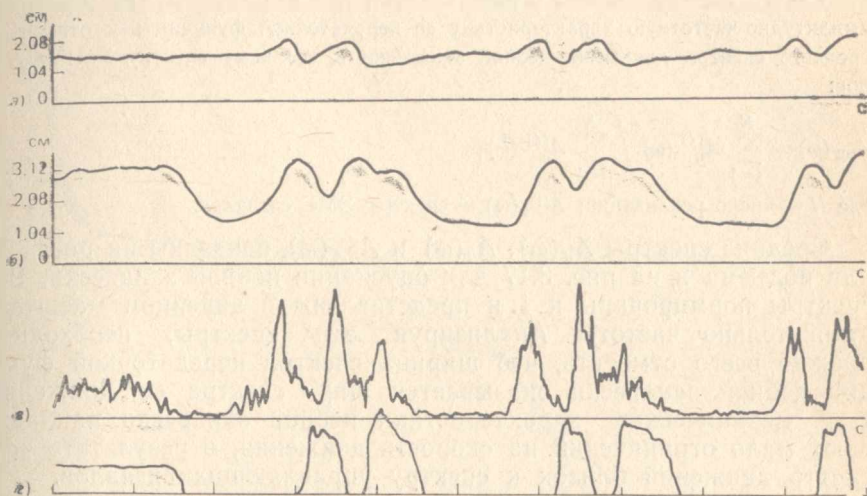


Рис. 3.15. Активность поднимающей мышцы и движение небной занавески и нижней губы. Звукосочетания ПИМПИ:

а — вертикальные движения нижней губы; б — вертикальные движения небной занавески; в — ЭМГ потенциалы поднимающей мышцы; г — огибающая речевого сигнала

Фурье. Если обозначить сигнал электрической активности мышцы через $e(t)$, а вертикальную координату небной занавески через $y(t)$, то амплитудно-частотная характеристика, например, движений небной занавески может быть найдена как

$$A_y(\omega) = \frac{1}{2\pi} \left| \int_0^T y(t) e^{-j\omega t} dt \right|,$$

где T — интервал анализа: $T = 512 \Delta t$; $\Delta t = 2,4$ мс. С целью уменьшения ошибок, связанных с конечным интервалом анализа, а также с учетом ступенчатобразного вида функций $e(t)$ и $y(t)$, для вычисления спектров были сформированы вспомогательные функции:

$$y_1(t) = \begin{cases} y(t), & t \leq t_0; \\ y_0, & t_0 < t < 2t_0; \\ y(2t_0 - t), & 2t_0 \leq t \leq T; \end{cases}$$

$$e_1(t) = \begin{cases} e(t), & -\delta t \leq t \leq t_0 - \delta t; \\ e_0, & t_0 - \delta t < t < 2t_0 - \delta t; \\ e(2t_0 - t), & 2t_0 - \delta t \leq T - \delta t. \end{cases}$$

Здесь величине с нулевым индексом y_0 соответствует установившееся значение сигнала $y(t)$ в момент t_0 , а e_0 — значение сигнала $e(t_0 - \delta t)$, где δt — задержка между началом электрической активности мышцы и началом движения небной занавески. При движении вниз принимается $\delta t = 0$, поскольку в этом случае понятие задержки не имеет физического смысла.

Считая систему «мышца — небная занавеска» линейной, можно определить амплитудно-частотную характеристику ее передаточной функции как отношение среднего спектра движений небной занавески к среднему спектру ЭМГ сигналов:

$$A_{н.з}(\omega) = \sum_{i=1}^N A_y^{(i)}(\omega) / \sum_{i=1}^N A_e^{(i)}(\omega),$$

где N — число реализаций; $A_e^{(i)}(\omega)$ — спектры ЭМГ сигналов.

Средние спектры $A_y(\omega)$, $A_e(\omega)$ и $A_{н.з}(\omega)$ показаны на рис. 3.16 для подъема и на рис. 3.17 для опускания небной занавески. Все спектры нормированы к 1 и представлены в линейном масштабе относительно частоты. Анализируя эти спектры, необходимо прежде всего отметить, что ширина спектра передаточной функции небной занавески оказывается шире спектра ее движений, т. е. механические характеристики небной занавески накладывают мало ограничений на скорость движения, в результате чего спектр движений близок к спектру управляющих сигналов.

Сравнивая спектры передаточной функции для различных групп движений — отдельных звуков, слогов и звукосочетаний при подъеме и опускании, обнаруживаем парадоксальное явление — ширина спектра при опускании небной занавески для отдельных звуков больше, чем ширина спектра при подъеме. Эта

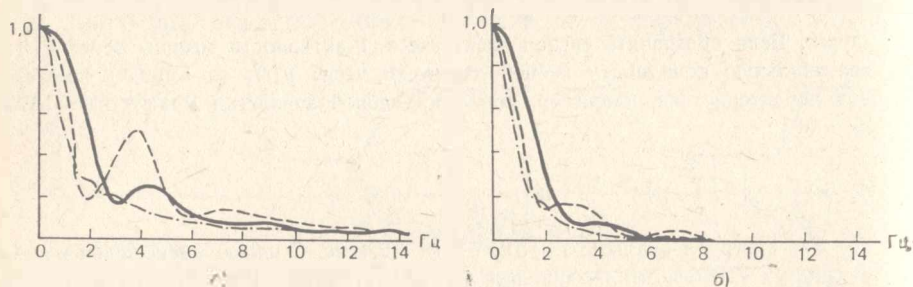
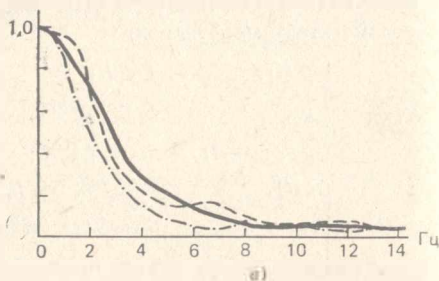


Рис. 3.16. Нормированные амплитудно-частотные характеристики ЭМГ поднимающей мышцы (а), движений (б) и передаточной функции небной занавески (в).

Движение вверх, — — — слоги; — — — — — бессмысленные звукосочетания; — · — — произвольные движения



разница становится еще более заметной при сравнении логарифмических спектров (рис. 3.18). В то же время спектры для слогов и звукосочетаний примерно одинаковы как для подъема, так и для опускания небной занавески. Парадоксальность этого явления заключается в его противоречии представлению о том, что актив-

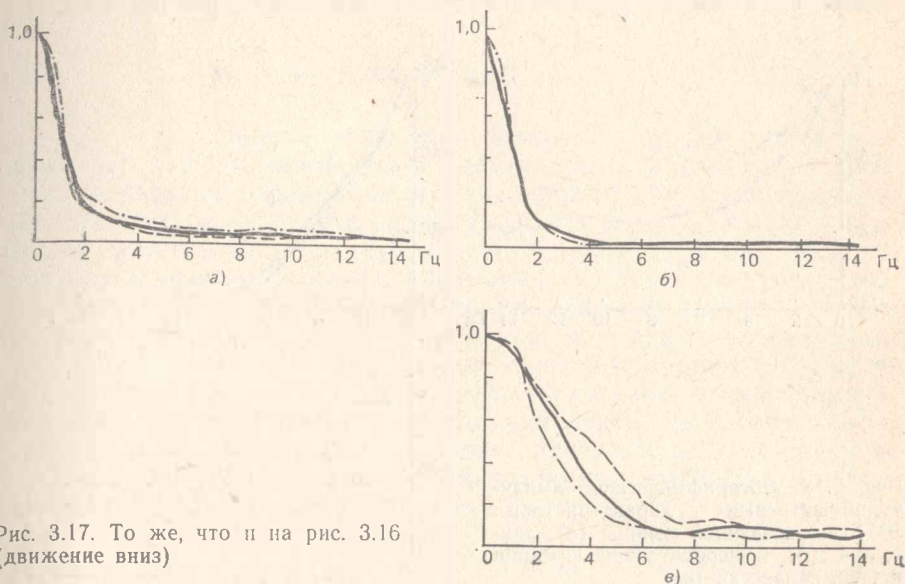


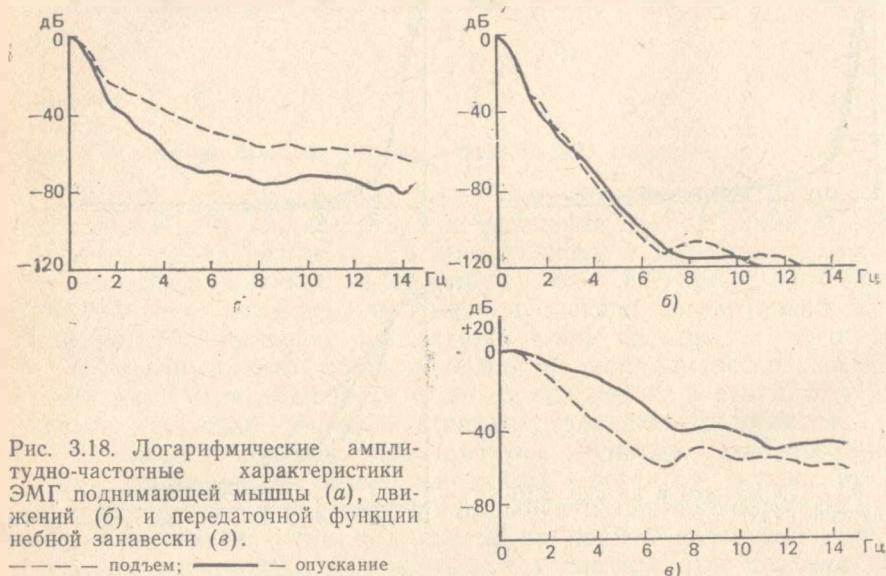
Рис. 3.17. То же, что и на рис. 3.16 (движение вниз)

ное сокращение мышцы, увеличивающее ее жесткость, должно расширять спектр передаточной функции и ускорять переходные процессы (см. гл. 1). Подъем небной занавески на конечных этапах к тому же подвержен ограничению при ее соприкосновении с задней стенкой прохода в носовую полость, а этот вид нелинейности должен был бы еще расширить спектр передаточной функции небной занавески при ее подъеме по сравнению со спектром при ее опускании.

Возможно несколько причин этого явления. Первой из них может быть неучтенная активность напрягающей и небно-язычной мышц при опускании небной занавески. И хотя оба электрода, отводящие потенциалы от небно-язычной мышцы, не регистрировали никакой активности в течение описываемых экспериментов, все же исключить такую возможность полностью нельзя. Правда, в этом случае возникает вопрос: почему предположительно незарегистрированная активность этой мышцы столь сильно различается для артикуляции отдельных звуков и звукосочетаний? Другая причина может заключаться в воздушном давлении на небную занавеску при вдохе через нос, который осуществляется только в серии с отдельными звуками.

Наконец, еще одно различие между этими сериями эксперимен-

тов состоит в том, что длительность интервала, в течение которого небная занавеска остается поднятой, была значительно больше в серии с артикуляцией отдельных звуков (1—2 с), чем в слогах и звукосочетаниях (около 0,2—0,4 с). Этот фактор представляется наиболее важным, если принять во внимание возможность существования больших релаксационных постоянных времени в мыш-



це (обсуждение этого вопроса см. в § 1.4). Если поднимающая мышца обладает спектром релаксационных постоянных времени, в число которых входят сравнительно большие величины, т. е. порядка секунд, то к концу интервала 1—2 с жесткость мышцы возрастет по сравнению с ее начальным состоянием перед подъемом небной занавески. Поскольку жесткость не может быстро уменьшиться, небная занавеска начинает опускаться при большем значении жесткости мышечных тканей, что и определяет расширение спектра ее передаточной функции.

Ширина спектра передаточной функции небной занавески на уровне фона (примерно —60 дБ от уровня на нулевой частоте) для звукосочетаний ЛИМПИ и ПИНПИ составляет примерно 10—12 Гц при подъеме небной занавески и примерно 7—9 Гц при ее опускании. При артикуляции этих звукосочетаний время, в течение которого небная занавеска задерживается в поднятом положении, сравнительно невелико, поэтому несколько большая ширина ее спектра при подъеме по сравнению со спектром опускания вполне соответствует представлению о роли повышения жесткости активно сокращающейся мышцы в переходных процессах управляемого ею объекта.

Оценка временных параметров движений небной занавески приводится в [3.31], где путем идентификации параметров обыкновенного дифференциального уравнения второго порядка был установлен диапазон значений постоянной времени (0,075—0,081 с) и коэффициент затухания (1,0—1,07).

УПРУГИЕ ДЕФОРМАЦИИ

Небная занавеска является упругим телом, а ее движения с формальной точки зрения представляют собой трехмерные упругие деформации, которые можно описать с помощью математического аппарата, представленного в гл. 1. Однако в процессах речеобразования существенным является лишь площадь прохода в носовую полость, зависящая от степени подъема небной занавески, а форма пространственного ее изгиба играет второстепенную роль. Это дает основание для использования однопараметрического уравнения движения небной занавески, где аргументом является угол поворота φ относительно неподвижных структур неба или вертикальной координаты центра ее тяжести, аналогично тому, как это было в задаче динамики нижней челюсти.

Используя эквивалентные (приведенные) значения момента инерции J (или массы m), упругости c и вязкости r , имеем уравнение второго порядка

$$J\ddot{\varphi} + r\dot{\varphi} + c\varphi = F_m(t), \quad (3.5)$$

где $F_m(t)$ — разность между моментами силы, создаваемой сокращением поднимающей мышцы, и аэродинамической силы, возникающей при протекании воздушного потока из ротовой полости в носовую. В этом уравнении коэффициент упругости c является функцией аргумента φ , поскольку после соприкосновения небной занавески с задней поверхностью прохода в носовую полость жесткость системы возрастает за счет дополнительного изгиба тканей небной занавески. Аналогичное явление происходит и с вязкостью системы, поэтому (3.5) можно представить на всем интервале значений угла φ двумя уравнениями, каждое из которых линейно на интервалах $\varphi_1 \leq \varphi_0$, $\varphi_2 > \varphi_0$, где φ_0 — угол, при котором небная занавеска соприкасается с задней поверхностью прохода в носовую полость. Каждое из этих уравнений имеет разное, но не зависящее от φ значение коэффициента упругости c и коэффициента вязкости r :

$$J\ddot{\varphi}_1 + r_1\dot{\varphi}_1 + c_1\varphi_1 = F_1(t), \quad \varphi_1 \leq \varphi_0; \quad (3.6)$$

$$J\ddot{\varphi}_2 + r_2\dot{\varphi}_2 + c_2\varphi_2 = F_2(t), \quad \varphi_2 > \varphi_0. \quad (3.7)$$

Для (3.7) существует граница применения, связанная с ограничением дальнейшего движения небной занавески после полного перекрытия прохода в носовую полость при некотором значении $\varphi_{кр}$, после чего аргумент φ_2 остается неизменным $\varphi_2 = \varphi_{кр}$, а даль-

нейшее сокращение поднимающей мышцы приведет лишь к сжатию тканей небной занавески и большей силе ее прижатия к стенке прохода в носовую полость.

Как уже говорилось, эффект назализации определяется в конечном счете площадью прохода в носовую полость, а рассчитать эту величину без больших погрешностей лишь по значению вертикальной координаты небной занавески из-за нелинейной зависимости между ними (особенно на заключительных стадиях подъема) затруднительно. Поэтому наиболее точное представление об изменении площади этого прохода и его соотношения с движениями других артикуляторных органов можно получить только с помощью измерений на кинорентгенограммах.

На рис. 3.19 для слога АМА показаны траектории расстояний между губами, между языком и задней поверхностью речевого тракта (измерения производились в горизонтальной плоскости, прохо-

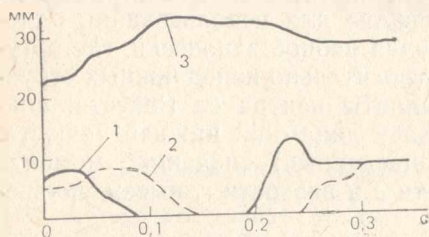


Рис. 3.19. Артикуляция слога АМА:
1 — расстояние между губами; 2 — ширина прохода в носовую полость; 3 — расстояние между языком и задней поверхностью речевого тракта (на уровне нижних зубов)

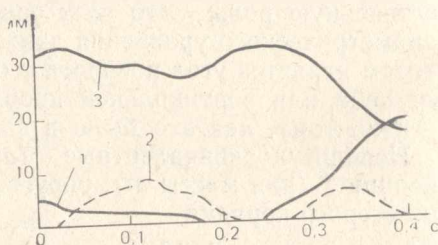


Рис. 3.20. Артикуляция слога АНА:
1 — расстояние от кончика языка до твердого неба; 2 — ширина прохода в носовую полость; 3 — расстояние от языка до задней поверхности речевого тракта (на уровне нижних зубов)

дящей через верхнюю поверхность нижних зубов), а также минимальной ширины прохода в носовую полость. Эти величины были измерены на кинорентгенограмме, снятой со скоростью 75 кадр./с для того же диктора, что и в предыдущих экспериментах. Слог АМА находился между слогами АДА и АНА, причем темп артикуляции был весьма высоким. Обращает на себя внимание то, что небная занавеска была опущена уже на первом гласном А, а затем проход в носовую полость закрылся почти в середине губной смычки и открылся вновь после ее взрыва, по-видимому, приготовляясь к артикуляции следующего слога. Аналогичная картина наблюдается и для слога АНА (рис. 3.20). Соответствующая конфигурация речевого тракта показана на рис. 3.21.

Таким образом, проход в носовую полость открывается еще до того, как губы или язык начинают движение к смычке, и закрывается или уменьшается до весьма малого значения сразу же после смычки. В описанных выше экспериментах на микролучевом рентгенооскопе также было обнаружено, что, например, в звукосочетании ПИМПИ небная занавеска поднимается почти одно-

временно с губной смычкой на звуке *М*. Эти наблюдения согласуются с данными [3.32] о том, что подъем небной занавески начинается в районе смычки иногда за несколько десятков миллисекунд до нее (за 60—80 мс). Измерения воздушного потока через ноздри [3.9] дают среднее опережение опускания небной занавески относительно начала смычки для назальных согласных в окружении гласных примерно 40—50 мс, в то время как для назальных, следующих за неназальными согласными, небная занавеска опускается с запаздыванием относительно начала смычки примерно на 27 мс, что определяется необходимостью обозначить начало смычки как сегмент, принадлежащий неназальному согласному. Если учесть, что движение языка или губ к смычке занимает примерно 40—80 мс, то, скорректировав данные относительно акустического сигнала, получим значение опережения, находящегося примерно в одном диапазоне для разных методов измерений.

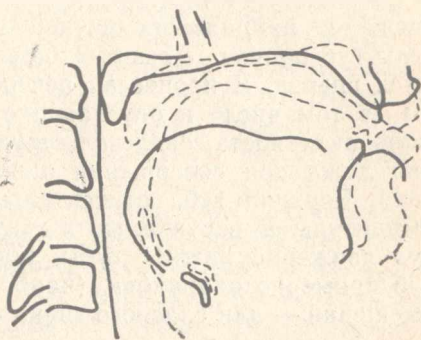


Рис. 3.21. Формы речевого тракта в слове АМА:

— первый гласный; — — — звук *М*

3.3. ГУБЫ

МЫШЕЧНЫЙ АППАРАТ

Движения губ управляются мышцами, часть из которых расположена симметрично относительно срединносагиттальной плоскости (рис. 3.22). В их число входят две ветви верхней поднимающей мышцы, поднимающие верхнюю губу, две ветви нижней опускающей мышцы, опускающие нижнюю губу, щечная, растягивающая углы губ в стороны, и мышца, опускающая углы губ. Сжатие и выпячивание губ осуществляется за счет сокращения кольцевой мышцы, имеющей верхнюю и нижнюю ветви. Подъем нижней губы также может происходить при сокращении подбородочной мышцы, закрепленной верхним концом на костной структуре нижней челюсти, а нижним — на

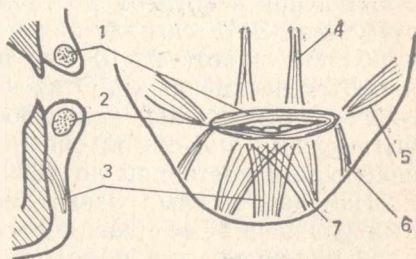


Рис. 3.22. Мышцы губ:

1 — верхняя ветвь кольцевой мышцы; 2 — нижняя ветвь кольцевой мышцы; 3 — подбородочная мышца; 4 — верхняя поднимающая мышца; 5 — щечная мышца; 6 — мышца, опускающая углы губ; 7 — нижняя опускающая мышца

мягких тканях подбородка, которые при сокращении этой мышцы поднимаются, толкая нижнюю губу вверх.

Правые и левые ветви мышц могут управляться независимо, придавая при своем сокращении различные формы губы. Степень выпячивания губ при сокращении кольцевой мышцей зависит от напряжения щечной — при ее расслабленном состоянии углы губ получают возможность сближаться, а сами губы выгибаются вперед. В процессах речеобразования все упомянутые мышцы (в том числе и опускающие углы губ) проявляют электрическую активность, свидетельствующую о необходимости координации даже при совершении одномерных движений губ — вверх — вниз. Верхняя губа относительно менее подвижна, чем нижняя: амплитуды ее вертикальных движений значительно меньше амплитуд движений нижней губы. Горизонтальные же смещения обеих губ примерно одинаковы (наибольшая амплитуда горизонтальных смещений — для гласного звука У).

ДИНАМИЧЕСКИЕ ХАРАКТЕРИСТИКИ

В [1.29] исследовались механические характеристики губ и связанных с ними тканей с помощью внешнего источника силы, скачкообразно растягивающего один угол губ при фиксированном положении другого. При этом обнаружилось, что переходные процессы характеризуются двумя сильно различающимися постоянными времени — 0,05 и 15 с. Это наблюдение свидетельствует о существовании широкого спектра релаксационных параметров живых тканей. Постоянная времени нарастания мышечного усилия (включая и нарастание электрического потенциала) равна 0,1 с. Используя вместо ступенчатого воздействия силу, изменяющуюся по синусоидальному закону с различным частотами, удалось измерить амплитудно-частотный и фазовый спектры. Различия между амплитудно-частотными характеристиками при нулевой активности кольцевой мышцы (полностью расслабленные губы) и при некотором постоянном уровне напряжения, контролируемом по ЭМГ сигналу кольцевой мышцы, могут достигать 10 дБ в области частот 2—10 Гц, тогда как фазовые характеристики остаются неизменными. Относительные изменения в жесткости губ за счет сокращения кольцевой мышцы при артикуляции могут быть весьма большими: они находятся в диапазоне от 200 в расслабленном состоянии до 2000 дн./см при сжатии губ [3.11].

Измерения силы, развиваемой кольцевой мышцей при различных удлинениях и степени электрической активности, показывают наличие линейности в довольно широком диапазоне и ясно демонстрируют зависимость эффективности этой мышцы от ее удлинения.

Амплитудно-частотный спектр вертикальных движений нижней губы совместно с движениями нижней челюсти вычислялся в [1.29] с помощью спектрографа, на который подавались сигналы

в процессе артикуляции на протяжении примерно 30—45 мин. Из этих исследований следует, что, начиная от частот 4—5 Гц, огибающая спектра совместных движений начинает падать с наклоном примерно в —6 дБ/окт., а вблизи частот 8—9 Гц наклон увеличивается до —18 дБ/окт. Увеличение скорости артикуляции поднимает уровень огибающей спектра в диапазоне частот 1—8 Гц (для разных дикторов по-разному). Спектр становится равномерным на уровне шумов примерно в области 20—30 Гц, но необходимо учесть, что поскольку он относится к совместному движению нижней губы и челюсти, то спектр собственных движений нижней губы должен быть уже.

При увеличении темпа артикуляции уровень ЭМГ активности кольцевой мышцы и скорость движения губ обычно возрастают [3.24]. Однако нет согласия в различных сообщениях о зависимости между смещениями и скоростью движения губы. Так, в [3.9] приводятся результаты измерений движений губы в звуко-сочетаниях типа гласный — согласный, которые свидетельствуют об увеличении скорости при увеличении амплитуды смещения. Вместе с тем в [3.49] указывается, что в противоположность нижней челюсти (см. предыдущий параграф) для нижней губы такой зависимости не было обнаружено. Эксперименты, проведенные совместно с Т. Геом и В. Эваном (Коннектикутский университет, США), показали, что это явление имеет более сложную природу, хотя в общем можно говорить о наличии положительной корреляции между скоростью и амплитудой смещения нижней губы. В этих экспериментах движение нижней губы отслеживалось с помощью системы фотоэлектрических датчиков, причем движения нижней челюсти вычитались из движений губ. Одновременно с помощью проводочных биполярных электродов регистрировались сигналы ЭМГ от подбородочной, кольцевой и опускающей мышц. В качестве речевого материала служили трехсложные бессмысленные звуко-сочетания с ударением на втором слоге, с одним и тем гласным *А* во всех слогах и различными согласными *Б, П, М, В, Ф, Д, Т, С, З, Ш* (например, *БАБАБА, ТАТАТА* и т. д.). Звуко-сочетания произносились в нормальном и ускоренном темпах диктором-носителем русского языка. На рис. 3.23 видно, что характер связи между скоростью движения нижней губы и амплитудой ее смещения неоднозначен и зависит от темпа, направления движения губы и фонетического качества артикулируемого звука.

На скорость движения и величину смещения артикуляторных органов влияют индивидуальные характеристики дикторов. Например, в [3.43] было обнаружено, что у одного из дикторов скорость движения небной занавески возрастала с увеличением темпа артикуляции, а у другого — оставалась неизменной. Поэтому сведения о соотношении между скоростью и смещением артикуляторных органов, очевидно, не носят абсолютного характера, а в известной мере определяются индивидуальной тактикой системы управления артикуляцией.

В соответствии с большей подвижностью нижней губы и актив-

ность нижней ветви кольцевой мышцы больше активности верхней ветви (рис. 3.24, изолированный звук *Б*, поверхностные электроды). Как видно из этого рисунка, в артикуляции взрывного согласного принимают участие также и опускающая, и подбородочная мышцы, причем отчетливо видны синхронные импульсы этих мышц при размыкании губ. В то же время для огубленных гласных, например *У*, эти мышцы не проявляют заметной активности.

Синхронность опускающей и подбородочной мышцы, по-видимому, объясняется необходимостью стабилизировать нижнюю губу за счет усилий, приложенных к ее углам с целью компенсации ее сдвига как твердого тела под влиянием усилия, развиваемого подбородочной мышцей. Такая же схема активности сохраняется и при артикуляции звуко сочетаний.

Разброс в значениях параметров нейромоторных команд — опережения по фазе движения губы и длительности — зависит от темпа артикуляции, причем изменения темпа больше сказываются на длительности команд, чем на опережении. Измеренные в этих экспериментах параметры довольно близки к указываемым в [3.49] опережениям для подбородочной мышцы 40 мс, и для опускающей мышцы 40—60 мс.

Длительность переходных процессов нижней губы зависит от

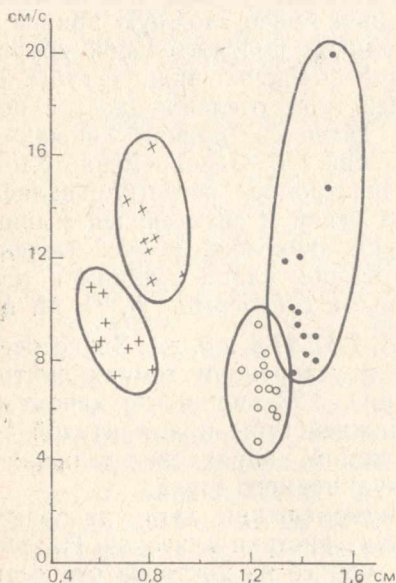


Рис. 3.23. Зависимость скорости движения губ от амплитуды ее смещений.

(+) — подъем; (O) — опускание; (X) — быстрый подъем; (•) — быстрое опускание

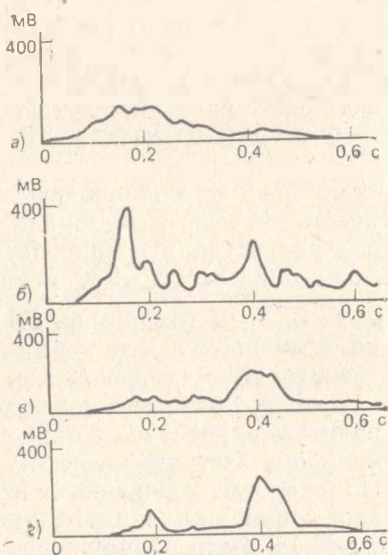


Рис. 3.24. Активность губных мышц (звук *Б*):

а — верхняя ветвь кольцевой мышцы; *б* — нижняя ветвь кольцевой мышцы; *в* — мышца, опускающая углы губ; *г* — подбородочная мышца

скорости артикуляции, положения слога в звукосочетании и направления ее движения — вверх или вниз (табл. 3.6). Из таблицы видно, что опускание губ происходит гораздо быстрее, чем ее подъем. Это объясняется как различиями в тактике управления движениями губ, так и возможным влиянием избыточного давления в ротовой полости, накапливаемого во время смычки. При взрыве смычки возникает воздушный поток с большой скоростью, которой может вызвать образование турбулентных шумов на губах и замаскировать фонетическое качество последующего звука. Если же при взрыве смычки губы быстро разойдутся на расстояние, на котором уже невозможно образование турбулентности, то длительность замаскированного сегмента звука окажется минимальной. Это соображение подкрепляется обнаружением ЭМГ активности депрессоров нижней губы во время взрыва не только губной смычки (где эта активность наибольшая), но и переднеязычной смычки.

Таблица 3.6

Вид движения	Зависимость длительности переходных процессов нижней губы, мс, от положения в звукосочетании		
	1	2	3
Подъем:			
норма	300	200	180
быстро	210	100	120
Опускание:			
норма	70	160	90
быстро	50	60	60

В проведенных экспериментах также выяснилось, что точность

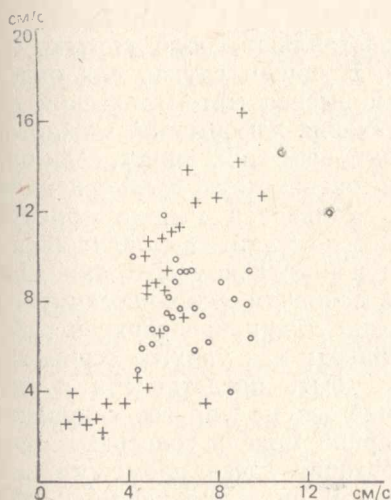


Рис. 3.25. Скорость движений нижней челюсти (ось абсцисс) и губ (ось ординат).

(+) — подъем; (•) — опускание

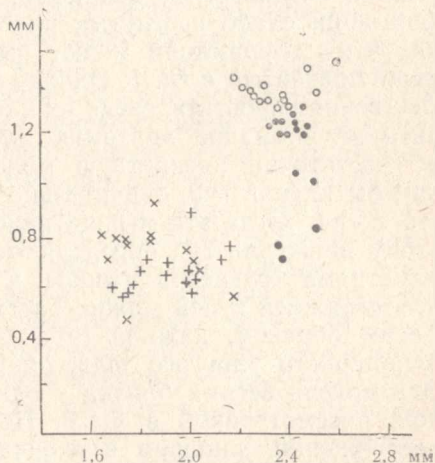


Рис. 3.26. Амплитуда смещений нижней челюсти (ось абсцисс) и губ (ось ординат)

(+) — подъем; (•) — опускание; (x) — быстрый подъем; (O) — быстрое опускание

стабилизации положения нижней губы значительно выше точности стабилизации скорости — относительная дисперсия для положения составляет примерно 5% при опускании и около 10% при подъеме, тогда как относительная дисперсия скорости — порядка 30% независимо от направления движения губы. Здесь надо принимать во внимание погрешность, связанную с определением собственного движения нижней губы путем вычитания движений нижней челюсти из совокупного движения губы и челюсти, так что фактически ошибки стабилизации положения губы при ее подъеме должны быть несколько меньше.

В этих экспериментах не был обнаружен эффект компенсации параметров движений нижней губы в зависимости от параметров движений нижней челюсти ни по скорости, ни по положению (рис. 3.25 и 3.26). Подобная компенсация, как известно, наблюдается в некоторых случаях речевой патологии или в специальных экспериментах, где фиксируется положение нижней челюсти или в него вносятся неожиданные изменения. Общая картина смещений нижней челюсти и собственных смещений нижней губы, наоборот, свидетельствует о положительной корреляции между ними, а в диапазоне смещений 1,6—2,6 см эта зависимость весьма близка к пропорциональной. Зависимость, приближающаяся к пропорциональной, наблюдается и в отношении скоростей движения этих артикуляторных органов.

УПРУГИЕ ДЕФОРМАЦИИ

Движения губ под воздействием распределенных усилий, создаваемых мышцами, представляют собой упругие деформации слабо изогнутых пластин. В общем случае для описания этих деформаций может быть применен математический аппарат, развитый в гл. 1. Однако уравнения деформаций можно существенно упростить, сведя их к одномерным колебаниям, если принять во внимание анатомию губ и разработать приближенную, но достаточно адекватную модель. Кольцевая мышца образует как бы каркас губ, значительно более жесткий, чем окружающие его ткани, когда эта мышца находится в активном состоянии. Поэтому движения губ определяются в основном механическими параметрами кольцевой мышцы, а мягкие ткани над верхней губой и под нижней губой можно рассматривать как упругое основание. Таким образом, каждая губа может быть представлена в виде одномерного упругого тела, опирающегося на упругое основание. Эта модель весьма близка к одномерной модели голосовых складок, рассмотренной в § 2.6. Не повторяя здесь промежуточных рассуждений, запишем в окончательном виде уравнение колебаний какой-либо губы, поскольку для другой губы отличия будут заключаться лишь в геометрии и механических параметрах.

Пусть ось x направлена вдоль губ во фронтальной плоскости, ось y — горизонтально в сагиттальной плоскости, по направлению

от губ к затылку, а ось z направлена вверх. Тогда уравнение колебаний губы

$$-\frac{\partial^2}{\partial x^2} \left[J_y (E + N) \frac{\partial^2 u}{\partial x^2} \right] - NS \frac{\partial^2 u}{\partial x^2} + cu + r \frac{\partial u}{\partial t} + \rho \frac{\partial^2 u}{\partial t^2} = F(x, t), \quad (3.9)$$

где $u(x, t)$ — отклонение средней линии губы от нейтрального положения; J_y — момент инерции сечения относительно оси y ; E — модуль упругости губы при расслабленном состоянии кольцевой мышцы; N — удельное натяжение, создаваемое кольцевой мышцей; S — площадь поперечного сечения; c — упругость подстилающего слоя; r — коэффициент вязкого трения; ρ — погонная плотность тканей; $F(x, t)$ — сила, создаваемая остальными мышцами. Силу инерции вращения поперечного сечения, а также поперечные силы сдвига, как и в § 2.6, можно учесть просто поправками на собственные числа уравнения (3.9). Характер движения губ зависит от степени закрепления их углов, т. е. от степени активности щечной и опускающей углы губ мышц. При напряженном состоянии этих мышц углы губ оказываются более или менее жестко закрепленными, и при сокращении остальных губных мышц губы совершают лишь движения вверх — вниз. При расслабленном состоянии угловых мышц углы губ приобретают возможность сближаться, так что сокращение кольцевой мышцы приводит не только к вертикальному движению, но и к выпячиванию губ. В этом случае определенную роль играет и то, что губы изогнуты, т. е. в расслабленном состоянии их средняя линия слегка выдается вперед. Этот изгиб невелик, и его можно не учитывать при составлении уравнения первого приближения.

Свойство подвижности углов губ влияет на выбор модели граничных условий для уравнения (3.9). Ясно, что импеданс закрепления углов губ описывается довольно сложной зависимостью от степени активности кольцевой мышцы и мышц, связанных с углами губ. Но характер этой зависимости не вполне ясен, и, пожалуй, нет необходимости в точном его описании, если принять во внимание приближенный характер формальной модели колебаний губ. Поэтому можно ограничиться выбором простейшей модели граничных условий, отражающей основной тип закрепления губ. В [3.1] было исследовано решение (3.9) для граничных условий типа шарнира, а в [3.4] — для условий типа жесткой заделки. Последняя модель дает более сложную форму для собственных функций и несколько более высокие значения собственных чисел. Формы первых гармоник для того и другого решений оказываются довольно близкими. Учитывая податливость концов губ, представляется целесообразным принять в качестве модели граничных условий простейший вариант — шарнирное закрепление, соответствующее требованию отсутствия смещений и нулевому значению момента поворота:

$$u(0, t) = u(l, t) = 0; \quad \left. \frac{\partial^2 u(x, t)}{\partial x^2} \right|_{x=0} = \left. \frac{\partial^2 u(x, t)}{\partial x^2} \right|_{x=l} = 0, \quad (3.10)$$

где l — длина губ.

Зависимость параметров уравнения (3.9) от координаты x довольно слабая, поэтому в первом приближении все коэффициенты можно принять постоянными. Движения губ вверх и вниз отличаются степенью активности, а следовательно, и жесткостью кольцевой мышцы. Например, при опускании нижней губы за счет сокращения опускающей мышцы можно принять $N = \text{const}$, тогда как ее подъем осуществляется при активном изменении натяжения кольцевой мышцы, так что N становится функцией времени. В первом случае (3.9) может рассматриваться как уравнение с постоянными параметрами, а во втором — зависимость $N(t)$ будем аппроксимировать линейным законом $N(t) = n_0 + n_1 t$.

Для решения (3.9) вновь используем метод разделения переменных, представляя $u(x, t) = U(x)T(t)$. Получим систему:

$$U^{IV} - \frac{NS}{(E+N)J_y} U'' - \lambda^2 U = 0; \quad (3.11)$$

$$T'' + 2gT' + \omega^2 T = 0, \quad (3.12)$$

где λ — собственные числа; $\omega^2 = [c + \lambda^2(E+N)]J_y/\rho$; $2g = r/\rho$.

Общее решение (3.11) есть

$$U(x) = a_1 \operatorname{ch} p_1 x + a_2 \operatorname{sh} p_1 x + a_3 \cos p_2 x + a_4 \sin p_2 x,$$

где p_1 и p_2 — корни характеристического уравнения. Принимая во внимание граничные условия (3.10), получаем:

$$a_1 = a_3 = 0; \quad a_2 \operatorname{sh} p_1 l + a_4 \sin p_2 l = 0; \quad p_1^2 a_2 \operatorname{sh} p_1 l - p_2^2 a_4 \sin p_2 l = 0.$$

Исключая тривиальный случай $a_2 = a_4 = 0$, имеем $\sin p_2 l = 0$, откуда $a_2 = 0$, и собственные функции для (3.11) есть

$$U_k(x) = a_4 \sin p_{2k} x,$$

где $p_{2k} = k\pi/l$, $k = 1, 2, \dots$, а a_4 определяется из условия нормирования собственных функций и равно в данном случае $a_4 = \sqrt{2/l}$. Квадраты собственных чисел найдем как

$$\lambda_k^2 = p_{2k}^2 \left[p_{2k}^2 + \frac{2NS}{(E+N)J_y} \right].$$

В случае постоянного натяжения кольцевой мышцы, например при опускании нижней губы, решение для (3.12) представляется обычной комбинацией затухающих колебаний, а частное решение (3.9), т. е. решение неоднородного уравнения без возбуждающей силы F , есть

$$u(x, t) = \sum_{k=1}^{\infty} (C_{1k} e^{-\tau_{1k} t} + C_{2k} e^{-\tau_{2k} t}) \sqrt{\frac{2}{l}} \sin p_{2k} x, \quad (3.13)$$

где $\tau_{1k} = -g + \sqrt{g^2 - \omega_k^2}$; $\tau_{2k} = -g - \sqrt{g^2 - \omega_k^2}$, а C_{1k} и C_{2k} находим из начальных условий. При линейном изменении натяжения кольцевой мышцы во время сближения губ аналогично решению, полученному для голосовых складок в § 2.6, имеем

$$T_k(t) = e^{-gt} \sqrt{n_0 + n_1 t} Z_{\frac{1}{3}}(\tilde{t}),$$

где $Z_{1/3}(\tilde{t}) = C_{3k}J_{1/3}(\tilde{t}) + C_{4k}N_{1/3}(\tilde{t})$ — функция Бесселя; а $J_{1/3}(\tilde{t})$ — цилиндрическая функция первого рода; $N_{1/3}(\tilde{t})$ — цилиндрическая функция второго рода; $\tilde{t}$ — новый аргумент, зависящий от $N(t)$. Переходные процессы губ занимают значительно больший интервал, чем период колебаний голосовых складок, поэтому, начиная с некоторого момента, выполняется условие $\tilde{t} \gg 1$, при котором становится справедливым асимптотическое представление функций Бесселя:

$$J_v(\tilde{t}) = \sqrt{\frac{2}{\pi \tilde{t}}} \cos\left(\tilde{t} - \frac{v\pi}{2} - \frac{\pi}{4}\right);$$

$$N_v(\tilde{t}) = \sqrt{\frac{2}{\pi \tilde{t}}} \sin\left(\tilde{t} - \frac{v\pi}{2} - \frac{\pi}{4}\right).$$

Частное решение неоднородного уравнения (3.9) составляется из свободных колебаний системы $u_{св}(x, t)$, зависящих от начальных условий, и вынужденных колебаний $u_{вын}(x, t)$:

$$u(x, t) = u_{св}(x, t) + u_{вын}(x, t) = \sum_{k=1}^{\infty} [T_{k\text{ св}}(t) + T_{k\text{ вын}}(t)] \sqrt{\frac{2}{l}} \sin p_{2k}x,$$

где, например, для опускания нижней губы $u_{св}(x, t)$ определяется по (3.13), а вынужденное движение для каждой гармоники получается как решение следующего уравнения:

$$T_k'' + 2gT_k' + \omega_k^2 T_k = 4F(t)/\rho k\pi.$$

Здесь предполагается, что усилия мышц распределены равномерно по длине губы. Амплитуды возбуждаемых гармоник быстро убывают с ростом их порядкового номера. Например, амплитуда второй гармоники примерно в 10 раз меньше амплитуды первой, а амплитуда третьей гармоники уже в 100 раз меньше амплитуды первой гармоники. Поэтому при малых смещениях губ достаточно хорошая аппроксимация их движений получается при использовании лишь одной — первой — гармоники. Это иллюстрируется рис. 3.27, где показаны результаты аппроксимации с помощью аналоговой вычислительной машины движений нижней губы, измеренных на кинорентгенограммах. При следующих параметрах: $E = 5 \cdot 10^6$ Па, $\rho = 1,05$ г/см³, $l = 4$ см, $J_y = 0,05$ см⁴, $g = -38$ см⁻¹ — была достигнута наилучшая аппроксимация. Во время губной смычки, особенно для

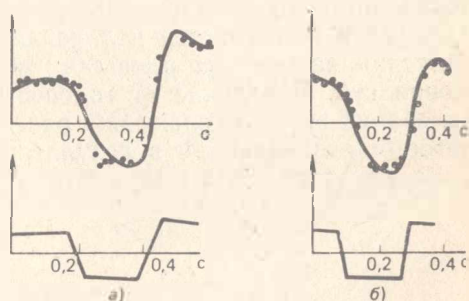


Рис. 3.27. Движения нижней губы (в условных единицах): а — слог АПЕ; б — слог АПА:
— измерения на кинорентгенограммах; ... — моделирование; внизу — управляющие команды

глухого согласного Π , в ротовой полости накапливается избыточное воздушное давление, способствующее раскрытию губ, как это обсуждалось выше.

Рассмотрим механизм влияния внутриротового давления на движение губ. При взрыве смычки падение избыточного давления P описывается по [В.3] следующим уравнением:

$$\frac{d}{dt} V \left[1 + \frac{P(t)}{P_0} \right] = -v(t) S_r(t), \quad (3.14)$$

где v — скорость воздушного потока; V — объем ротовой полости; P_0 — статическое давление в атмосфере; S_r — площадь губного отверстия; $S_r(t) = \int_0^l [u_1(x, t) - u_2(x, t)] dx$; $u_1(x, t)$ — форма изгиба верхней губы; $u_2(x, t)$ — форма изгиба нижней губы. Учитывая предыдущие результаты, касающиеся быстрого падения амплитуды высших гармоник в составе функций $u_1(x, t)$ и $u_2(x, t)$, можно приближенно записать, считая форму изгиба обеих губ идентичной с точностью до постоянного коэффициента b ,

$$S_r(t) \approx \sqrt{\frac{2}{l}} \int_0^l T_1(t) (b+1) \sin \frac{\pi x}{l} dx = \frac{\sqrt{2l}}{\pi} (b+1) T_1(t).$$

Скорость воздушного потока через губное отверстие, в свою очередь, может быть найдена путем решения аэродинамического уравнения, составленного в § 2.4:

$$\rho_0 h \frac{d(S_r v)}{dt} + \kappa_1 S_r v + \frac{\kappa_2 \rho_0}{2} S_r v^2 = P S_r, \quad (3.15)$$

где h — длина губного отверстия вдоль оси y ; κ_1 — коэффициент вязкого трения; κ_2 — коэффициент динамического сопротивления. В данном случае скорость движения губ s' мала по сравнению с площадью губного отверстия, что дает основания для применения аналитического решения (3.15), соответствующего квазистационарному режиму, исследованному в § 2.4. Решив систему (3.14), (3.15) относительно избыточного давления в ротовой полости, найдем дополнительную силу, действующую на губы в продолжение начального интервала взрыва губной смычки. Как и в случае автоколебаний голосовых складок, уравнение движения губ после взрыва смычки оказывается нелинейным, поскольку сила избыточного давления зависит от площади губного отверстия, т. е. от положения губ. В отличие от голосовых складок, не возникает автоколебаний губ, поскольку они расходятся на такое расстояние, при котором возникающая в результате эффекта Бернулли сила уже мала для того, чтобы заставить губы вновь сближаться.

3.4. ЯЗЫК

СТРОЕНИЕ И МЫШЕЧНЫЙ АППАРАТ

Изменения формы и положения языка обеспечивают артикуляцию большинства звуков речи. Это определяет важность исследования движений языка и необходимость построения адекватных математических моделей. В отличие от рассмотренных ранее артикуляторных органов: нижней челюсти, небной занавески и даже губ, — принципиальной и существенной особенностью языка является то, что его форма есть результат упругих деформаций, возникающих под влиянием усилий внешних и внутренних мышц, и именно его форма влияет на фонетическое качество артикулируемых звуков. Таким образом, язык есть система с распределенными параметрами.

В мышечном аппарате языка различают две группы мышц — внешних и внутренних. Внешние мышцы одним концом прикреплены к костным структурам, а другим — к языку. В их число входят подбородочно-язычная, шило-язычная, подъязычно-язычная (рис. 3.28). В эту группу иногда относят небно-язычную мышцу, уже известную нам по исследованиям небной занавески, и язычно-фарингиальную (которая, строго говоря, не является самостоятельной мышцей, а представляет собой верхний отдел мышцы, сжимающей фаринкс). Внутренние мышцы расположены только внутри языка и не имеют непосредственной связи с костями. В их число входят верхняя и нижняя продольная мышцы, поперечная и вертикальная мышцы. На положение языка и направление усилий его внешних мышц также влияет положение подъязычной косточки, которая сдвигается впе-

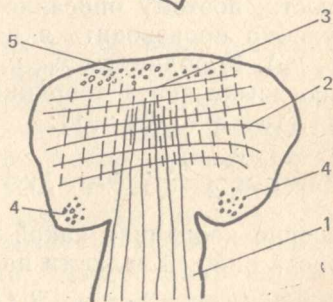
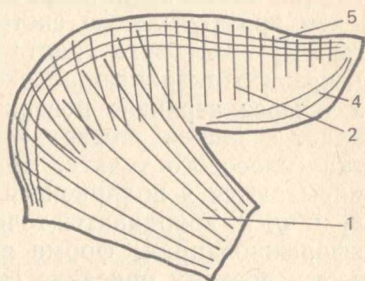
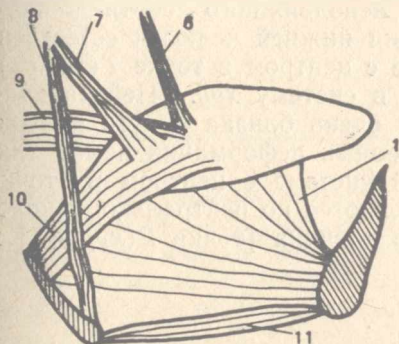


Рис. 3.28. Мышцы языка:

1 — подбородочно-язычная; 2 — вертикальная; 3 — поперечная; 4 — нижняя продольная; 5 — верхняя продольная; 6 — небно-язычная; 7 — шило-язычная; 8 — шило-подъязычная; 9 — язычно-фарингиальная; 10 — подъязычно-язычная

ред при сокращении челюстно-подъязычной мышцы и назад или вниз под влиянием сокращения мышц гортани.

Направление усилий, создаваемых внешними мышцами языка, обеспечивает смещение языка как твердого тела в любую сторону в ротовой полости, и в совокупности с сокращениями внутренних мышц эти усилия могут придавать весьма сложные формы поверхности языка. Связь активности той или иной мышцы с каким-либо конкретным движением языка или артикуляцией звуков речи исследовалась в [2.97, 3.39, 3.41, 3.46, 3.52]. При этом выяснилось, в частности, что различные мышечные волокна подбородочно-язычной мышцы активизируются в разной степени при артикуляции разных звуков [3.41].

ДВИЖЕНИЯ ЯЗЫКА КАК ТВЕРДОГО ТЕЛА

Язык может менять свою форму как в сагиттальной плоскости головы, вертикально проходящей через ось симметрии голосового тракта, так и в поперечном направлении. Роль этих поперечных деформаций еще не выявлена в достаточной степени, и большинство исследователей отдает предпочтение описанию деформаций языка в сагиттальной плоскости. В этой плоскости конфигурация голосового тракта определяется положением языка относительно неподвижных участков тракта и собственными деформациями. Смещения языка как единого целого задаются вращениями нижней челюсти относительно точки ее закрепления и горизонтальными и вертикальными смещениями корня языка. Модели этих движений были разработаны в [3.5].

Введем прямоугольную систему координат $хоу$, в которой опишем форму голосового тракта (рис. 3.29). Плоскость $хоу$ проходит через вертикальную плоскость симметрии голосового тракта, ось x — через верхнюю поверхность голосовых складок, а ось y совпадает с задней поверхностью неподвижного вертикального участка голосового тракта. Вращения нижней челюсти естественно описываются в полярной системе с центром в точке $T(x_T, y_T)$, где x_T и y_T — координаты точки T в систему $хоу$. Нейтральная (недеформированная) форма языка очень близка к дуге полуокружности, поэтому описание собственных деформаций языка также удобно производить в полярной системе с центром в точке $P(x_P, y_P)$. Координаты x_P, y_P вычисляются по нейтральному положению языка через координаты его кончика [точка $F(x_F, y_F)$] и корня [точка $V(x_V, y_V)$]:

$$x_P = x_V + \frac{x_F - x_V}{2}; \quad y_P = y_V + \frac{y_F - y_V}{2}.$$

Смещение координат какой-либо точки языка $A(x_A, y_A)$ за счет поворота нижней челюсти на угол $\varphi_{н.ч}$ есть:

$$x'_A = x_T + (x_A - x_T) \cos \varphi_{н.ч} + (y_A - y_T) \sin \varphi_{н.ч}; \quad (3.16)$$

$$y'_A = y_T - (x_A - x_T) \sin \varphi_{нч} + (y_A - y_T) \cos \varphi_{нч}, \quad (3.16)$$

где x'_A , y'_A — новые координаты точки A . Угол поворота нижней челюсти $\varphi_{нч} = 2 \arcsin \frac{UD}{TD}$, где U — точка на верхнем зубе, с которой совмещается точка D нижнего зуба при полном сжатии зубов. Координаты точки корня языка $V(x_V, y_V)$ также изменяются за счет поворота нижней челюсти. Однако эта точка имеет и самостоятельное движение при сокращении соответствующих мышц, смещающих корень языка вверх — вниз или вперед — назад. Такое смещение корня языка оказывает влияние на положение всех остальных точек языка, причем характер их смещения отличается от смещения, вызванного поворотом нижней челюсти. Для правильного расчета влияния движений корня языка необходимо отделить его самостоятельные движения от смещений за счет поворота нижней челюсти.

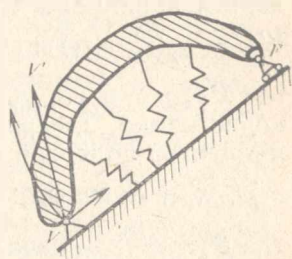
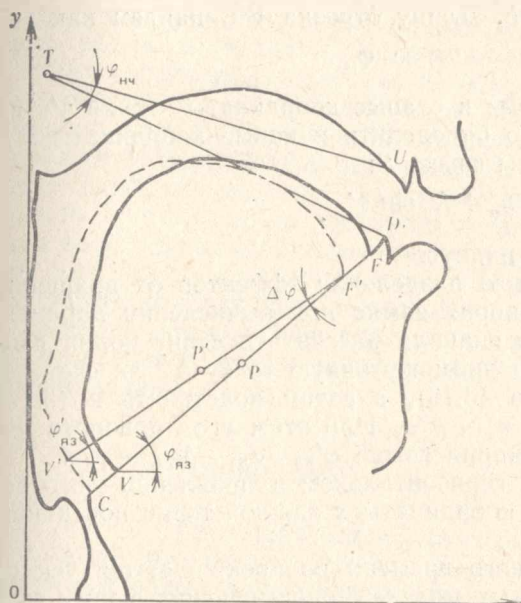


Рис. 3.30. Модель сдвига языка при смещении его корня

Рис. 3.29. Координатные системы моделей движений языка

Механика движений корня языка не столь проста, как поворот нижней челюсти, и для ее описания необходима некоторая модель. Эта модель изображена на рис. 3.30. Кончик языка на некотором протяжении свободен и может скользить по подстилающей поверхности, что условно показано в виде скользящего шарнира. Остальная часть языка находится на упругом основании, изображенном в виде пружин. Принимая такую модель, любое смещение корня языка из точки V в точку V' можно рассматривать как суперпозицию двух движений: сдвига языка вдоль линии VF и поворота относительно точки F . Эти смещения вызыва-

ют появление упругих сил реакции подстилающей поверхности, которые могут в какой-то степени изменить форму языка. Однако при анализе рентгенограмм отделить это изменение формы от изменений, вызванных сокращением внешних и внутренних мышц языка, довольно трудно. Если не учитывать деформаций, а рассматривать лишь смещение языка как твердого тела, то для описания влияния движений корня языка достаточно вычислить новое положение центра полярных координат $P(x_P, y_P)$.

Обозначим угол наклона линии VF в исходной декартовой системе xoy для нейтрального положения языка через $\varphi_{яз}$, а поворот этой линии при движении корня через $\Delta\varphi_{яз}$. Тогда новое значение угла наклона линии VF есть $\varphi'_{яз} = \varphi_{яз} - \Delta\varphi_{яз}$, где $\Delta\varphi_{яз} =$

$$= \arctg \frac{y_F - y_V}{x_F - x_V}; \quad \Delta\varphi_{яз} = \arcsin \frac{V'C}{V'F'}, \quad \text{а точка } C \text{ — проекция нового}$$

положения корня языка V' на линию VF . Пренебрегая растяжением или сжатием языка под действием сил, приложенных к его корню, принимаем $V'F' = VF$. Длину отрезка $V'C$ найдем как

$$V'C = -(x'_V - x_V) \sin \varphi_{яз} + (y'_V - y_V) \cos \varphi_{яз}, \quad (3.17)$$

где x'_V, y'_V и x_V, y_V — новые и старые координаты точки V на корне языка. Теперь можно определить и новые координаты x'_P, y'_P центра полярной системы языка P :

$$x'_P = x'_V + R_0 \cos \varphi'_{яз}; \quad y'_P = y'_V + R_0 \sin \varphi'_{яз},$$

где $R_0 = VF/2$ — начальный радиус.

В связи с необходимостью разделения эффектов от вращения нижней челюсти и сдвига корня языка при вычислении параметров этого смещения нужно сначала рассчитать сдвиг корня языка из точки V в некоторую промежуточную точку V'' за счет поворота нижней челюсти по (3.16), а затем подставить в (3.17) вместо x_V, y_V координаты x''_V, y''_V . При этом угол поворота линии VF за счет смещения корня языка $\varphi'_{яз} = \varphi_{яз} - \Delta\varphi_{яз} - \varphi_{н.ч.}$

Сдвиг нижней челюсти в горизонтальном направлении учитывается просто приращениями координаты x каждой точки поверхности языка.

На протяжении длительного времени не прекращаются усилия по разработке количественных моделей, описывающих форму языка в процессах речеобразования. Так, в [3.44] была предпринята попытка описания формы языка как суперпозиции его конфигураций для некоторых гласных звуков. В [3.36] было обращено внимание на самостоятельную роль нижней челюсти как артикуляторного органа в формировании конфигурации речевого тракта. Модель языка в виде упругого эллипсоида, деформируемого сокращением мышц, была предложена в [3.27]. Эта модель основывалась на правильной трактовке физики деформаций языка, но ввиду математических трудностей, как известно, теория упругости предлагает решение в замкнутой форме лишь для ограниченного числа трехмерных конфигураций, в том числе и для эллипсо-

ида. Поэтому, обладая преимуществом аналитического решения, эта модель все же оказывается слишком далекой от реального объекта. С целью преодоления аналитических трудностей в [3.45] был использован метод численного анализа упругих деформаций языка, представленного в сагиттальной плоскости *хоу* в виде 14 четырехугольных областей. Задавая усилия в направлении реально существующих мышц, на этой модели удалось получить весьма близкие к естественным формы и временные переходы языка, однако время расчетов на универсальной ЭВМ было чрезвычайно большим. Для ускорения вычислений в [3.34] применили метод конечных элементов, причем язык был представлен в виде набора из 86 тетраэдров, т. е. моделировалась уже не плоская, а трехмерная структура. Дальнейшее развитие этой модели выявило интересные особенности деформаций языка, связанные с практической нежизнеспособностью его объема [3.23]

В [3.1, 3.2, 3.4, 3.5] был развит другой метод анализа деформаций языка, ориентированный на его упругие колебания. Такой подход в большей степени учитывает динамический характер процессов речеобразования. Представление упругих колебаний языка через собственные функции некоторого уравнения в частных производных дает одновременно как форму языка, так и соответствующий этой форме переходный процесс и избавляет от необходимости решения дифференциального уравнения для каждой артикуляционной ситуации. Таким образом, открываются возможности не только для интенсивного исследования этой модели, но и для ее практического использования при синтезе речи. Не меньшую роль эта модель может играть и в решении обратной задачи — восстановления конфигурации голосового тракта по речевому сигналу, так как учет кинематики и динамических ограничений позволяет в значительной степени доопределить эту некорректную задачу.

КОНЧИК ЯЗЫКА

Гистологические исследования строения языка [3.40, 3.45] показывают, что его кончик, занимающий примерно четверть от общей длины языка, свободен от волокон подбородочно-язычной мышцы, которые входят в тело языка симметрично относительно среднесагиттальной плоскости, перекрывая примерно 30% ширины языка (см. рис. 3.28). Поэтому изгиб кончика языка вверх или вниз относительно тела языка осуществляется лишь за счет сокращения верхней или нижней продольной мышцы соответственно. Микроручевой рентгеноскоп дает возможность одновременной записи электрической активности верхней и нижней мышц и различных движений кончика языка. В наших экспериментах потенциалы верхней продольной мышцы измерялись в двух точках, расположенных справа и слева от среднесагиттальной плоскости языка.

На рис. 3.31 и 3.32 показаны движения кончика языка в экспериментах с заданием либо коснуться, либо не касаться твердого неба. В обоих случаях активность верхней мышцы начиналась примерно за 50 мс до начала подъема кончика языка. Соответствующие спектры ЭМГ сигналов, движений кончика языка и амплитудно-частотные характеристики его передаточной функции

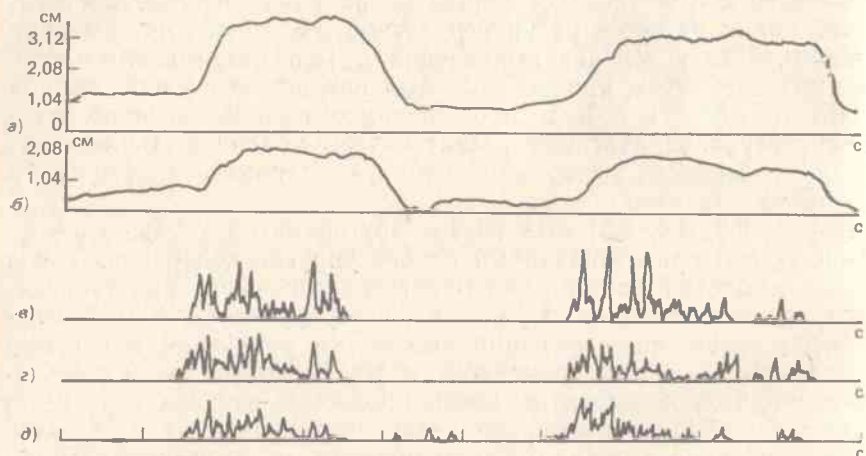


Рис. 3.31. Движения кончика языка при задании не касаться твердого неба: а — вертикальное смещение; б — горизонтальное смещение; в, г, д — ЭМГ верхней продольной мышцы.
Интервал между метками 0,43 с

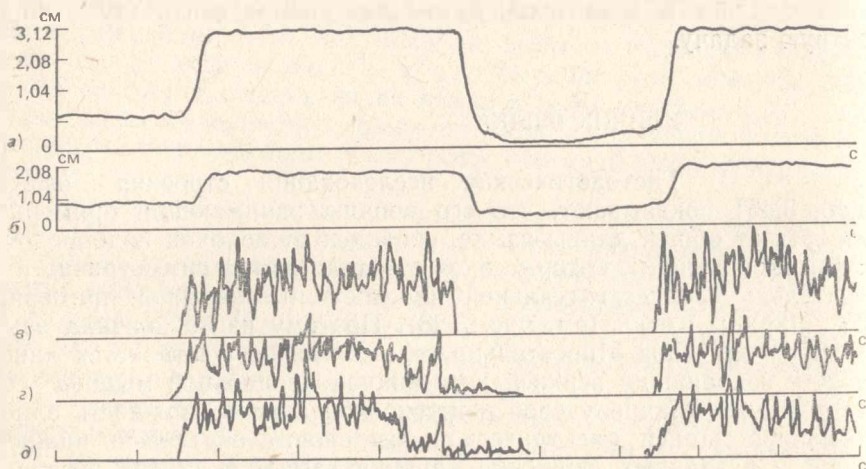


Рис. 3.32. Движения кончика языка с касанием твердого неба (обозначения те же, что и на рис. 3.31)

показаны на рис. 3.33. Как и следовало ожидать, АЧХ передаточной функции кончика языка при касании твердого неба шире, чем в том случае, когда кончик языка не доходит до твердого неба, поскольку ограничения дальнейшему движению равносильны введению нелинейного элемента в передаточную функцию. Поэтому при практически одинаковых спектрах входных сигналов (ЭМГ

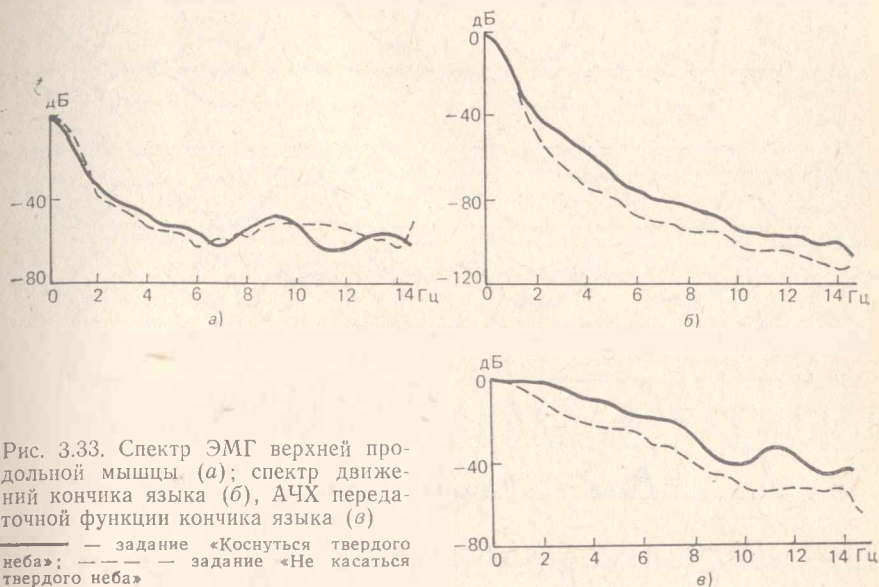


Рис. 3.33. Спектр ЭМГ верхней продольной мышцы (а); спектр движения кончика языка (б), АЧХ передаточной функции кончика языка (в)

верхней продольной мышцы) спектры движений кончика языка и АЧХ его передаточной функции при касании твердого неба оказываются шире.

Участие кончика языка в артикуляции переднеязычных согласных звуков несомненно. Однако и в этом случае артикуляция представляет собой результат координированных движений различных артикуляторных органов, причем одна и та же цель, например переднеязычная смычка, может достигаться различными средствами. Так, на рис. 3.34 и 3.35 показаны движения кончика языка и нижней челюсти при артикуляции последовательности слогов *ТА* в нормальном и ускоренном темпах. Видно, что при нормальном темпе артикуляции амплитуда вертикальных движений кончика языка имеет заметную величину, отражающую степень его самостоятельного движения. При быстром темпе артикуляции вертикальная координата кончика языка примерно постоянна, тогда как горизонтальное смещение происходит синхронно с акустическим сигналом. Здесь обращает на себя внимание горизонтальное движение нижней челюсти, амплитуда которого сравнима с амплитудой ее вертикальных колебаний, так что в бы-

стром темпе артикуляции переднеязычная смычка в значительной степени осуществляется за счет движений нижней челюсти. Причиной этого явления может состоять в относительной слабости продольных мышц языка, не обеспечивающих необходимой скоро-

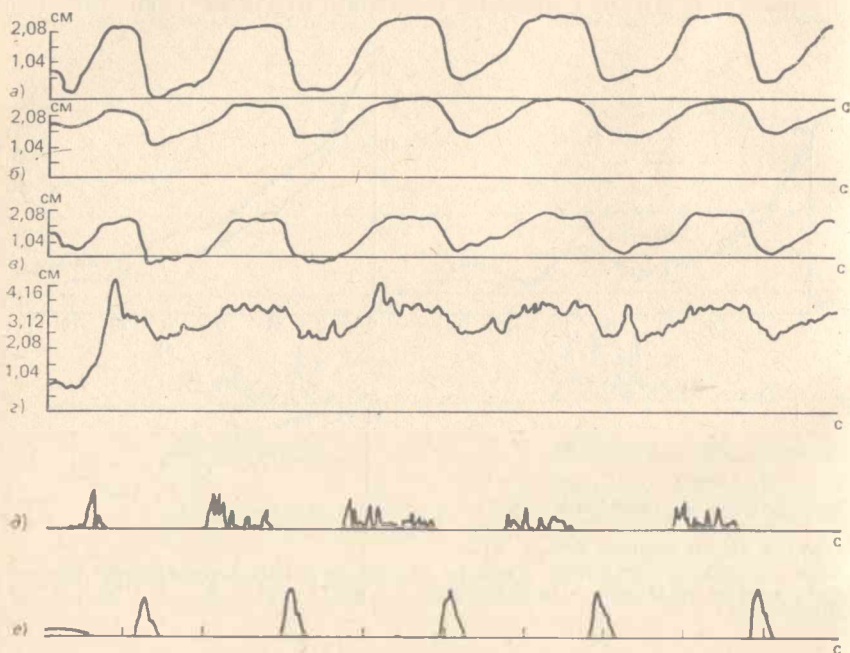


Рис. 3.34. Последовательность слогов *ТА*, нормальный темп:

а — вертикальное смещение нижней челюсти; *б* — горизонтальное смещение нижней челюсти; *в* — вертикальное смещение кончика языка; *г* — горизонтальное смещение кончика языка; *д* — ЭМГ жевательной мышцы; *е* — огибающая акустического сигнала.
Интервал между метками — 0,43 с

сти его подъема при быстром темпе артикуляции, что и вызывает вспомогательную активность нижней челюсти. Так же, как и в случае исследований ЭМГ сигналов жевательной мышцы для быстрого темпа осмысленных звуко сочетаний, увеличение скорости артикуляции может приводить к изменению нейромоторной деятельности мышечного аппарата нижней челюсти, компенсируя дефицит мощности мышц языка.

Из-за особенностей строения языка и возможности самостоятельного движения его кончика представляется целесообразным рассмотреть механику движений кончика языка как отдельного артикуляторного органа. Введем систему координат такую, что ось *x* направлена вдоль плоскости симметрии языка от позвоночника к зубам, ось *y* — вверх, а ось *z* — в поперечном направлении. Проведя условную границу, отделяющую кончик языка от его

тела, получим возможность описать его как упругую пластинку. При сокращении верхней продольной мышцы создаются момент, изгибающий кончик языка вверх, и усилие, тянущее его назад. Сдвиг кончика назад не представляет в данный момент интереса, поскольку он практически не изменяет формы языка и может быть

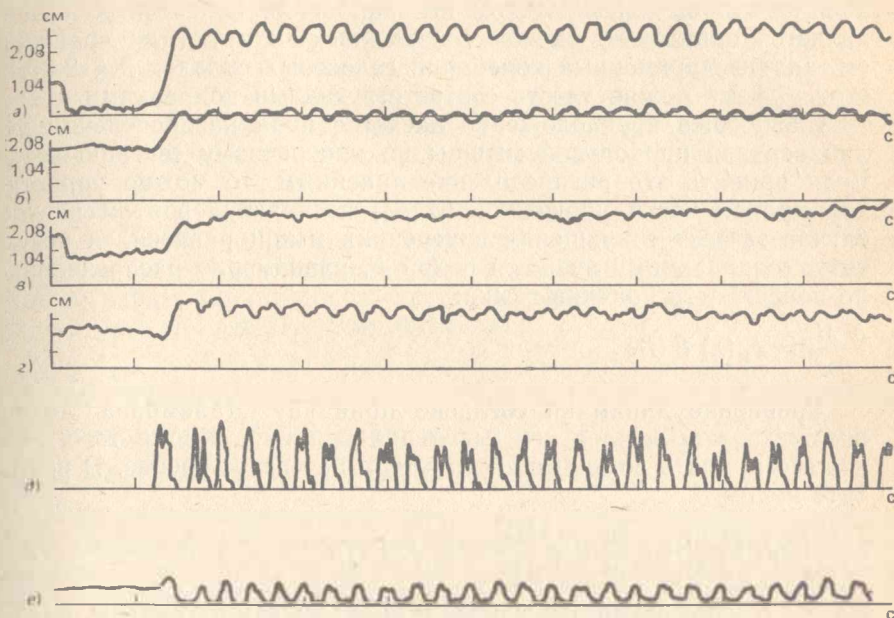


Рис. 3.35. Движения кончика языка и нижней челюсти. Последовательность слогов ТА, быстрый темп (обозначения те же, что и на рис. 3.34)

при необходимости скомпенсирован усилиями передней части подбородочно-язычной мышцы. Рассмотрим поэтому лишь изгиб кончика, составив уравнение статики, требующее равновесия изгибающих моментов:

$$\frac{\partial^2}{\partial x^2} \left(E J_z \frac{\partial^2 u}{\partial x^2} \right) = - \frac{\partial^2 M}{\partial x^2}, \quad (3.18)$$

где u — отклонение средней линии кончика языка от равновесного положения; E — модуль упругости тканей языка; J_z — момент инерции поперечного сечения относительно оси z ; M — момент, создаваемый верхней продольной мышцей.

Сделаем несколько предположений, принципиально не изменяющих модели механики упругих деформаций, но существенно упрощающих процесс анализа. Прежде всего откажемся от описания поперечных изгибов кончика языка, т. е. будем рассматривать его не как пластинку, а как стержень. Описание поперечного изгиба, вообще говоря, не представляет затруднений с формальной

точки зрения, но делает выкладки более громоздкими. Выбор одномерной модели изгиба кончика языка оправдывается экспериментально установленной симметрией его движений (см., например, [3.28]), и тем, что поперечные изгибы кончика не существенны при артикуляции переднеязычных согласных звуков. Единственным негативным результатом игнорирования поперечных изгибов является занижение оценок собственных частот (для сравнения см. анализ двухмерных колебаний голосовых складок в § 2.7), но этот эффект можно учесть соответствующими поправками.

Следующее предположение касается вида распределения усилий верхней продольной мышцы по поперечному сечению языка. Если принять это распределение линейным, то можно упростить вид правой части уравнения (3.18), поскольку в этом частном случае эффект сокращения внутренних мышц равносителен результату, создаваемому внешней силой F , равномерно распределенной по поверхности кончика языка:

$$\frac{\partial^2 M}{\partial x^2} = J_z(x) F(t).$$

Уравнение динамики согласно принципу Даламбера должно включать в себя член, учитывающий инерцию. Кроме того, необходимо принять во внимание и потери на вязкое трение. В результате имеем

$$\frac{\partial^2}{\partial x^2} \left[EJ_z \frac{\partial^2 u}{\partial x^2} \right] + r \frac{\partial u}{\partial t} + \rho \frac{\partial^2 u}{\partial t^2} = -J_z F(t). \quad (3.19)$$

Здесь ρ — погонная плотность тканей языка; $\rho = \rho_T S(x)$; $S(x)$ — площадь поперечного сечения; r — погонный коэффициент вязкости, также зависящий от поперечного сечения. Разделим переменные в (3.19), представив $u(x, t) = U(x)T(t)$, и исследуем однородное уравнение, распадающееся на два уравнения, каждое из которых зависит лишь от одной координаты:

$$(J_z U'')'' - \rho \lambda^2 U = 0; \quad (3.20)$$

$$T'' + 2gT' + \lambda^2 ET = 0. \quad (3.21)$$

В (3.20) момент инерции $J_z = J_z(x)$ и погонная плотность $\rho = \rho(x)$, так что все коэффициенты зависят только от пространственного аргумента и не зависят от времени. В (3.21) $2g = r/\rho$, и если принять, что r зависит от площади, так же как и ρ , т. е. $r = r_T S(x)$, где r_T — показатель вязкого трения, то g оказывается не зависящим от пространственного аргумента x . Модуль упругости $E = E(t)$, поскольку кончик языка поднимается за счет сокращения внутренней мышцы, что приводит к изменению жесткости тканей. Таким образом, коэффициенты (3.21) не зависят от x .

Граничные условия на конце языка соответствуют свободному положению, а в месте соединения с телом языка в первом приближении примем жесткое закрепление. Это даст следующую систему:

$$u(0, t) = \frac{\partial u(x, t)}{\partial x} \Big|_{x=0} = 0; \quad \frac{\partial^2 u(x, t)}{\partial x^2} \Big|_{x=l} = \frac{\partial^3 u(x, t)}{\partial x^3} \Big|_{x=l} = 0, \quad (3.22)$$

где l — длина кончика языка.

Степень жесткости связи кончика языка с телом языка зависит от напряжения внутренних мышц, и требование равенства нулю производной $u'(0)$ в области их соединения, конечно, идеализирует реальную ситуацию, но оно оправдано существенным упрощением дальнейшего анализа, тогда как связанная с этим погрешность в определении собственных частот может быть скорректирована при сравнении с реально наблюдаемыми движениями кончика языка.

Чтобы получить аналитическое решение (3.19) с граничными условиями (3.22), примем в качестве первого приближения, что площадь поперечного сечения S и момент инерции J_z постоянны. Тогда корни характеристического уравнения для (3.20) есть $p, -p, jp, -jp$, где j — мнимая единица, и

$$p = \sqrt[4]{\lambda^2 \frac{\rho}{J_z}}.$$

а общее решение (3.19) имеет вид

$$U(x) = a_1 \operatorname{ch} px + a_2 \operatorname{sh} px + a_3 \cos px + a_4 \sin px.$$

Учитывая граничные условия (3.22), получаем трансцендентное уравнение для поиска корней $\operatorname{ch} pl \cos pl = -1$, откуда имеем $p_k = q_k/l$, где $q_1 = 0,597\pi$; $q_2 = 1,494\pi$; $q_k = (k-0,5)\pi$, $k > 2$. Собственные функции затем выражаются как

$$U_k(x) = a_k (\operatorname{ch} p_k x - \cos p_k x) + b_k (\operatorname{sh} p_k x - \sin p_k x).$$

Коэффициенты a_k и b_k зависимы:

$$-b_k = a_k \frac{\operatorname{ch} q_k + \cos q_k}{\operatorname{sh} q_k + \sin q_k} = a_k \frac{\operatorname{sh} q_k - \sin q_k}{\operatorname{ch} q_k + \cos q_k}.$$

Из условия нормировки $\int_0^l U_k^2(x) dx = 1$ получим: $a_k = 0,707 \sqrt{2/l}$;

$b_1 = -0,518 \sqrt{2/l}$; $b_2 = -0,721 \sqrt{2/l}$; $b_k = -a_k$, $k > 2$. Таким образом, общее решение (3.19) есть

$$u(x, t) = \sum_{k=1}^{\infty} [T_{k \text{ св}}(t) + T_{k \text{ вын}}(t)] U_k(x),$$

где $T_{k \text{ св}}(t)$ — свободные колебания системы, соответствующие решению (3.21) при заданных начальных условиях, а $T_{k \text{ вын}}(t)$ — вынужденное движение, соответствующее решению уравнения

$$T_k'' + 2g T_k' + \lambda_k^2 E T_k = F_k(t),$$

где

$$F_k(t) = \frac{2}{l} \int_0^l J_z F(t) U_k(x) dx = - \frac{4 b_k J_z}{q_k} F(t).$$

При необходимости можно уточнить общее решение для модели первого приближения, принимая во внимание различные факторы реального объекта. Прежде всего учтем влияние энергии вращения поперечного сечения и сдвига по [2.10], поскольку при существующем отношении геометрических параметров кончика языка (его толщины, ширины и длины) эти факторы довольно заметны:

$$\lambda_k = \rho_k^2 \sqrt{\frac{J_z}{\rho}} \left[1 - \frac{\rho_k J_z}{2S} \left(1 + \frac{F}{\eta G} \right) \right],$$

где G — модуль сдвига; η — коэффициент формы сечения.

Форма кончика языка в плоскости xoz напоминает параболу, поэтому зависимость площади поперечного сечения от координаты можно выразить как $S(x) = S_0(1 - x^2/l^2)$, где S_0 — площадь поперечного сечения на границе между кончиком и телом языка. Для такой функции $S(x)$ можно рассчитать поправки $\delta\lambda_k$ к собственным числам λ_k методом возмущения [2.2]:

$$\lambda_k^{(1)} = \lambda_k + \delta\lambda_k,$$

где $\delta\lambda_k = \int_0^l (x^2 U''_k)'' U_k dx / \int_0^l (x U_k)^2 dx$. По оценке [2.10] для параболического закона изменения площади поперечного сечения значение первого собственного числа λ_1 возрастает примерно вдвое по сравнению с телом, имеющим постоянную площадь сечения S_0 . Вследствие относительно малой длины кончика языка по сравнению с его шириной можно предположить, что, так же как и для губ, основную роль в его упругих деформациях играет первая гармоника, тогда как вклад высших гармоник (для $k > 2$) будет сравнительно невелик.

Особый режим движений кончика языка возникает при артикуляции звука P . В этом случае, так же как и для голосовых складок, появляются автоколебания, вызванные переменным потоком воздуха в щели, образованной кончиком языка и твердым небом. Уравнение для скорости воздушного потока здесь то же, что для голосовой щели и для губ. Площадь щели $S_{щ}(x, t)$ определяется формой твердого неба $u_\tau(x, z)$ и кончика языка $u(x, t)$:

$$S_{щ}(x, t) = \int_0^d [u_\tau(x, z) - u(x, t)] dz,$$

где d — ширина щели.

В этом случае вынуждающая сила для уравнения движений кончика языка помимо мышечных усилий содержит и аэродинамическую компоненту, в результате чего уравнение колебаний принимает следующий вид:

$$\frac{\partial^2}{\partial x^2} \left(EJ_z \frac{\partial^2 u}{\partial x^2} \right) + \rho \frac{\partial^2 u}{\partial t^2} + r \frac{\partial u}{\partial t} = -J_z F(t) + c_x \rho_0 S_{щ} \frac{v^2}{2} - \Delta P S_{щ}, \quad (3.23)$$

где c_x — коэффициент «подъемной» силы; S_n — площадь верхней поверхности кончика языка; ΔP — избыточное давление в ротовой полости; v — скорость воздушного потока. Уравнение (3.23) является нелинейным, поскольку аэродинамическая компонента вынуждающей силы зависит от степени подъема кончика языка. Для его решения может быть применен метод последовательных приближений, описанный в § 2.6.

Движения кончика языка совершаются за счет сокращения внутренних мышц, поэтому модуль упругости его тканей изменяется во времени и вновь может быть использовано решение в беселевых функциях, как это было описано для голосовых складок и губ.

УПРУГИЕ ДЕФОРМАЦИИ ТЕЛА ЯЗЫКА

Приступим к рассмотрению деформаций языка в целом, причем в качестве деформируемого тела примем лишь ту часть языка, которая ограничена сверху и снизу продольными мышцами. Подбородочно-язычную мышцу будем считать упругим основанием языка с переменной жесткостью и источником внешних распределенных усилий. Такой подход сильно упрощает аналитические построения, а погрешность за счет отброшенной массы подбородочно-язычной мышцы оказывается сравнительно малой (по оценке [2.10] ошибка в определении частоты первого резонанса в случае равенства масс тела и упругого основания менее 1%).

Средняя линия описанного таким образом тела языка весьма близка к дуге полуокружности, так что уравнение моментов в общем виде $K = -M/EJ_z$, где K — кривизна деформированной средней линии, удобно представить в полярных координатах (R, φ) , что дает по [1.5]

$$\frac{R^2 - RR' + 2R'^2}{(R^2 + R'^2)^{3/2}} = -\frac{M}{EJ_z}, \quad (3.24)$$

где $R(\varphi)$ — радиус средней линии; M — изгибающий момент; E — модуль упругости; J_z — момент инерции поперечного сечения относительно оси z . Пусть средняя линия языка $R(\varphi)$ в нейтральном положении лишь незначительно отличается от дуги полуокружности с радиусом R_0 , т. е.

$$R(\varphi) = R_0 + \varepsilon \psi(\varphi), \quad 0 \leq \varphi \leq \pi, \quad (3.25)$$

где $\psi(\varphi)$ — поправочная функция; ε — малый коэффициент. Подставляя (3.25) в (3.24) и рассматривая лишь малые деформации с точностью до членов первого порядка, находим

$$\frac{\partial^2 u}{\partial \varphi^2} + u = -\frac{Mu^2}{EJ_z},$$

где $u(\varphi)$ — деформация средней линии. Далее, учитывая инерционность и вязкое трение, получаем уравнение колебаний

$$\frac{1}{[R_0 + \varepsilon \psi(\varphi)]} \frac{\partial^2}{\partial \varphi^2} \left(EJ_z \frac{\partial^2 u}{\partial \varphi^2} + u \right) + cu + r \frac{\partial u}{\partial t} + \rho \frac{\partial^2 u}{\partial t^2} = - \frac{\partial^2 M}{\partial \varphi^2},$$

где жесткость упругого основания $c \approx E_n/R_0$; E_n — модуль упругости подбородочно-язычной мышцы; ρ — погонная плотность тканей; r — погонный коэффициент вязкого трения. Пренебрегая членами второго порядка малости, получаем

$$\begin{aligned} & \frac{1}{R_0^2} \frac{\partial^2}{\partial \varphi^2} \left(EJ_z \frac{\partial^2 u}{\partial \varphi^2} + u \right) + \left[1 + \frac{2\varepsilon}{R_0} \psi(\varphi) \right] \left(cu + r \frac{\partial u}{\partial t} + \rho \frac{\partial^2 u}{\partial t^2} \right) = \\ & = - \left[1 + \frac{2\varepsilon}{R_0} \psi(\varphi) \right] \frac{\partial^2 M}{\partial \varphi^2}. \end{aligned} \quad (3.26)$$

Гистологические и рентгенографические исследования формы языка в нейтральном положении убеждают в том, что коэффициент ε очень мал, так что уравнение (3.26) вполне можно рассматривать как задачу с малым возмущением, а практически, по-видимому, достаточно ограничиться решением невозмущенной задачи, представленной в виде уравнения

$$\frac{1}{R_0^2} \frac{\partial^2}{\partial \varphi^2} \left(EJ_z \frac{\partial^2 u}{\partial \varphi^2} + u \right) + cu + r \frac{\partial u}{\partial t} + \rho \frac{\partial^2 u}{\partial t^2} = - \frac{\partial^2 M}{\partial \varphi^2}. \quad (3.27)$$

Как видно из рис. 3.28, площадь поперечного сечения языка увеличивается к его корню. Поэтому представим момент инерции поперечного сечения как $J_z(\varphi) = J_{z0}[1 + \xi(\varphi)]$, где J_{z0} — постоянная величина; $\xi(\varphi)$ — некоторая функция. Примем в качестве первого приближения, что площадь и момент инерции поперечного сечения языка постоянны. Влияние же переменного сечения обсудим несколько позже. Снова применим метод разделения переменных и попытаемся определить вид граничных условий для тела языка. Граничные условия для кончика языка, как и прежде, примем соответствующими свободному положению:

$$\left. \frac{\partial^2 u(\varphi, t)}{\partial \varphi^2} \right|_{\varphi=\varphi_l} = \left. \frac{\partial^3 u(\varphi, t)}{\partial \varphi^3} \right|_{\varphi=\varphi_l} = 0,$$

где φ_l — значение угла в полярной системе координат для кончика языка ($\varphi_l \approx \pi$).

Граничные же условия для корня языка определяются степенью напряженности мышц, управляющих его положением. Таким образом, можно предположить, что эти граничные условия не постоянны, а зависят от произносимого звука. Тем не менее можно попытаться найти доминирующий вид граничных условий. В описываемых ниже экспериментах сравнивалось жесткое и шарнирное закрепление, причем выяснилось, что реальные граничные условия близки к последнему виду закрепления.

Граничным условием для шарнирного закрепления соответствуют следующие собственные функции:

$$U_k(\varphi) = \operatorname{sh} p_k \varphi + \frac{p_k^2}{q_k^2} \sin q_k \varphi \frac{\operatorname{sh} p_k \varphi + p_k \operatorname{ch} p_k \varphi}{\sin \pi q_k + q_k \cos \pi q_k},$$

где $q_k^2 = 1 + p_k^2$. Собственные числа p_k здесь зависят от длины языка или максимального угла φ_l его радиуса-вектора. Согласно принятой системе координат для нейтральной формы языка $\varphi_l = \pi$, и соответствующие этому значению φ_l собственные числа p_k есть $p_1 = 0,8544$; $p_2 = 2,0347$; $p_3 = 3,1006$; $p_4 = 4,1353$; $p_5 = 5,1569$.

Как видно, в постановке задачи описания упругих колебаний языка имеется целый ряд неизвестных факторов, таких как зависимость коэффициентов уравнения от пространственной координаты и времени, вид граничных условий. Кроме того, существует принципиальная трудность в измерении распределения мышечных усилий и вызванных ими деформаций языка. Все это не позволяет прямо проверить любую теоретическую модель путем сравнения измеренных и теоретических вычисленных деформаций языка при заданных параметрах и усилиях. Однако существует другой, косвенный, путь оценки работоспособности модели. Этот путь состоит в выборе таких коэффициентов и граничных условий для дифференциального уравнения (3.27), чтобы его можно было решить аналитически, получив при этом явную зависимость для собственных функций U_k . Затем, пользуясь методом Ритца, формально представить истинные (неизвестные) собственные функ-

ции $\bar{U}_k$ как суперпозицию от U_i : $\bar{U}_k = \sum_{i=1}^N b_{ik} U_i$. Постулируем, что наиболее близким к действительности является такое множество собственных функций $\{U_i\}$ (и связанная с ним модель), которое имеет минимальное N при заданной ошибке ϵ аппроксимации поверхности языка этими функциями.

Таким образом, можно описать деформации языка через собственные функции модели некоторого приближения $\{U_i\}$, не вычисляя истинных собственных функций $\{\bar{U}_k\}$. При этом становится несущественным вид зависимости параметров дифференциальных уравнений от пространственной координаты.

Эксперименты по аппроксимации поверхности языка собственными функциями уравнений (3.19) и (3.27) проводились на материале рентгенограмм двух видов. Одна серия рентгенограмм содержала контуры речевого тракта для девяти гласных звуков шведского языка, произносимых протяжно и изолированно. Другая серия снималась для отдельных звуков и слогов русской речи, произносимых в довольно быстром темпе со скоростью 75 кадр./с и временем экспозиции 3 мс. Речевой материал содержал все звуки русского языка. Для оконтуривания выбирались кадры, непосредственно следующие за подъемом мягкого неба для всех звуков, за исключением назальных, где, наоборот, выбирался первый кадр, следующий за раскрытием прохода в носовую полость.

Таким образом, экспериментальный материал позволял сопоставить два языка и двух дикторов, а также сравнить результаты аппроксимации для статического и динамического произнесения звуков.

Прежде всего проверке были подвергнуты модели движения языка как твердого тела под влиянием поворота нижней челюсти и смещения корня языка. Для этого на рентгенограммах шведских гласных были выбраны два звука, отличающиеся наибольшим расстоянием между положениями нижней челюсти и корня языка. Это оказались звуки *ж* и *и*. Из положения, соответствующего звуку *и*, язык был перенесен в положение, соответствующее звуку *ж*, не меняя своей формы. Затем были произведены вычисления формы поверхности языка для положения нижней челюсти и корня языка, соответствующих звуку *ж* и *и*. Результаты вычислений сопоставлялись с исходной формой поверхности языка для звука *ж* и *и* (рис. 3.36). Расхождение между вычисленной и истинной формами языка оказалось чрезвычайно малым; еще меньше оно для других звуков, у которых отклонение положений нижней челюсти и корня языка не столь велико. Это подтверждает высокую точность моделей, использованных для описания эффектов вращения нижней челюсти и смещения корня языка.

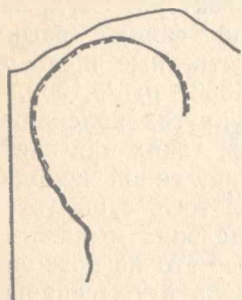


Рис. 3.36. Аппроксимация поверхности языка (— — —); измерения на кинорентгенограммах (—)

При аппроксимации поверхности языка использовались одна (первая) собственная функция для кончика языка и несколько собственных функций для языка в целом. При этом для кончика языка граничные условия закрепления принимались как жесткие, а для языка в целом — либо как жесткие, либо как шарнирные. Перед аппроксимацией из истинной формы языка $\tilde{\Phi}(\varphi)$ вычиталась форма нейтрального положения $\Phi_0(\varphi)$. Полученная функция $\Phi(\varphi) = \tilde{\Phi}(\varphi) - \Phi_0(\varphi)$ аппроксимировалась суммой собственных функций $\Phi^*(\varphi) = \sum_{k=1}^N c_k \psi_k(\varphi)$ исходя из критерия минимума некоторой нормы разности

$\Delta\Phi(\varphi) = \Phi(\varphi) - \Phi^*(\varphi)$.

Для шведских гласных отсутствовала рентгенограмма нейтрального состояния голосового тракта. Она была получена искусственно путем усреднения изображений всех гласных, кроме сильно продвинутого назад звука *и*.

Рассматривались два вида метрик для вычисления меры близости исходной поверхности $\Phi(\varphi)$ и аппроксимирующей ее $\Phi^*(\varphi)$. Оценка расстояния в метрике L_2 через $\rho_{L_2}(\Phi, \Phi^*) =$

$$\Delta\Phi(\varphi) = \Phi(\varphi) - \Phi^*(\varphi).$$

Для шведских гласных отсутствовала рентгенограмма нейтрального состояния голосового тракта. Она была получена искусственно путем усреднения изображений всех гласных, кроме сильно продвинутого назад звука *и*.

Рассматривались два вида метрик для вычисления меры близости исходной поверхности $\Phi(\varphi)$ и аппроксимирующей ее $\Phi^*(\varphi)$. Оценка расстояния в метрике L_2 через $\rho_{L_2}(\Phi, \Phi^*) =$

$= \left\{ \int_0^{\varphi_l} [\Delta\Phi(\varphi)]^2 d\varphi \right\}^{1/2}$ позволяет аналитически определить минимизируемый функционал и сравнительно просто реализуется на ЭВМ. Здесь φ_l — значение угла радиуса-вектора языка, которое он принимает на кончике языка (угол отсчитывается от корня языка). Однако во многих случаях качество аппроксимации при использовании этой метрики неудовлетворительно. В рассматриваемой задаче аппроксимация поверхности языка среднеквадратическим методом оказалась заметно хуже, чем при использовании оценки степени близости $\Phi(\varphi)$ и $\Phi^*(\varphi)$ в метрике C , где расстояние измеряется как $\rho_C(\Phi, \Phi^*) = \max |\Delta\Phi(\varphi)|$. В метрике C требуется гораздо большее число вычислительных операций при поиске минимума расстояния, и этот метод зависит от конкретной схемы поиска, начальных условий и ограничений на диапазон значений параметров. В данном случае применялся метод минимизации, в котором посредством малых пробных смещений коэффициентов c_k определялось наиболее выгодное направление, а затем в этом направлении делался большой скачок.

Представляется интересным рассматривать порознь погрешность аппроксимации для полной поверхности языка и его передней части. Это связано с большой чувствительностью акустических процессов к изменениям формы именно передней части языка. Поэтому вычислялись два вида ошибок:

$$\varepsilon_1 = \max |\delta\Phi(\varphi)| \operatorname{sign} [\delta\Phi(\varphi)], \quad 0 \leq \varphi \leq \varphi_l; \quad (3.28)$$

$$\varepsilon_2 = \max |\delta\Phi(\varphi)| \operatorname{sign} [\delta\Phi(\varphi)], \quad 0,5 \varphi_l \leq \varphi \leq \varphi_l, \quad (3.29)$$

$$\text{где } \operatorname{sign} [\delta\Phi(\varphi)] = \begin{cases} +, & \delta\Phi(\varphi) \geq 0; \\ -, & \delta\Phi(\varphi) < 0; \end{cases} \quad \delta\Phi(\varphi) = \Delta\Phi(\varphi)/\Phi_0(\varphi).$$

Процесс минимизации проводился относительно всей поверхности языка, и для найденного минимума определялись величины ε_1 и ε_2 . При минимизации в метрике L_2 абсолютные ошибки по (3.28) и (3.29) для шведских гласных были соответственно $|\varepsilon_1|_{L_2} = 17,5\%$ и $|\varepsilon_2|_{L_2} = 8,4\%$, тогда как при минимизации в метрике C эти ошибки были значительно меньше: $|\varepsilon_1|_C = 5,5\%$ и $|\varepsilon_2|_C = 4,5\%$. На основании этого сопоставления остальные эксперименты были проведены только для минимизации в метрике C . Поверхность языка на рентгенограммах измерялась в полярной системе координат с отсчетами через 5° .

Рассматриваемая задача аппроксимации является некорректной по Адамару, т. е. небольшим отклонением измеренных поверхностей $\Phi(\varphi)$ могут соответствовать большие отклонения коэффициентов c_k . Это оказывает влияние на решение вопроса о пригодности той или иной модели упругих деформаций языка. Кроме того, разным множествам коэффициентов $\{c_k\}$ соответствуют различные переходные процессы. Неустойчивость коэффициентов c_k лежит в основе самой постановки задачи аппроксимации. Эта неустойчивость может возникать не только в результате случайных

ошибок измерений поверхности языка $\Phi(\varphi)$, но и вследствие ограниченной точности вычислений.

Для решения неустойчивых задач в [3.8] был предложен так называемый метод регуляризации, который состоит в добавлении специального демпфирующего члена в минимизируемый функционал μ :

$$\mu[\Delta\Phi, \alpha, \{c_k\}] = \|\Delta\Phi\|^2 + \alpha\|c_k\|^2, \quad (3.30)$$

где $\|\cdot\|$ — норма; α — некоторая функция от ошибок измерения, которая, будучи правильно выбранной, гарантирует устойчивость решения задачи аппроксимации относительно коэффициентов c_k . Согласно [3.8] норма регуляризирующего члена $\|c_k\|^2$ вычислялась в метрике L_2 как сумма квадратов коэффициентов c_k . Норма основного члена функционала (3.30), как указывалось выше, вычислялась в большинстве случаев в метрике C . Значение коэффициента регуляризации α , в принципе, может быть вычислено по исходным данным задачи (по распределению ошибок измерений), однако процесс этих вычислений весьма трудоемок. В описываемых экспериментах осуществлялся независимый перебор по α в некотором диапазоне значений. Такой перебор в определенном смысле лучше, чем альтернативный вариант в виде включения величины α в процессе минимизации, так как он позволяет оценить влияние коэффициента регуляризации на устойчивость решения и ошибку аппроксимации. На рис. 3.37 показана зависимость ошибок аппроксимации от величины коэффициента регуляризации α для трех шведских гласных: y , u , e . Из этого рисунка видно, что увеличение α в большей или меньшей степени ухудшает точность аппроксима-

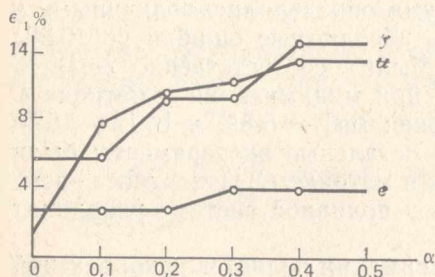


Рис. 3.37. Зависимость модуля ошибки аппроксимации от коэффициента регуляризации для передней части языка

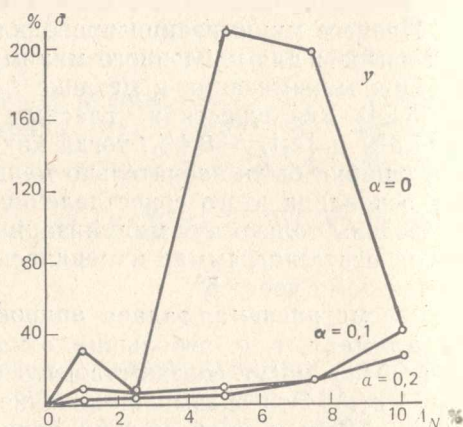


Рис. 3.38. Зависимость среднеквадратического отклонения σ коэффициентов c_k при собственных функциях языка от относительного уровня шумов

ции. Это ухудшение наблюдается в большинстве случаев, хотя изредка встречаются примеры улучшения точности в результате использования регуляризации. Однако можно пожертвовать в некоторой степени точностью аппроксимации для того, чтобы достичь стабильности динамических характеристик при синтезе речи.

Необходимость регуляризации иллюстрируется рис. 3.38, где показана величина среднеквадратического отклонения σ коэффициентов c_k при воздействии случайных шумов на исходную поверхность $\Phi(\varphi)$ для шведской гласной y . Искаженная поверхность $\Phi_N(\varphi) = \Phi(\varphi) \{1 + A_N[0,5 - \xi(\varphi)]\}$, где $\xi(\varphi)$ — случайная величина, равномерно распределенная на интервале $0-1$; а A_N — относительный уровень шума, %. За исходную точку при вычислении среднеквадратического отклонения σ принималось множество $\{c_k\}$ при отсутствии шумов ($A_N=0$). Из рисунка видно, что диапазон отклонений $\{c_k\}$ без регуляризации недопустимо велик даже при весьма умеренном уровне шумов порядка 5%. Последующий спад σ объясняется случайной природой результатов однократного статистического испытания. Видно также, что даже сравнительно небольшой коэффициент регуляризации $\alpha=0,1$ хорошо стабилизирует коэффициенты c_k .

Надлежащая величина α в каждом случае должна была бы, в принципе, выбираться путем многократных статистических испытаний при наложении случайного шума с интенсивностью, равной ошибке измерений. Но вместо этой трудоемкой процедуры производилась минимизация ошибки при разных значениях α ($0-0,4$) и выбиралось наименьшее из тех значений α , для которых среднеквадратическое отклонение было минимальным.

Регуляризирующий член в рассматриваемой задаче имеет не только математическое обоснование, но и физическую интерпретацию — он связан с энергетическими затратами мышц на формирование той или иной поверхности языка. Включение энергии затрат в минимизируемый функционал, таким образом, связано с принципом экономии усилий, широко обсуждающимся в физиологии управления мышечными движениями.

В описываемых экспериментах были рассчитаны только собственные неуправляемые переходные процессы языка от нейтральной формы к форме, соответствующей какому-либо звуку. Принимались следующие значения параметров уравнений: площадь поперечного сечения языка $S=2 \text{ см}^2$; погонная плотность $\rho = \rho_\tau S$, где $\rho_\tau = 1,1 \text{ г/см}^3$; момент инерции $J_z = 0,4 \text{ см}^4$ — модуль упругости $E = 10^6 \text{ Па}$; погонная упругость $c = 25 \text{ г/см} \cdot \text{с}^2$; погонное сопротивление $r = 170 \text{ г/см} \cdot \text{с}$. Переходные процессы вычислялись как свободные колебания, соответствующие сумме решений обыкновенных дифференциальных уравнений второго порядка, взятых с коэффициентами c_k . Корни этих уравнений

$$\beta_{1,2k} = -\frac{r}{2\rho} \pm \sqrt{\left(\frac{r}{2\rho}\right)^2 - \left(\frac{c}{\rho} + \lambda_k^4\right)},$$

где $\lambda^4_k = \frac{EJ_z}{\rho} (p_k q_k)^2$; p_k — собственные числа соответствующих собственных функций, форма которых для разных типов граничных условий и их переходные процессы показаны на рис. 3.39 и 3.40.

Прежде чем рассмотреть результаты экспериментов, необходимо выяснить возможные источники ошибок, которые могут повлиять на точность аппроксимации.

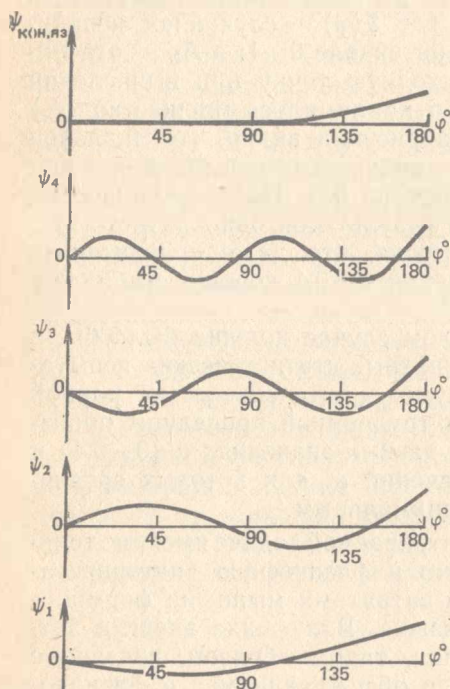


Рис. 3.39. Собственные функции языка и его кончика

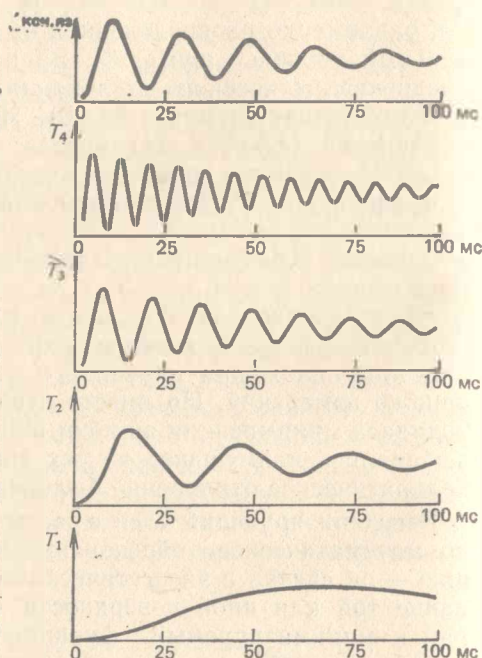


Рис. 3.40. Переходные процессы собственных колебаний языка и его кончика

Контуры речевого тракта на рентгенограммах измерялись с увеличением примерно в 10 раз по сравнению с естественными размерами речевого тракта. Точность измерения равнялась 1 мм, так что в этом случае основную роль играли ошибки в оконтуривании полутонных рентгенограмм. Эти ошибки примерно равны 3—5%. Кроме того, форма языка была найдена как средняя линия между контурами середины языка и его краями, т. е. прогиб языка вблизи его плоскости симметрии не принимался во внимание.

Следующий источник ошибок заключается в способе выбора нейтральной формы языка, которая вычиталась из формы языка для всех звуков. Неправильный выбор нейтральной формы языка

приводит к искажению аппроксимируемой поверхности, однако дать какие-либо количественные оценки в этом случае весьма затруднительно. Можно лишь предположить, что если форма, принятая за нейтральную, содержит элементы некоторой деформации, то смещенная в результате этого оценка коэффициентов c_k при собственных функциях больше повлияет на характер переходного процесса, чем на точность аппроксимации. Использование регуляризации также ухудшает точность в среднем не более чем на 1%, хотя для отдельных звуков это ухудшение может быть и большим. Наконец, точность аппроксимации может уменьшиться за счет того, что процедура аппроксимации сходится не к глобальному минимуму, а к одному из множества локальных минимумов. В процессе экспериментов иногда наблюдалась существенная разница в результатах поиска минимума ошибки аппроксимации при различных начальных условиях. Поскольку общий результат аппроксимации оказался вполне приемлемым, не было сделано попытки к поиску глобального минимума, что позволило сократить время эксперимента во много раз¹.

Результаты вычислений ошибки аппроксимации при различных условиях представлены в табл. 3.7 и 3.8. В табл. 3.7 даны значения средних ошибок, полученные на материале русского языка, для двух видов граничных условий — жесткого и шарнирного закрепления языка. Кроме того, учитывается регуляризация для различного числа собственных функций, применяемых для аппроксимации. Ошибки указываются для отдельно произносимых звуков, а также для звуков, произнесенных в составе слогов (общие средние ошибки).

Рассматривая табл. 3.7, отмечаем, что нет принципиальной разницы между точностью аппроксимации для изолированных звуков и звуков в слогах. По-видимому, при довольно быстром темпе про-

Таблица 3.7

Ошибки	Ошибка аппроксимации для русских звуков, %, при закреплении											
	шарнирном						жестком					
	с регуляризацией		без регуляризации				с регуляризацией		без регуляризации			
	N=5		N=5		N=4	N=3	N=5		N=5		N=4	
	звуки	общие	звуки	общие	звуки	общие	звуки	общие	звуки	общие	звуки	общие
	звуки	общие	звуки	общие	звуки	общие	звуки	общие	звуки	общие	звуки	общие
ε_1	7,4	7,5	6,8	7,6	7,7	11,5	9,2	10,4	8,8	9,9	10,7	
ε_2	6,3	6,4	5,9	6,2	6,3	9,7	6,8	7,6	6,3	7,1	8,5	

¹ Возможный выигрыш в точности при поиске глобального минимума может быть проиллюстрирован результатом экспериментов по аппроксимации формы языка для английских гласных [3.2]. В этих экспериментах ошибка аппроксимации на передней части языка не превышала 5% (при ошибке в области корня 8—10%) даже в случае использования всего трех собственных функций, соответствующих граничным условиям жесткого закрепления, которые, как выяснилось в последующем, обеспечивают не самый лучший результат.

Таблица 3.8

Гласные	Ошибка	Ошибки аппроксимации для гласных, %, при закреплении			
		с регуляризацией	без регуляризации	с регуляризацией	без регуляризации
Русские	$ \varepsilon_1 $	7,5	7,7	8,6	8,1
	$ \varepsilon_2 $	6,7	7,2	7,7	6,4
Шведские	$ \varepsilon_1 $	5,5	5,7	9,2	9,3
	$ \varepsilon_2 $	4,5	4,6	5,9	6,0

изнесения различие в скорости артикуляции отдельных звуков и звуков в слогах, которое могло бы повлиять на граничные условия закрепления языка, невелико, хотя этот вывод не претендует на окончательность ввиду ограниченного числа исследованных слогов. Невелика разница и между формами языка для одного и того же звука, произнесенного изолированно и в слоге.

В табл. 3.8 сравнивается «статическое» произнесение шведских гласных с «динамическим» русских гласных. Гласные произносились изолированно, но в разных темпах: шведские — протяжно, а русские — быстро. При вычислениях использовалось фиксированное число собственных функций $N=5$, тогда как вид граничных условий различался, а регуляризация либо применялась, либо нет.

Наибольшее влияние на точность аппроксимации оказывают вид граничных условий и число собственных функций N (напомним, что для кончика языка всегда использовалась только первая собственная функция). Закрепление корня языка в виде шарнира, очевидно, лучше соответствует реальным условиям, чем жесткое закрепление. В действительности при повороте корня языка относительно своего конца возникает не нулевой момент сопротивления, но существует линейное его смещение. Шарнир не создает сопротивления повороту, но препятствует линейным смещениям, так что этот вид граничных условий не вполне соответствует действительному положению дел, хотя, несомненно, выбор шарнирного закрепления более адекватен, чем жесткого. Например, при числе собственных функций $N=4$ для граничных условий типа шарнир точность аппроксимации все еще остается выше, чем при $N=5$ для жесткого закрепления.

Роль гибкости закрепления корня языка становится еще более заметной при сравнении общей средней ошибки аппроксимации ε_1 и ошибки на передней части языка ε_2 . Как видно из табл. 3.8, ошибка на передней части языка мало отличается для разных видов граничных условий, тогда как вблизи корня языка, т. е. там, где влияние условий закрепления наиболее велико, различие в точности аппроксимации становится весьма заметным. Сопоставляя истинные поверхности языка и их аппроксимации, нетрудно

заметить, что ошибки аппроксимации вблизи корня языка отрицательны. Это говорит о недостаточной степени «колебательности» используемых собственных функций, однако преодолеть этот недостаток простым добавлением собственных функций, имеющих большее число пересечений нуля, вряд ли возможно. Об этом свидетельствует сравнение ошибок для разного числа собственных функций: если при переходе от $N=3$ к $N=4$ ошибка уменьшается в среднем примерно на 4%, то использование пяти собственных функций вместо четырех дает выигрыш лишь около 1%, поэтому увеличение числа собственных функций больше шести нецелесообразно. Для достижения лучшей аппроксимации в области корня языка нужно было бы использовать более мягкие относительно смещения граничные условия.

При рассмотрении табл. 3.8 можно заметить расхождение значений ошибок при разных типах граничных условий. Так, для шведских гласных различие по общей ошибке ϵ_1 достигает почти 4%, тогда как для русских гласных оно около 1%. Необходимо принять во внимание и факторы, непосредственно не касающиеся граничных условий. Это, во-первых, отличия в артикулярном строе шведского языка, характеризующегося более «задним» (по терминологии фонетистов) ненапряженным способом генерирования звуков. В результате меньшего напряжения создается и меньшая деформация формы языка, что соответственно способствует лучшей аппроксимации. Другая особенность, касающаяся специфики этого случая, состоит в том, что нейтральная форма языка для шведских гласных была вычислена путем усреднения имеющих рентгенограмм гласных, так что она была ближе к исследуемым гласным, чем истинная нейтральная форма. Однако этот фактор, как и особенность звукового строя шведского языка, мог бы объяснить только лучшее качество аппроксимации шведских гласных по сравнению с русскими гласными при шарнирном закреплении. При жестком закреплении как для русских, так и для шведских гласных значения ошибок близки, так что рассмотренные особенности шведских гласных лишь частично оправдывают такую разницу в точности аппроксимации при разных типах граничных условий. Вероятно в данном случае играет роль темп артикуляции — для шведских гласных в граничных условиях отсутствует динамический элемент, тогда как для быстро артикулируемых русских звуков изменение жесткости мышц, управляющих положением корня языка (и, возможно, изменением жесткости самого языка), приводят к синхронному изменению формы собственных колебаний. Граничные условия можно было бы уточнить, воспользовавшись измерениями импеданса мышечных тканей, но он зависит от места измерения, испытываемого человека, и амплитуды колебаний. Кроме того, следует знать скорость нарастания и степень напряженности мышц, управляющих положением корня языка. Если учесть параметры импеданса закрепления корня языка при минимизации ошибки аппроксимации, то вполне возможно, что удастся сократить ошибки аппроксимации до очень малых

значений. Однако такой процесс минимизации потребует значительных затрат машинного времени при проведении эксперимента и вычисления новых собственных функций для каждого значения параметров граничных условий. Следует помнить, что полученные результаты указывают на вполне удовлетворительную точность математической модели, а ошибки в 6—7% на практике (в частности, для синтеза речи) приемлемы, если учесть, например, значительную неопределенность в выборе способа вычисления площади поперечного сечения голосового тракта по его форме в саггитальной плоскости. Кроме того, наибольшие ошибки приходится на область вблизи корня языка, обычно прикрытую надгортанником, и к тому же влияние этой области на формирование акустических характеристик речевого сигнала не столь велико, как влияние передней части языка.

Значения ошибок ϵ_1 и ϵ_2 , коэффициента регуляризации α и коэффициентов c_k для звуков русского языка приведены в табл. 3.9. Соответствующие этим коэффициентам расчетные формы голосового тракта (штрих) показаны на рис. 3.41. Сопоставляя коэффициенты c_k для безударного (первого) звука A в слогах $АБА$, $АДА$, $АГА$, можно заметить, что эти коэффициенты отличаются друг от друга, как и соответствующие им аппроксимационные кривые. Это еще раз подчеркивает неустойчивый характер решения задачи аппроксимации. Если же обратиться к переходным процессам для вычисленных конфигураций языка, то окажется, что, несмотря на различие коэффициентов c_k , переходные процессы для безударного A в разных слогах имеют много общего. К тому же определенные различия в форме голосового тракта для одного и того же звука, в одной и той же позиции, но для разных слогов вполне

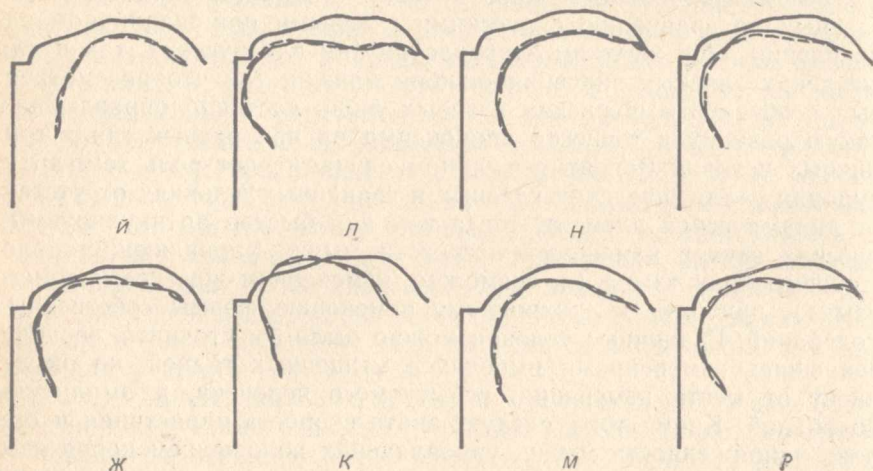


Рис. 3.41. Аппроксимация поверхности языка (— — —); измерения на кинорентгенограммах (——) для некоторых звуков русского языка

Таблица 3.9

Звук		Ошибка		Параметры аппроксимации для звуков русского языка					
		общая ϵ_1	на передней части языка ϵ_2	Коэффициент регуляризации	Коэффициенты при собственных функциях				кончика языка
					языка				
					1	2	3	4	
О	изолированно	-8,3	-8,5	0,1	-0,85	0,325	0,3	-0,0875	0,0
У	—»—	-6,6	5,7	0,2	0,475	0,05	0,2625	0,075	0,025
И	—»—	-7,3	-5,9	0,4	-0,2125	-0,0375	0,1125	-0,0375	0,0
Ы	—»—	-8,0	-8,0	0,4	-0,3125	0,175	0,175	0,0	-0,3
Э	—»—	-8,9	6,2	0,4	0,3375	0,0875	0,225	-0,05	0,0
Г	—»—	6,9	6,5	0,1	-0,7125	-0,15	0,2875	-0,025	0,6625
Д	—»—	-9,1	-7,8	0,2	0,4125	0,2	0,1125	-0,025	-0,1375
З	—»—	-8,6	-5,5	0,4	0,3375	0,3	0,3	-0,08	0,0
Ж	—»—	-7,9	6,0	0,4	0,1125	0,1125	0,0875	-0,0875	0,3
И	—»—	7,1	5,0	0,05	2,0	0,1	0,1	0,0	-2,3
К	—»—	-5,1	-4,9	0,05	-1,2	-0,275	0,4125	-0,1	1,4
Л	—»—	-8,3	7,6	0,1	-0,425	0,5625	0,3375	-0,1125	1,1
М	—»—	-6,1	-4,7	0,4	0,3125	-0,0875	0,175	0,075	0,2625
Н	—»—	-4,5	-4,2	0,05	1,0	0,325	0,15	0,0125	-0,8
С	—»—	-8,8	6,8	0,3	0,1875	0,1875	0,125	0,05	0,1875
Р	—»—	-6,3	5,5	0,05	-0,035	0,3	0,35	0,0	1,3
Т	—»—	-5,4	-5,3	0,4	0,35	0,3125	0,1375	-0,0315	0,025
Х	—»—	-8,6	7,5	0,1	-0,525	0,45	0,1125	-0,1875	0,35
Ш	—»—	-5,8	-5,0	0,4	0,125	0,1875	0,1625	-0,075	0,0
Б	—»—	-7,1	-5,3	0,1	0,65	-0,3125	0,025	0,0	-0,0125
В	—»—	-5,5	3,8	0,4	0,3625	-0,225	0,1125	-0,0125	0,325
Г	—»—	-4,3	2,9	0,2	0,4	-0,425	0,1875	0,075	0,125
Д	—»—	-9,9	-7,5	0,2	0,9125	-0,425	0,15	-0,1	-0,025
А	(в слоге АБА)	-6,0	-6,0	0,3	0,075	0,45	0,1125	-0,0625	-0,3
Б	(в слоге АБА)	-5,4	-5,4	0,2	-0,225	0,375	0,15	-0,1	-0,3
А	(в слоге ИА)	-12,6	-12,6	0,4	-0,05	0,2	-0,05	0,075	-0,1625
И	(в слоге ИО)	9,1	-6,2	0,05	1,5	-0,3	-0,05	-0,225	-0,9
О	(в слоге ИО)	-11,4	-10,3	0,4	-0,1	-0,225	-0,35	-0,0625	-0,075
А ₁	(в слоге А ₁ Д ₂)	-7,0	-6,5	0,4	0,225	0,25	0,1	0,0375	-0,15
Д	(в слоге А ₁ Д ₂)	-7,2	-6,1	0,4	0,6	0,2375	0,2125	0,0125	0,0625
А ₁	(в слоге А ₁ Г ₂)	-8,1	-7,1	0,2	0,1875	0,2875	0,2625	-0,0625	-0,1875

допустимы и отмечались неоднократно при исследовании рентгенограмм.

В дополнение к приведенному анализу результатов описанных экспериментов следует вспомнить сделанное в начале данного раздела предположение о постоянстве площади поперечного сечения тела языка и его момента инерции вдоль пространственной координаты. Как видно из рис. 3.28, толщина языка увеличивается у его корня. Можно оценить влияние этого увеличения, воспользовавшись методом возмущения и задавшись законом изменения толщины языка. Если принять линейный закон изменения площади

поперечного сечения какого-либо упругого тела, то по [2.10] значение его первого собственного числа увеличивается примерно в 1,5 раза по сравнению со значением первым собственного числа тела постоянного сечения. Таким образом, утолщение языка к его корню оказывает заметное влияние на значения собственных чисел, и в этом явлении заключается резерв повышения точности аппроксимации упругих деформаций языка.

Наряду с усовершенствованием методики анализа упругих деформаций языка возможен и другой подход, состоящий в статистическом анализе его формы. Так, в [3.26] методом факторного анализа были найдены две главные компоненты, связанные с движениями корня языка вверх—вниз и вперед—назад, с помощью которых достигается удовлетворительное описание многих реально встречающихся форм языка. Статистический анализ собственных функций измеренных на рентгенограммах форм поверхности языка дает примерно четыре-пять компонент [3.38, 3.47], довольно близких по форме к собственным функциям, получаемым в результате анализа упругих деформаций.

В заключение упомянем о надгортаннике, основная функция которого состоит в закрывании дыхательного горла во время приема пищи. Роль надгортанника в артикуляции исследована очень мало. Лишь в [3.35] сообщается о наблюдениях, позволяющих сделать заключение о его участии в артикуляции гласной А, фарингиальных взрывных, а также при шепоте.

АКУСТИКА РЕЧЕОБРАЗОВАНИЯ

4.1. УРАВНЕНИЕ РЕЧЕВОГО ТРАКТА

В теории речеобразования исследования акустических процессов играют важную роль [В.3, В.4, В.5, 4.57]. На основе методов физической акустики и электромеханических аналогий были установлены многие важные свойства акустических процессов в речевом тракте. Эти исследования почти исключительно касались статики речеобразования, тогда как накопленный в дальнейшем экспериментальный материал и более детальный теоретический анализ стали наводить на мысль о важной роли параметрических и нестационарных явлений. Поэтому возникла необходимость такого развития акустической теории речеобразования, которое учитывало бы податливость стенок речевого тракта, переменные граничные условия со стороны как губ, так и голосовых складок, изменения во времени площади поперечного сечения речевого тракта, взаимодействие голосового источника с речевым трактом и т. д.

Трактовка акустической системы как системы с распределенными или сосредоточенными параметрами зависит от соотношения ее геометрических размеров с длинами волн возбуждаемых в ней колебаний. Если некоторый характерный геометрический размер системы l гораздо меньше определенной длины волны λ_0 , т. е. $l \ll \lambda_0$, то для всех волн, длина которых больше λ_0 , акустическая система может считаться сосредоточенной. Конкретизируя это условие, например, как $10l \leq \lambda_0$, найдем граничную частоту колебаний f_0 , ниже которой речевой тракт можно рассматривать как систему с сосредоточенными параметрами: $f_0 \leq c_0/10l$, где c_0 — скорость звука в речевом тракте. Поскольку длина речевого тракта обычно не превышает 20 см (среднее ее значение по [В.3] равно 17,5 см), граничная частота $f_0 \approx 175$ Гц. Это означает, что во всем диапазоне резонансных частот речевой тракт должен описываться как система с распределенными параметрами.

Акустические колебания в системах с распределенными параметрами описываются уравнениями, тесно связанными с волновым уравнением. В свободной среде волновое уравнение выводится исходя из некоторых предположений относительно свойств среды и характера колебаний. Предполагается, что среда обладает свойством неразрывности и однородности, что скорости постоянных потоков в среде гораздо меньше скорости звука, а амплитуды звуковых колебаний малы, т. е. скорости частиц газа малы по сравнению с газокинетической скоростью молекул. Кроме того, допуска-

ется, что акустические колебания носят безвихревой характер, т. е. существует потенциал скорости. При этих предположениях трехмерное волновое уравнение для свободной среды относительно потенциала скорости Φ имеет вид

$$\frac{\partial^2 \Phi}{\partial x^2} + \frac{\partial^2 \Phi}{\partial y^2} + \frac{\partial^2 \Phi}{\partial z^2} = \frac{1}{c_0^2} \frac{\partial^2 \Phi}{\partial t^2}, \quad (4.1)$$

где v — скорость акустических колебаний: $v = -\text{grad } \Phi$; x, y, z — декартовские координаты; t — время; c_0 — скорость звука в воздухе. Границы применимости волнового уравнения определяются помимо ограничений, накладываемых упомянутыми предположениями, также величинами пространственных и временных дифференциалов: пространственные дифференциалы должны быть больше размеров молекул, но меньше длины волны, а временные — меньше периода колебаний [4.14].

Для адиабатических процессов скорость звука определяется как

$$c_0 = \sqrt{\gamma P_0 / \rho_0},$$

где $\gamma = 1,41$; P_0 — атмосферное давление; ρ_0 — плотность воздуха. Скорость звука не зависит от давления, поскольку при его изменении отношение P_0/ρ_0 остается постоянным, но зависит от температуры:

$$[c_0^2 = \frac{\gamma R}{M} T,$$

где T — температура по Кельвину; R — универсальная газовая постоянная: $R = 8,315 \cdot 10^3$ Дж/кмоль·град; M — молекулярный вес газа (для кислорода $M = 32$). Для воздуха при температуре 0°C скорость звука равна $c_0 = 333$ м/с, а для речевого тракта она оказывается несколько выше — около 350 м/с.

Для цилиндрической трубы волновое уравнение удобно записать в цилиндрической системе координат:

$$\frac{\partial^2 \Phi}{\partial r^2} + \frac{1}{r} \frac{\partial \Phi}{\partial r} + \frac{1}{r^2} \frac{\partial^2 \Phi}{\partial \varphi^2} + \frac{\partial^2 \Phi}{\partial z^2} = \frac{1}{c_0^2} \frac{\partial^2 \Phi}{\partial t^2}, \quad (4.2)$$

где r — радиус; φ — азимут. Применяя метод разделения переменных и совершая подстановку $\Phi = R(r)\Psi(\varphi)Z(z)T(t)$, получаем систему:

$$\frac{1}{c_0^2 T} \frac{\partial^2 T}{\partial t^2} = -K_t^2, \quad (4.3)$$

$$\frac{1}{Z} \frac{\partial^2 Z}{\partial z^2} = -K_z^2, \quad (4.4)$$

$$\frac{1}{\Psi} \frac{\partial^2 \Psi}{\partial \varphi^2} = -K_\varphi^2, \quad (4.5)$$

$$\frac{1}{R} \left(\frac{\partial^2 R}{\partial r^2} + \frac{1}{r} \frac{\partial R}{\partial r} \right) = \frac{K_\varphi^2}{r^2} - K_r^2, \quad (4.6)$$

где собственные числа временных мод акустических колебаний для (4.3)

$$K_t^2 = K_z^2 + K_r^2. \quad (4.7)$$

Общее решение (4.2) в виде осевых бегущих и стоячих волн для других направлений имеет вид

$$\Phi = \sum_{K_\Phi} \sum_{K_z} \sum_v [A_{K_\Phi} \cos(K_\Phi \Phi + \theta_{K_\Phi}) J_{K_\Phi}(K_{r\nu} r) + B_{K_\Phi} \cos(K_\Phi \Phi + \theta'_{K_\Phi}) N_{K_\Phi}(K_{r\nu} r)] e^{\pm i K_z z}, \quad (4.8)$$

где J_{K_Φ} — функции Бесселя первого рода; N_{K_Φ} — функции Неймана [4.14].

Если на оси цилиндрической трубы нет источников, то поле на оси должно быть конечно, поэтому из решения (4.8) исчезает второй член, так как $N_{K_\Phi}(0) = \infty$. Частоты радиальных резонансов определяются характеристиками боковых стенок трубы. Для абсолютно жестких стенок нормальная радиальная составляющая скорости звуковых колебаний при $r=a$ должна равняться нулю (a — внутренний радиус трубы), т. е.

$$i \frac{A_{K_\Phi}}{\rho_0 c_0} \frac{\partial}{\partial r} J_{K_\Phi}(K_{r\nu} r) \Big|_{r=a} = 0, \quad (4.9)$$

откуда $K_{r\nu} = \alpha_{K_\Phi \nu} / a$, $\nu = 1, 2, \dots$, где $\alpha_{K_\Phi \nu}$ — корни уравнения $J'_{K_\Phi} = 0$. Для нулевой моды ($K_\Phi = 0$) первый корень $\alpha_{01} = 3,83$ и частота первого радиального резонанса $f_{01} = 0,61 c_0/a$.

Для трубы прямоугольного сечения частота радиального резонанса $f_{01} = 0,5 c_0/L$, где L — наибольшая сторона прямоугольника [4.7]. Решение, удовлетворяющее этим граничным условиям для цилиндрической трубы, есть

$$\Phi_{m\nu} A_m \cos(m\Phi + \theta_m) J_m\left(\alpha_{m\nu} \frac{r}{a}\right) e^{\pm i K_{m\nu} z},$$

$$\text{где } K_{m\nu} = \sqrt{K_t^2 - K_{r\nu}^2} = K_t \sqrt{1 - \left(\frac{f_{m\nu}}{f_c}\right)^2}, \quad f_{m\nu} = \frac{\alpha_{m\nu} c_0}{2\pi a},$$

$$m = 1, 2, \dots$$

Если $f < f_{m\nu}$, то величина $K_{m\nu}$ становится мнимой и в трубе не возникают цилиндрические волны высшего порядка. Для цилиндрической трубы с площадью поперечного сечения 5 см^2 (примерно равной средней площади поперечного сечения речевого тракта) нижняя граничная частота возникновения цилиндрических волн $f_{10} = 0,293 c_0/a \approx 8 \text{ кГц}$. Для $m > 1$ граничная частота f_{m0} еще выше. Для трубы с той же площадью поперечного сечения, но с прямоугольным сечением или для большей площади поперечного сечения граничная частота первого радиального резонанса будет несколько меньше, но в основном диапазоне частот речевого сигнала

ла цилиндрические волны не возникают, а распространяются лишь плоские.

Трубы, геометрические размеры поперечного сечения которых меньше длины волны первого радиального резонанса $\lambda_{01} = c_0/f_{01}$, называются «узкими» трубами. Если же радиус a или наибольшая сторона прямоугольного сечения гораздо меньше λ_{01} , то такие трубы называются «очень узкими» трубами и они обладают тем свойством, что их изгибы или даже изломы не влияют на характер распространения волн [4.7]. Геометрические размеры речевого тракта обычно удовлетворяют условию «очень узкой» трубы, поэтому в первом приближении можно пренебречь влиянием его изгиба на распространение волн, что значительно упрощает задачу вывода и решения уравнения речевого тракта.

Для трубы с податливыми стенками, как это имеет место в речевом тракте, радиальные колебания присутствуют всегда независимо от соотношения геометрических размеров трубы и длин волн $\lambda_{m\nu}$. В этом случае граничное условие (4.9) принимает вид

$$P/V_r|_{r=a} = Z_w, \quad (4.10)$$

где P — акустическое давление; V_r — радиальная скорость акустических колебаний; Z_w — акустический импеданс стенок. Поскольку (4.2) записано относительно потенциала скорости Φ , то, рассматривая периодические процессы и представляя $\Phi = \Phi_r e^{-j\omega t}$, а также принимая во внимание, что

$$V_r = -\frac{\partial \Phi}{\partial r}; \quad (4.11)$$

$$P = \rho_0 \frac{\partial \Phi}{\partial t}, \quad (4.12)$$

получаем граничное условие для трубы с податливыми стенками

$$\frac{j\omega\rho_0}{Z_w} J_m(K_{rm}r)|_{r=a} = -\frac{\partial}{\partial r} J_m'(K_{rm}r)|_{r=a}. \quad (4.13)$$

Податливость стенок трубы приводит к тому, что уменьшение столба воздуха в трубе под воздействием давления обуславливается не только сжимаемостью самой среды, но и смещением стенок. Анализ физики этого явления позволяет перейти от (4.13) к более наглядному представлению, характеризующему зависимость скорости звука в трубе от свойств ее стенок в виде

$$c_0 = \frac{\bar{c}_0}{\sqrt{1 + \frac{\Delta S}{S} \frac{\rho_0 \bar{c}_0^2}{P}}}, \quad (4.14)$$

где ΔS — изменение площади поперечного сечения трубы за счет податливости стенок под воздействием давления P ; $\bar{c}_0$ — скорость звука в трубе с абсолютно жесткими стенками [4.7]. Выражение (4.14) является другой формой отображения эффекта повышения частот резонансных колебаний в трубе с податливыми стенками.

Для цилиндрической трубы $\Delta S/S = \pi[(a + \Delta a)^2 - a^2]/\pi a^2$, и при малых смещениях стенок, т. е. при $\Delta a \ll a$, $\Delta S/S \approx 2 \Delta a/a$.

Для прямоугольного сечения со сторонами a и b при $a \gg b$ $\Delta S/S = [(a + \Delta a)(b + \Delta b) - ab]/ab \approx \Delta b/b$.

Для цилиндрической трубы скорость радиальных колебаний стенок $V_r = d\Delta a/dt$. Если колебания стенок — гармонические, то $\Delta a = \Delta a_0 e^{-j\omega t}$ и проводимость стенок $Y_w = 1/Z_w$ можно выразить как $Y_w = -j\omega a_0/P$, и тогда уравнение (4.14) принимает вид

$$c_0^2 = \frac{\bar{c}_0^2}{1 - \frac{2Y_w \rho_0 \bar{c}_0^2}{j\omega a}} \quad (4.15)$$

Вводя удельную безразмерную проводимость $\bar{Y}_w = Y_w \rho_0 \bar{c}_0$, получаем

$$c_0^2 = \frac{\bar{c}_0^2}{1 - \frac{2\bar{Y}_w \bar{c}_0}{j\omega a}} \quad (4.16)$$

Можно показать, что (4.15) справедливо и для прямоугольного сечения, если за a принять наименьшую сторону. Если ширина прямоугольного сечения остается постоянной при изменении его площади, то характер изменения высоты a будет другим, чем характер изменения радиуса a круглого сечения, и изменение скорости звука в этих трубах будет по-разному зависеть от площади поперечного сечения [4.17]. Это свойство существенно для речевого тракта, поперечное сечение которого часто принимает форму овала, который лучше аппроксимируется не кругом, а прямоугольником со сторонами, отличающимися в несколько раз.

Как видно из (4.16), степени и знак приращения скорости звука c_0 в трубе с податливыми стенками зависят от проводимости стенок Y_w . Область применения (4.16) ограничивается толщиной стенок трубы, так как если она сравнима с длиной радиальной волны, то свойства стенок не исчерпываются их импедансом, а зависят от распределения звукового давления и скорости частиц вдоль трубы [4.15]. Стенки речевого тракта могут считаться тонкими по сравнению с длиной волны радиальных колебаний, и (4.16) справедливо для оценки акустических процессов в речевом тракте.

Импеданс стенок можно рассчитать по формуле для многослойных систем

$$Z_w = j\rho_w c_w \frac{Z_H - j\rho_w c_w \operatorname{tg} K_0 h}{Z_H \operatorname{tg} K_0 h + j\rho_w c_w} \quad (4.17)$$

где ρ_w — плотность стенок; c_w — скорость звука в материале стенок, равная $1,54 \cdot 10^5$ см/с [4.43]; h — толщина стенок; K_0 — волновое число в свободной среде ($K_0 = \omega/c_0$); Z_H — импеданс, на который нагружена стенка (для полубесконечной свободной среды $Z_H = \rho_0 c_0$) [4.7]. Стенку толщиной h , малой по сравне-

нию с длиной волны, можно считать системой с сосредоточенными параметрами, характеристики которой зависят от импеданса нагрузки Z_n . Если Z_n мал, то стенка является сосредоточенной массой, а если Z_n велик, — то сосредоточенной упругостью. Для сосредоточенной массы имеем $mdV/dt = P$, где $m = \rho_w h$ — поверхностная плотность вещества. Для гармонических колебаний давления $P = P_0 e^{-j\omega t}$ импеданс $Z_m = -j\omega m$. В этом случае суммируются импедансы стенок и нагрузки: $Z_w = Z_n - j\omega m$.

Для сосредоточенной упругости $E_w \xi / h = P$, где E_w — модуль упругости материала стенок; ξ — их смещение. Упругий импеданс $Z_{yn} = jc_{yn} / \omega$, где $c_{yn} = E_w / h$. В этом случае суммируются проводимости $Y_w = Y_n - j\omega / c_{yn}$. Если сосредоточенная масса нагружена на сосредоточенную упругость, опертую на жесткую стенку, то импеданс $Z_w = jc_{yn} / \omega - j\omega m$. Принимая во внимание потери в материале стенок, получаем запись импеданса в общем виде

$$Z_w = r + j \left(\frac{c_{yn}}{\omega} - \omega m \right). \quad (4.18)$$

Ткани боковых стенок речевого тракта имеют различные характеристики. Зубы, например, можно считать абсолютно жесткой поверхностью; язык, губы, щеки и мягкое нёбо более или менее податливы, а твердое нёбо имеет костную основу, на которой находится сравнительно тонкий слой мягкой ткани, поэтому очень трудно рассчитать средний импеданс тканей стенок речевого тракта. Непосредственные измерения импеданса тканей человека демонстрируют большой разброс параметров в зависимости от места измерения, амплитуды колебаний, формы измерительных контактов, возраста человека и т. д. В [4.25, 4.34] исследовался мастоид — участок поверхности головы за ухом, аналогичный по своему строению твердому нёбу. Установлено, что до 3 кГц импеданс мастоида носит упругий характер, а затем меняется на инерционный. Активную и реактивную компоненты этого импеданса R и X можно аппроксимировать как

$$R = \frac{1,15 \cdot 10^6}{\omega} + 0,01 \omega; \quad X = \frac{4,08 \cdot 10^6}{\omega^2} - 0,0106 \omega. \quad (4.19)$$

В [4.33] было обнаружено, что безразмерный удельный импеданс мягких тканей в области живота имеет инерционный характер:

$$Z_w = 200 - j 0,02 \omega. \quad (4.20)$$

Если принять, что импеданс мягких тканей речевого тракта также имеет инерционный характер, то, записывая его как $Z_w = r - j\omega m$, получаем из (4.15)

$$c_0^2 \left[1 - \frac{2 \bar{c}_0 m}{a(r^2 + \omega^2 m^2)} + j \frac{2 \bar{c}_0 r}{a \omega(r^2 + \omega^2 m^2)} \right] = \bar{c}_0^2. \quad (4.21)$$

Переписав (4.21) в виде $c_0^2 = (U_w - jV_w) \bar{c}_0^2$, воспользуемся биномом Ньютона и для малых поправок к скорости звука получим

$$c_0 \approx \bar{c}_0 \left(\sqrt{U_w} - j \frac{V_w}{2 \sqrt{U_w}} \right), \quad (4.22)$$

где U_w и V_w — активная и реактивная компоненты комплексной величины, обратной содержанию квадратных скобок в (4.21). Поправки на скорость звука $\Delta c = (\bar{c}_0 - c_0)/\bar{c}_0$, рассчитанные в [4.17] для импедансов типов (4.19) и (4.20), показаны на рис. 4.1. В этих расчетах сечение однородной трубы считалось прямоугольным с наибольшей стороной, равной 5 см. Как видно из этого рисунка, импеданс типа (4.19) в полосе частот до 3 кГц понижает

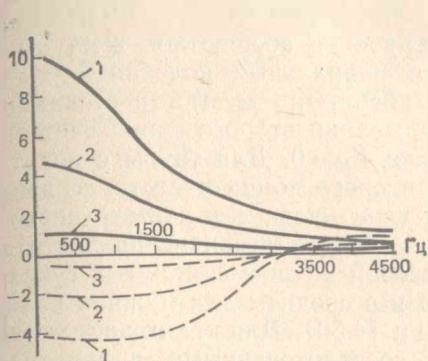


Рис. 4.1. Поправки на скорость звука в однородной трубе с прямоугольным сечением:

— — инерционный импеданс стенок; — — упруго-инерционный импеданс.

Площадь трубы: 1 — 1 см²; 2 — 2 см²; 3 — 4 см²

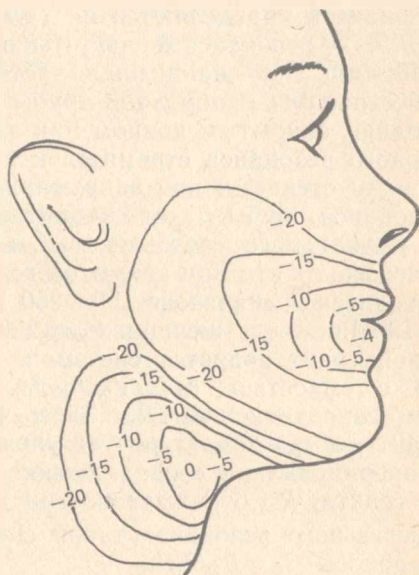


Рис. 4.2. Излучение через стенки речевого тракта (изолинии в дБ)

скорость звука, а выше 3 кГц — повышает ее. Импеданс инерционного типа повышает скорость звука и соответственно повышает резонансные частоты трубы. Для импеданса любого типа влияние податливости стенок увеличивается при уменьшении частоты и площади поперечного сечения трубы. При неограниченном росте частоты $\omega \rightarrow \infty$ скорость звука в трубе с податливыми стенками стремится к скорости звука в трубе с абсолютно жесткими стенками $c_0 \rightarrow \bar{c}_0$, т. е. влияние податливости падает с ростом частоты.

Импеданс (4.20) одно время широко использовался в расчетах податливости стенок речевого тракта, но затем было признано, что он неточно описывает характеристики тканей речевого тракта. В [4.39] были проведены непосредственные измерения импеданса тканей шеи и щек, которые для расслабленных щек дали: $r = 800$ г/с·см²; $m = 2,1$ г/см²; $c_{уп} = 8,45 \cdot 10^4$ дн/см³ (частота механического резонанса $F_w = 32$ Гц); для напряженных щек: $r = 1060$ г/с·см²; $m = 1,5$ г/см²; $c_{уп} = 3,33 \cdot 10^4$ дн/см³ (частота меха-

нического резонанса $F_w=60$ Гц), а для шеи: $r=2320$ г/с·см²; $m=2,1$ г/см²; $c_{yp}=4,91 \cdot 10^5$ дн/см³ (частота механического резонанса $F_w=72$ Гц).

Эти данные нельзя непосредственно применять к расчету радиальных колебаний в речевом тракте, поскольку они приводят к слишком низким резонансным частотам по сравнению с экспериментально измеренными. Дело в том, что при наличии радиальных колебаний волновые числа, а следовательно, и частоты резонансов определяются по (4.7), или $F_i^2 = F_{i0}^2 + F_w^2$, $i=1, 2, \dots$, где F_{i0} — резонансные частоты в трубе с абсолютно жесткими стенками; F_w — наинизшая частота радиальных колебаний. Как известно, для однородной трубы с абсолютно жесткими стенками и одним закрытым концом при закрывании второго конца частота первого резонанса стремится к нулю: $F_{i0} \rightarrow 0$. Для трубы с податливыми стенками при закрывании второго конца частота первого резонанса трубы будет стремиться не к нулю, а к частоте первого радиального резонанса: $F_i \rightarrow F_w$. Первая резонансная частота речевого тракта при закрытой голосовой щели и сомкнутых губах, находится в диапазоне 150—250 Гц и в среднем для мужчин $F_w \approx 190$ Гц, а для женщин $F_w \approx 220$ Гц [4.30]. Для таких частот речевой тракт является системой с сосредоточенными параметрами с емкостью $c_{p,t} = V_{p,t} / \rho_0 c^2$, индуктивностью L_w и активным сопротивлением R_w . Здесь $V_{p,t}$ — объем воздуха в речевом тракте, а $c_{p,t}$ характеризует упругость воздуха в этом объеме, L_w пропорциональна сосредоточенной удельной массе стенок речевого тракта, R_w отражает потери на деформацию стенок. Частота радиального резонанса такой системы $F_u = 1/2\pi \sqrt{L_w c_{p,t}}$, а ширина полосы $B_w = R_w / 2\pi L_w$.

Эксперименты с возбуждением колебаний в речевом тракте с помощью внешнего генератора позволили установить существование двух областей наиболее интенсивных колебаний стенок речевого тракта — около губ и в нижней части тракта (рис. 4.2 по [4.30]). Таким образом, эквивалентная схема речевого тракта была более точно представлена в виде двух параллельных сосредоточенных RLC -элементов, и по данным измерений были вычислены параметры среднего акустического импеданса стенок: $L_w = 0,011$ г/см⁴ для мужчин и $0,014$ г/см⁴ для женщин [4.30]. Пересчитывая данные [4.39] в акустический импеданс, получаем среднюю величину индуктивности, равную $0,0135$ г/см⁴, что довольно близко к среднему значению L_w для мужчин и женщин. Активное сопротивление R_w по данным [4.30] в среднем равно $5,2$ г/с см⁴ для мужчин и $8,3$ г/с см⁴ для женщин. Приводя данные [4.30, 4.39] к безразмерному импедансу, находим, что активное сопротивление находится в диапазоне $r \approx 30—80$, а реактивный член $m \approx 0,06—0,09$ с.

Изменение плотности воздуха согласно (4.15) изменяет скорость звука, что было использовано в [4.6] в экспериментах с речью водолазов, которые дышали кислородно-гелиевой смесью. На 150

основании этих экспериментов было найдено, что импеданс стенок речевого тракта в среднем носит инерционный характер: $Z_w = 15000 - j16\omega$ кг/м²·с. Рассчитанное по этим данным среднее значение частоты радиального резонанса в речевом тракте $F_w = 230$ Гц, причем для разных гласных звуков русской речи диапазон изменений F_w составлял 166—390 Гц. Результаты измерений близки данным [4.30, 4.39] для реактивного члена безразмерного импеданса стенок, величина которого $m = 0,04$ с, тогда как активный член $r = 400$ заметно отличается, что может быть связано с повышенным давлением, использовавшимся в этих экспериментах.

Влияние податливости стенок на акустические характеристики трубы не ограничивается смещением резонансных частот и увеличением потерь. Как видно из (4.16), в такой трубе скорость звука для разных частот различна, т. е. возникает дисперсия колебаний с разными частотами. Для стенок инерционного типа скорость звука тем больше, чем ниже частота колебаний, поэтому низкочастотные колебания опережают высокочастотные. Это опережение невелико, например, при относительной разнице в скорости распространения двух частот в 1% разница во времени прибытия фронта более быстрой волны при длине трубы в 17,5 см составит 5 мкс, а при разнице в 10% — 50 мкс.

Однако фазовая скорость распространения колебаний с уменьшением частоты сигнала сначала увеличивается, а затем начинает уменьшаться и при некоторой критической частоте $f_{кр}$ становится равной нулю, а при $f < f_{кр}$ становится мнимой. Для импеданса стенок инерционного типа

$$f_{кр} = \frac{1}{2\pi m} \sqrt{\frac{2m\bar{c}_0}{a} - r^2}. \quad (4.23)$$

Таким образом, при $f < f_{кр}$ нет бегущих волн, а существуют лишь неоднородные. Зависимость $f_{кр}(a)$ для разных параметров m и r показана на рис. 4.3, где область «запирания» расположена ниже и левее каждой кривой. Как видно из этого рисунка, активные потери существенно влияют на критическую величину параметра a , выше которой пропускаются колебания любых частот. Кривые на рис. 4.3 рассчитаны для однородной трубы с податливыми стенками, где a — либо радиус круглого сечения, либо наименьшая сторона прямоугольника. Если сужение трубы происходит локально, как это часто бывает в речевом тракте, то явление «запирания» также должно возникать, хотя, возможно, характер зависимости $f_{кр}(a)$ будет другим. Физический смысл явления «запирания» состоит в том, что стенки трубы начинают колебаться в противофазе с акустической волной, препятствуя ее распространению, по-

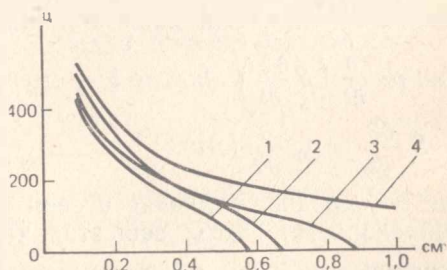


Рис. 4.3. Частота запирания как функция характерного геометрического размера сечения и параметров импеданса стенок:

1 — $m = 0,08$, $r = 100$; 2 — $m = 0,06$, $r = 80$; 3 — $m = 0,08$, $r = 80$; 4 — $m = 0,06$, $r = 50$.

этому и локальное сужение также при известных условиях должно приводить к прекращению распространения волн определенной частоты.

Импеданс стенок речевого тракта неоднороден, поэтому зависимость (4.23) может непосредственно характеризовать условия «запирания» и в голосовой щели, и при сужении в области мягкого нёба или губ, т. е. там, где импеданс стенок — инерционного типа. В области твердого нёба, импеданс которого носит упругий характер в полосе частот до 3 кГц, лишь ткани языка имеют импеданс инерционного типа, поэтому при образовании узких щелей в этой области влияние податливости стенок на распространение волн носит более сложный характер. Этот вопрос требует дальнейшего исследования ввиду важности частотных характеристик переднеязычных щелевых и взрывных звуков.

Как мы видели, податливость стенок речевого тракта приводит к появлению ряда эффектов, особенно заметных на низких частотах. Поэтому в общем случае уравнение речевого тракта зависит от двух пространственных переменных и должно представляться в виде уравнения (4.2). В качестве некоторого приближения можно пользоваться одномерным уравнением, а для расчета поправок на собственные частоты и затухания, а также определения критической частоты «запирания» — выражениями (4.22) и (4.23).

В речевом тракте площадь поперечного сечения S является функцией пространственной координаты x вдоль его оси и времени t . Эта зависимость приводит к изменению формы волнового уравнения. Выведем уравнение речевого тракта, не принимая пока во внимание различные виды потерь. Возьмем элемент газа, заключенный между плоскостями с координатами x и $x + \Delta x$. Если в речевом тракте существует звуковое давление $P(x)$, то на левую плоскость в момент t действует сила PS , а на правую — сила $PS + \frac{\partial(PS)}{\partial x} dx$. Осевая компонента реакции стенок есть $P_{\text{ст}} \frac{\partial S}{\partial x} dx$, где $P_{\text{ст}}$ — давление у стенок. Результирующая сила, действующая на элемент газа,

$$PS + P_{\text{ст}} \frac{\partial S}{\partial x} dx - PS - \frac{\partial(PS)}{\partial x} dx = \left[(P_{\text{ст}} - P) \frac{\partial S}{\partial x} - S \frac{\partial P}{\partial x} \right] dx.$$

Считая, что $P_{\text{ст}}$ существенно не отличается от среднего давления в сечении, т. е. фактически допуская, что в речевом тракте распространяются лишь плоские волны, получаем результирующую силу $S \frac{\partial P}{\partial x} dx$. Эта сила должна уравниваться силой инерции $\rho_0 \frac{\partial}{\partial t} \left(S \frac{\partial \xi}{\partial t} \right) dx$, где ξ — смещение. Отсюда получаем

$$-S \frac{\partial P}{\partial x} = \rho_0 \frac{\partial}{\partial t} (S v), \quad (4.24)$$

где v — среднее значение осевой компоненты вектора колебательной скорости. Поток вещества через левую границу выбранного сечения есть $\rho_0 v S$, а через правую — $\rho_0 v S + \frac{\partial}{\partial x} (\rho_0 v S) dx$. Разност-

ный поток должен быть равен скорости изменения массы элемента газа $\frac{S}{c_0^2} \frac{\partial P}{\partial t} dx$, откуда получаем второе уравнение, связывающее значения давления и скорости:

$$-S \frac{\partial P}{\partial t} = \rho_0 c_0^2 \frac{\partial (Sv)}{\partial x}. \quad (4.25)$$

Воспользовавшись определением потенциала скорости $v = -\partial\Phi/\partial x$, $P = \rho_0 \partial\Phi/\partial t$, из (4.25) получим

$$\frac{1}{S} \frac{\partial}{\partial x} \left(S \frac{\partial \Phi}{\partial x} \right) = \frac{1}{c_0^2} \frac{\partial^2 \Phi}{\partial t^2}. \quad (4.26)$$

Это уравнение известно как уравнение Вебстера, или уравнение рупора. Перепишем (4.24) и (4.25) в виде:

$$-\frac{\partial (Sv)}{\partial t} = \frac{S}{\rho_0} \frac{\partial P}{\partial x}; \quad -\frac{\partial (Sv)}{\partial x} = \frac{S}{\rho_0 c_0^2} \frac{\partial P}{\partial t}. \quad (4.27); (4.28)$$

Возьмем от (4.27) производную по x , а от (4.28) — производную по t и приравняем правые части. Получаем уравнение речевого тракта относительно звукового давления

$$\frac{\partial}{\partial x} \left(S \frac{\partial P}{\partial x} \right) = \frac{1}{c_0^2} \frac{\partial}{\partial t} \left(S \frac{\partial P}{\partial t} \right). \quad (4.29)$$

Если площадь поперечного сечения S не изменяется во времени, то $\partial S/\partial t = 0$ и (4.29) переходит в уравнение Вебстера:

$$\frac{1}{S} \frac{\partial}{\partial x} \left(S \frac{\partial P}{\partial x} \right) = \frac{1}{c_0^2} \frac{\partial^2 P}{\partial t^2}. \quad (4.30)$$

Аналогично из системы

$$-\frac{\partial P}{\partial x} = \frac{\rho_0}{S} \frac{\partial (Sv)}{\partial t}; \quad (4.31)$$

$$-\frac{\partial P}{\partial t} = \frac{\rho_0 c_0^2}{S} \frac{\partial (Sv)}{\partial x}, \quad (4.32)$$

взяв производную по t в (4.31) и производную по x в (4.32), получим уравнение относительно колебательной скорости

$$\frac{\partial}{\partial x} \left[\frac{1}{S} \frac{\partial (Sv)}{\partial x} \right] = \frac{1}{c_0^2} \frac{\partial}{\partial t} \left[\frac{1}{S} \frac{\partial (Sv)}{\partial t} \right]. \quad (4.33)$$

При неизменной во времени площади поперечного сечения S имеем

$$\frac{\partial}{\partial x} \left[\frac{1}{S} \frac{\partial (Sv)}{\partial x} \right] = \frac{1}{c_0^2} \frac{\partial^2 v}{\partial t^2}. \quad (4.34)$$

Для вывода уравнения относительно смещения частиц рассмотрим изменение объема элемента газа. При прохождении акусти-

ческой волны левая граница элемента газа сместится на $\xi(x)$, а правая — на $\xi(x) + \frac{\partial \xi}{\partial x} dx$. Объем газа в деформированном элементе есть $S dx + \frac{\partial(S\xi)}{\partial x} dx$, а его относительное изменение равно $\frac{1}{S} \frac{\partial(S\xi)}{\partial x}$. Тогда изменение давления в этом элементе по [4.13]

$$P = - \frac{\rho_0 c_0^2}{S} \frac{\partial(S\xi)}{\partial x}. \quad (4.35)$$

Возьмем в (4.35) производную по x , умножим обе части равенства на S и приравняем его правую часть правой части (4.24):

$$S \frac{\partial}{\partial x} \left[\frac{1}{S} \frac{\partial(S\xi)}{\partial x} \right] = \frac{1}{c_0^2} \frac{\partial}{\partial t} \left(S \frac{\partial \xi}{\partial t} \right). \quad (4.36)$$

При $\partial S / \partial t = 0$ имеем

$$\frac{\partial}{\partial x} \left[\frac{1}{S} \frac{\partial(S\xi)}{\partial x} \right] = \frac{1}{c_0^2} \frac{\partial^2 \xi}{\partial t^2}. \quad (4.37)$$

При анализе процессов речеобразования иногда удобно использовать разные представления уравнения речевого тракта, выбирая их из (4.26), (4.29), (4.33) и (4.36). Уравнение (4.26) во многих отношениях удобнее остальных. Здесь форма уравнения Вестера получается без предположения о неизменной площади поперечного сечения S во времени. Кроме того, как будет видно из дальнейшего, основное преимущество (4.26) состоит в удобстве записи граничных условий, в которые часто одновременно входят и давление, и скорость. Решив (4.26), можно получить звуковое давление и скорость, продифференцировав потенциал скорости Φ по t и x . Для однородной трубы с сечением $S = \text{const}$ все уравнения (4.26) — (4.37) приобретают вид волнового уравнения

$$\frac{\partial^2 \Phi}{\partial x^2} = \frac{1}{c_0^2} \frac{\partial^2 \Phi}{\partial t^2}, \quad (4.38)$$

где Φ — любая переменная — потенциал скорости, давление, смещение или скорость.

Существуют определенные основания для того, чтобы в ряде случаев принимать $\partial S / \partial t \approx 0$, поскольку скорости движения артикуляторных органов, как мы видели в гл. 3, значительно меньше скорости акустических колебаний, и за период колебаний даже первого (низшего) резонанса площадь сечения S обычно изменяется незначительно. Это обстоятельство позволяет применить к (4.26) метод разделения переменных, что существенно облегчает процесс его анализа и решения. При таком предположении очевидно, что уравнения (4.26) и (4.30) полностью эквивалентны. Необходимо отметить, что встречаются артикуляционные ситуации, при которых скорость изменения площади поперечного сечения S больше скорости движения артикуляторных органов. Это относится, например, к движениям средней части языка, открывающей или закрывающей щель между зубами, подключая или отключая дополнительные объемы, находящиеся между

зубами и щеками. Когда язык упирается своими краями в свод твердого неба, то при его дальнейшем подъеме площадь поперечного сечения в этой области также изменяется быстрее, чем скорость движения языка, вследствие изогнутой формы свода твердого неба. При некоторых способах решения уравнения речевого тракта ненулевая скорость изменения площади поперечного сечения может привести к возникновению искажений в рассчитанной речевой волне.

Уравнение речевого тракта составляет краевую задачу, и для ее решения необходимо задать начальные и граничные условия. Условия излучения волны со стороны губ и свойства голосовой щели описываются ниже, а сейчас остановимся на методах решения уравнения речевого тракта. Если акустическая система велика по сравнению с длиной волны, так что в ней доминируют неустановившиеся процессы, то предпочтительно решение в виде бегущих волн. Если же система мала по сравнению с длиной волны, то простейшие результаты дает решение в виде стоячих волн [4.14]. В первом случае обычно используют метод Даламбера, во втором — метод Фурье.

Известны аналитические решения лишь для небольшого числа функций S , входящих в уравнение Вебстера. В простейшем случае, когда $S = \text{const}$, имеем одномерное волновое уравнение типа (4.38). Как нетрудно проверить непосредственной подстановкой, решение (4.38) может быть представлено в виде волн, распространяющихся по трубе в противоположных направлениях: $\varphi(x, t) = f_1(x - c_0 t) + f_2(x + c_0 t)$. Для бесконечной трубы функции f_1 и f_2 определяются начальными условиями $\varphi(x, 0) = \varphi_0(x)$, $\partial\varphi(x, 0)/\partial t = \varphi_1(x)$ и решение есть

$$\varphi(x, t) = \frac{\varphi_0(x - c_0 t) + \varphi_0(x + c_0 t)}{2} + \frac{1}{2c_0} \int_{x - c_0 t}^{x + c_0 t} \varphi_1(x) dx. \quad (4.39)$$

Это решение в виде бегущих волн называется решением Даламбера. Для трубы ограниченной длины l с заданными граничными условиями ищут такое продолжение функций f_1 и f_2 , которое удовлетворяло бы граничным условиям в точках 0 и l , и определяют такое начальное возмущение бесконечной системы, чтобы движение на интервале $(0, l)$ удовлетворяло граничным условиям, а остальные участки можно было бы отбросить [4.16].

Для речевого тракта, однако, более удобно решение в виде стоячих волн, которое получается с помощью метода разделения переменных $\varphi(x, t) = X(x)T(t)$. Для однородной трубы, описываемой уравнением (4.38), применение этого метода дает систему:

$$X'' + \lambda^2 X = 0; \quad (4.40)$$

$$T'' + c_0^2 \lambda^2 T = 0, \quad (4.41)$$

где λ — собственные числа системы. В (4.40) производные берутся по x , а в (4.41) — по t . Общие решения для (4.40) и (4.41) есть:

$$X(x) = A \cos \lambda x + B \sin \lambda x; \quad (4.42)$$

$$T(t) = C \cos c_0 \lambda t + D \sin c_0 \lambda t. \quad (4.43)$$

Постоянные A и B в (4.42) определяются из граничных условий.

Возьмем в качестве переменной потенциал скорости. Тогда если на конце труба закрыта ($x=0$), то скорость частиц на задней стенке должна быть равна нулю и $\partial\Phi/\partial t=0$. Если на конце труба открыта ($x=l$), то акустическое давление равно нулю, т. е. $\partial\Phi/\partial x=0$. Это соответствует граничным условиям $X'(0)=X(l)=0$, подставляя которые в (4.42), находим $B=0$ и $\cos \lambda l=0$, что дает бесконечное множество собственных чисел $\lambda_k=(2k-1)\pi/2l$, $k=1, 2, \dots$. Пространственные моды такой системы есть $X_k(x)=A_k \cos \lambda_k x$, и решение (4.38) представляется как

$$\Phi(x, t) = \sum_{k=1}^{\infty} A_k \cos \lambda_k x (C_k \cos \lambda_k c_0 t + D_k \sin \lambda_k c_0 t),$$

где постоянные C_k и D_k находятся из начальных условий:

$$C_k = \frac{2}{l} \int_0^l \Phi_0(x) X_k(x) dx; \quad (4.44)$$

$$D_k = \frac{2}{l} \int_0^l \Phi_1(x) X_k(x) dx. \quad (4.45)$$

Если второй конец трубы также закрыт, то $X_k(x)=\cos \lambda_k x$, $\lambda_k=k\pi/l$, $k=1, 2, \dots$. Аналогично можно найти решения для трубы с обоими открытыми концами.

Если площадь поперечного сечения трубы изменяется по экспоненциальному закону $S(x)=S_0 e^{\alpha x}$, то уравнение Вебстера принимает вид

$$\frac{\partial^2 \Phi}{\partial x^2} + \alpha \frac{\partial \Phi}{\partial x} = \frac{1}{c_0^2} \frac{\partial^2 \Phi}{\partial t^2}. \quad (4.46)$$

Разделяя переменные в (4.46), получаем систему:

$$\begin{aligned} X'' + \alpha X' + \lambda^2 X &= 0; \\ T'' + \lambda^2 c_0^2 T &= 0. \end{aligned} \quad (4.47)$$

Общее решение (4.47) есть

$$X(x) = e^{-\frac{\alpha}{2}x} \left(A \cos \sqrt{\lambda^2 - \frac{\alpha^2}{4}} x + B \sin \sqrt{\lambda^2 - \frac{\alpha^2}{4}} x \right),$$

и в зависимости от граничных условий можно найти собственные числа λ_k и собственные функции $X_k(x)$ этого уравнения. Замкнутые решения получаются также и для семейства функций $S(x)$, описывающихся уравнением $\partial S/\partial x = a_n S^n$, которое при $n \neq 1$ имеет решение [4.20]:

$$S(x) = S_0 \left(1 + \beta \frac{1}{1-n} x \right)^{\frac{1}{1-n}},$$

где S_0 — начальное значение функции $S(0)$, а $\beta \frac{1}{1-n} = (1-n) a_n / S_0^{1-n}$.

В интервале дробных значений $0 \leq n \leq 1$ при $n = (m-1)/m$, $m=1, 2, \dots$, получаем семейство так называемых степенных рупоров с площадью поперечного сечения, изменяющейся по закону

$$S(x) = S_0(1 + \beta_m x)^m. \quad (4.48)$$

Подставляя (4.48) в уравнение Вебстера и разделяя переменные, получаем:

$$X'' + \frac{m}{x + \frac{1}{\beta_m}} X' + \lambda^2 X = 0; \quad (4.49)$$

$$T'' + \lambda^2 c_0^2 T = 0.$$

Общее решение уравнения (4.49), которое является разновидностью уравнения Бесселя, следующее:

$$X(x) = \frac{1}{\left(x + \frac{1}{\beta_m}\right)^{\frac{m-1}{2}}} H_{\frac{m-1}{2}} \left[\lambda \left(x + \frac{1}{\beta_m}\right) \right],$$

где $H_{\frac{m-1}{2}}$ — цилиндрическая функция Ганкеля $(m-1)/2$ -го порядка. Для четных значений m эта функция выражается через экспоненциальные функции, и в практически важном случае конического рупора, площадь поперечного сечения которого $S(x) = S_0(1 + \beta_2 x)^2$, решение (4.49) представляется как

$$X(x) = \sqrt{\frac{2}{\pi}} \frac{A \cos[\lambda(x + \gamma)] + B \sin[\lambda(x + \gamma)]}{(x + \gamma) \sqrt{\lambda}}, \quad (4.50)$$

где $\gamma = 1/\beta_2$; β_2 — коэффициент раскрытия конического рупора. По сравнению с решением для однородной трубы (4.42) решение (4.50) оказывается ненамного сложнее, однако использование конических секций для аппроксимации формы речевого тракта значительно выгоднее, чем использование цилиндрических секций. Поскольку функция площади поперечного сечения речевого тракта такова, что для нее невозможно аналитически решить уравнение речевого тракта, в ряде случаев прибегают к аппроксимации речевого тракта в виде последовательного соединения цилиндрических секций. Известно, что линейная интерполяция произвольной функции гарантирует значительно меньшую ошибку, чем ступенчатая, тогда как повышение порядка интерполяции, как правило, снижает ошибку интерполяции очень медленно, но повышает ее сложность [4.12]. Это означает, что использование конических секций вместо цилиндрических для аппроксимации формы речевого тракта существенно повышает точность либо при заданной точности уменьшает число требуемых секций. Последнее весьма важно для понижения сложности алгоритмов расчета акустических характеристик речевого тракта. Повышение порядка интерполяции за счет использования степенных рупоров приводит к ус-

ложению вычислительных процедур, которое, по-видимому, не окупается повышением точности аппроксимации.

Аналитические решения обладают рядом преимуществ, поэтому был разработан ряд методов для приближенного решения уравнения речевого тракта. Например, в [4.57] было предложено использовать метод малых возмущений, состоящий в поиске поправок на собственные числа и функции, если реальная система незначительно отличается от системы, для которой решение уже известно [4.10]. Собственные числа и функции решаемого уравнения представляются рядами:

$$\lambda_k = \lambda_k^{(0)} + \sum_{m=1}^{\infty} \varepsilon^m \lambda_k^{(m)}; \quad \varphi_k = \varphi_k^{(0)} + \sum_{m=1}^{\infty} \varepsilon^m \varphi_k^{(m)},$$

где $\lambda_k^{(0)}$ и $\varphi_k^{(0)}$ — собственные числа и функции исходной модели; m — порядок приближения; ε — малый параметр. Подставляя эти ряды в исходное уравнение, приравняем нулю члены при одинаковых степенях, умножим каждое уравнение на $\varphi_k^{(0)}$ и проинтегрируем по всей длине речевого тракта. Тогда получаем выражение для вычисления собственных чисел m -го приближения

$$\lambda_k^{(m)^2} = \int_0^l S_1(x) \varphi_k^{(m-1)'} \varphi_k^{(0)'} dx - \lambda_k^{(m-1)} \int_0^l S_1(x) \varphi_k^{(0)^2} dx - \\ - S_1(x) \varphi_k^{(m-1)'} \varphi_k^{(0)} \Big|_0^l,$$

где

$$S_1(x) = \frac{S(x)}{S_0(x)} - 1,$$

а $S_0(x)$ — площадь поперечного сечения исходной модели, для которой известны аналитическое решение и собственные числа и функции $\lambda_k^{(0)}$, $\varphi_k^{(0)}$. Если в качестве нулевого приближения для речевого тракта взять однородную трубу с площадью $S_0 = \text{const}$ и собственными числами $\lambda_k^{(0)} = \frac{(2k-1)\pi}{2l}$, $k = 1, 2, \dots$, то собственные числа в первом приближении

$$\lambda_k^{(1)} = \lambda_k^{(0)} \sqrt{1 - \frac{2}{l} \int_0^l \left[\frac{S(x)}{S_0} - 1 \right] \cos 2\lambda_k^{(0)} x dx - \left[\frac{S(x)}{S_0} - 1 \right] \cos 2\lambda_k^{(0)} l \sin \lambda_k^{(0)} l}.$$

Расчеты показывают, что путем надлежащего выбора S_0 в ряде случаев можно достичь довольно малой ошибки (рис. 4.4), однако были обнаружены и такие конфигурации речевого тракта, для которых погрешность оценки частоты некоторых резонансов остается сравнительно большой при любом значении S_0 в диапазоне от 0 до 10 см² (рис. 4.4 для гласных *И*, *Ы*). Этот метод не используется, так как при минимизации ошибки для каждого собственного числа требуются разные величины S_0 и пока не найден алгоритм их выбора. При постоянной величине $S_0 = 0,7 S_{\text{ср}}$, где $S_{\text{ср}}$ — средняя площадь речевого тракта, средняя ошибка (по сравнению с численным методом решения) составляет для пяти

русских гласных 11,5%, но максимальные ошибки доходят до $-25 - +39\%$. Это неудивительно, поскольку метод возмущений требует, чтобы максимальное отклонение функции $S(x)$ от S_0 не превышало 10%, а это условие соблюдается весьма редко.

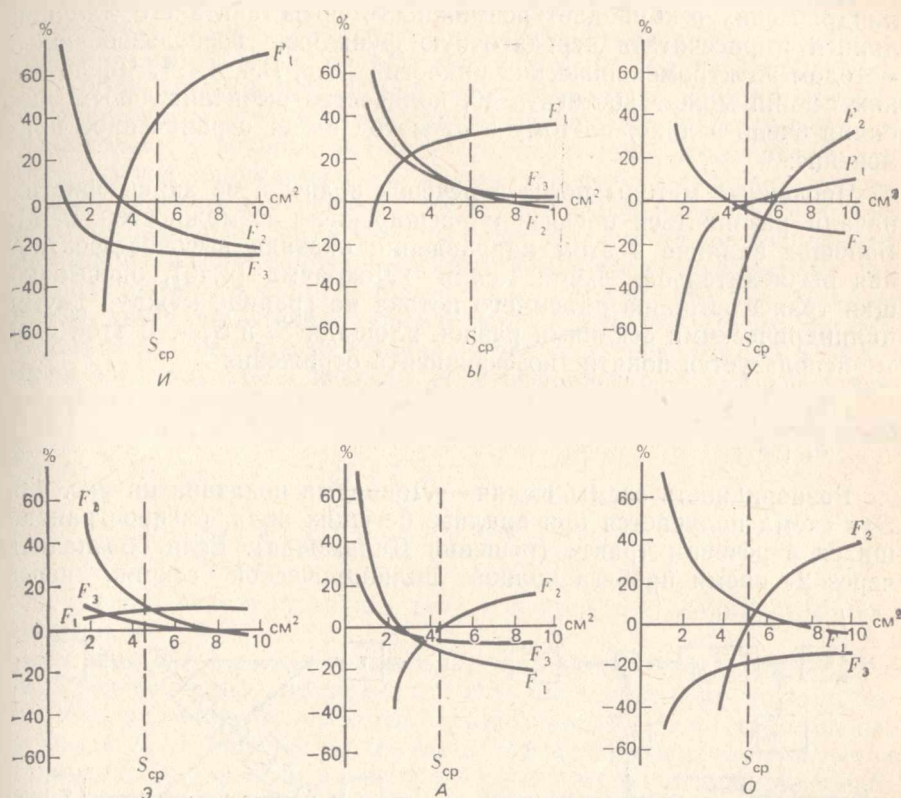


Рис. 4.4. Зависимость ошибок вычисления собственных частот речевого тракта методом возмущения от площади

Более удачная модификация метода возмущений была предложена в [4.36, 4.46, 4.55] на основе разложения логарифма площади $\ln S(x)$ в ряд Фурье:

$$\ln S(x) = \ln S_0 + \sum_{j=1}^{\infty} a_j \cos j \frac{\pi x}{l}. \quad (4.51)$$

Логарифм площади в (4.26) появляется, если представить это уравнение как

$$\frac{\partial^2 \Phi}{\partial x^2} + \frac{1}{S} \frac{\partial S}{\partial x} \frac{\partial \Phi}{\partial x} = \frac{1}{c_0^2} \frac{\partial^2 \Phi}{\partial t^2}$$

и записать второй член в виде логарифмической производной

$$\frac{1}{S} \frac{\partial S}{\partial x} = \frac{\partial (\ln S)}{\partial x}.$$

Коэффициенты разложения a_i оказываются в явном виде связанными с собственными числами уравнения Вебстера, однако в этом методе требуется использовать большое число членов разложения (4.51).

Аппроксимация речевого тракта в виде последовательности цилиндрических секций дает возможность представить его длинной линией и рассчитать передаточную функцию, воспользовавшись методом электромеханических аналогий [В.3, В.4, 4.42]. Число таких секций может достигать 40, количество вычислительных операций очень велико, поэтому этот метод имеет ограниченное применение.

Численные методы расчета речевой волны и ее характеристик начали развиваться в связи с расширением возможностей ЭВМ. Большое влияние в этом направлении оказала высокоэффективная вычислительная схема Келли — Локбаума [4.40], описывающая трансформацию объемного потока на границе между двумя цилиндрическими секциями разной площади S_i и S_{i+1} . В этой схеме используется понятие коэффициента отражения

$$\kappa_{от} = \frac{S_i - S_{i+1}}{S_i + S_{i+1}}.$$

Разновидность схемы Келли — Локбаума показана на рис. 4.5. Эта схема получается при анализе бегущих волн, распространяющихся в речевом тракте (решения Даламбера). Если обозначить через 2τ время пробега волной цилиндрической секции, через

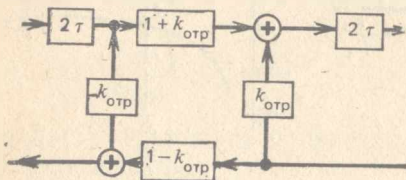


Рис. 4.5. Схема Келли — Локбаума

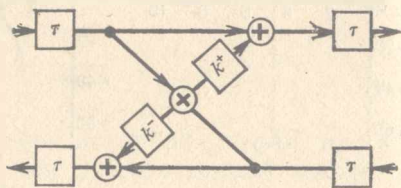


Рис. 4.6. Рекурсивная схема, учитывающая изменения сечения речевого тракта во времени

$u^+(t-\tau)$ волну, проходящую через правую границу секции в положительном направлении (от начала к концу речевого тракта), через $u^-(t+\tau)$ волну, проходящую через правую границу секции в обратном направлении, и аналогично через $u^+(t+\tau)$ и $u^-(t-\tau)$ волны, проходящие через левую границу секции, то, решая (4.24) и (4.25) с помощью решения Даламбера, получаем рекурсивную схему:

$$u_{m-1}^+(t+\tau) = \kappa_{от} u_{m-1}^-(t-\tau) + (1 + \kappa_{от}) u_m^+(t-\tau); \quad (4.52)$$

$$u_m^-(t+\tau) = (1 - \kappa_{от}) u_{m-1}^-(t-\tau) - \kappa_{от} u_m^+(t-\tau),$$

где m — порядковый номер секции; u — объемная скорость акустического потока [4.11]. Эта схема показана на рис. 4.5, но существует целый ряд эквивалентных ей схем.

В дальнейшем была установлена связь между коэффициентами отражения в последовательности цилиндрических секций и коэффициентами линейного предсказания речевой волны, и схема Келли — Локбаума получила широкое применение в синтезаторах. Однако синтез речи по этой схеме сопровождается характерными искажениями, причина которых состоит в том, что за время пробега волной цилиндрической секции поперечное сечение соседних секций немного изменяется и их граничные условия оказываются несогласованными. Несмотря на относительно медленное изменение площади поперечного сечения речевого тракта во времени, вычислительные особенности схемы (4.52) таковы, что это приводит к заметному искажению речевой волны. Для учета этих искажений более сложная рекурсивная схема была предложена в [4.44] (рис. 4.6). Эта схема использует коэффициенты отражения, зависящие от времени:

$$\kappa_{m-1, n}^+ = \left(\frac{S_{m-1, n-1}}{S_{m, n}} - 1 \right) / \left(\frac{S_{m-1, n-1}}{S_{m, n-1}} - 1 \right);$$

$$\kappa_{m-1, n}^- = \left(1 - \frac{S_{m, n-1}}{S_{m-1, n}} \right) / \left(1 + \frac{S_{m, n-1}}{S_{m-1, n-1}} \right),$$

где первый индекс означает номер секции, а второй — момент времени.

Рекурсивные схемы типа (4.52) дают результат вычислений в виде речевой волны, но во многих случаях необходимо сначала вычислить резонансные частоты и затухания, а также собственные функции. С этой целью может быть эффективно использован метод Галеркина, который, в отличие от других методов приближенного решения, не накладывает никаких ограничений на коэффициенты дифференциального уравнения [4.8]. В основе метода Галеркина лежит разложение искомых собственных функций $\psi_k(x)$ краевой задачи в ряд по собственным функциям $\psi_i^{(0)}(x)$ некоторого известного решения:

$$\psi_k(x) = \sum_{i=1}^k a_i \psi_i^{(0)}(x). \quad (4.53)$$

Разделяя переменные в уравнении (4.26), получаем систему:

$$(S X')' + \lambda^2 S X = 0;$$

$$T'' + \lambda^2 c_0^2 T = 0. \quad (4.54)$$

Подставим (4.53) в (4.54) и потребуем ортогональности (4.54) к каждой из функций нулевого приближения $\psi_j^{(0)}(x)$, т. е. выполнения для каждой k -й искомой собственной функции $\psi_k(x)$ условия

$$\int_0^l [(S \psi_k')' + \lambda_k^2 S \psi_k] \psi_j^{(0)} dx = 0, \quad (4.55)$$

где l — длина речевого тракта. Таким образом, для k собственных функций имеем систему из k уравнений. Чтобы избежать дифференцирования по x функции площади поперечного сечения $S(x)$, возьмем первый интеграл в (4.55) по частям и обозначим:

$$A_{ij} = S \psi_i^{(0)'} \psi_j^{(0)'} \Big|_0^l - \int_0^l S \psi_i^{(0)'} \psi_j^{(0)'} dx; \quad B_{ij} = \int_0^l S \psi_i^{(0)} \psi_j^{(0)} dx.$$

В результате получим систему уравнений, линейную относительно коэффициентов a_i :

$$\sum_{j=1}^k (A_{ij} + \lambda^2 B_{ij}) a_j = 0, \quad i = 1, 2, \dots, k. \quad (4.56)$$

Как известно, система уравнений (4.56) разрешима только в том случае, когда ее определитель равен нулю. Поскольку вместе с λ_i в (4.56) содержится $2k$ неизвестных, присоединим недостающее уравнение, получаемое из условия нормирования собственных функций

$$\int_0^l \psi_k^2(x) dx = 1. \quad (4.57)$$

Тогда система (4.56), (4.57) позволяет найти собственные числа λ_k и собственные функции ψ_k . Порядок приближения k ограничивается снизу числом искоемых собственных функций. Например, если желательно рассчитать параметры пяти первых мод, то и порядок приближения не может быть ниже пяти. Как будет видно из дальнейшего, для назальных и щелевых звуков речи в полосе частот до 5 кГц размещается до восьми резонансов, тогда как для гласных звуков в этой полосе обычно находится пять. Ограничение сверху на порядок приближения может быть найдено из условия ограничения отклонения от единичной диагональной матрицы, составленной из скалярных произведений

$$s_{ij} = \int_0^l \psi_i(x) \psi_j(x) dx,$$

т. е. меры неортогональности множества функций $\{\psi_k\}$. Для назальных звуков, например, уже при $k=10$ максимальное значение s_{ij} лежит в диапазоне 0,04—0,06, что вполне удовлетворительно. Таким образом, при минимально потребном значении приближения $k=8$ фактически необходимо использовать лишь немного большее число собственных функций нулевого приближения $k=10$ [4.1]. Для гласных звуков при $k=5$ условие ортогонализации приводит к $k=9$. В последнем случае имеется возможность сократить объем вычислений и использовать минимально требуемое $k=5$ за счет применения корректировки вычисленных собственных значений [4.2]. Такая возможность появляется благодаря тому, что наложение любых ограничений на систему, например, в виде усечения ряда (4.53) приводит к повышению вычисленных собственных частот системы. Эмпирически подбирая коррекцию для вычисленных λ_k ,

удается снизить порядок приближения почти в 2 раза — с 9 до 5. При этом оказывается, что порядок приближения очень слабо влияет на вычисленные значения коэффициентов затухания резонансных колебаний.

Итак, для речевого тракта число решаемых уравнений методом Галеркина равно 5—10, что эквивалентно минимальному числу уравнений, решаемых при аппроксимации речевого тракта последовательностью цилиндрических секций. При этом точность решения значительно выше. Для достижения требуемой точности при аппроксимации речевого тракта цилиндрическими секциями необходимо использовать до 30—40 секций и соответственно решать до 30—40 уравнений, поэтому метод Галеркина выглядит весьма привлекательно. К тому же, при использовании этого метода предварительно не нужно вычислять параметры емкости и индуктивности секций, как это делается при использовании метода длинных линий.

Собственные функции нулевого приближения $\psi_i^{(0)}(x)$ могут быть получены из решения краевой задачи для однородной трубы. Как будет видно из § 4.2, вследствие необходимости удовлетворения граничным условиям возникает задача выбора площади поперечного сечения однородной трубы S_0 , но практика показывает, что в методе Галеркина вариации S_0 в противоположность методу возмущений сравнительно мало влияют на результат вычислений.

Считается, что собственные числа λ_k медленно изменяются во времени, и это дает основание для решения (4.41) как уравнения с постоянными коэффициентами. Эксперименты с синтетической речью показали, что качество синтезированного сигнала оказывается удовлетворительным в том случае, если собственные числа остаются постоянными в течение примерно 5 мс, а затем скачком обновляются. В некоторых случаях скорость изменения площади поперечного сечения трубы может быть достаточно большой. Иногда скорость изменения собственных чисел λ_k может быть значительно больше скорости изменения $S(x, t)$ (см. § 4.2). Поэтому если аппроксимировать траектории собственных чисел во времени $\lambda_k(t)$ не ступенчатой функцией, а каким-либо временным законом, то при тех же перцептивных характеристиках речевого сигнала увеличатся интервалы времени, через которые необходимо обновлять параметры λ_k . Это приведет к уменьшению числа вычислительных операций, связанных с определением $\lambda_k(t)$ и $\psi_k(x, t)$ по $S(x, t)$ на единицу времени синтезированного речевого сигнала.

Ранее выяснилось преимущество линейной интерполяции относительно числа вычислительных операций при заданной точности аппроксимации. Обозначим $\omega_k^2 = \lambda_k^2 c_0^2$ и представим изменение $\omega_k^2(t)$ в интервале $t_{i-1} - t_i$ в виде линейного закона:

$$\omega_k^2(t) = a_k^{(i)} + b_k^{(i)} t,$$

где $a_k^{(i)}$ и $b_k^{(i)}$ в общем случае различны для разных моментов времени t_i . Заменим аргумент $\bar{t} = a_k + b_k t$, в результате чего вместо (4.41) получим

$$b_k^{(i)*} T_k'' + \bar{t} T_k = 0. \quad (4.58)$$

Это уравнение по [2.2] есть частный случай уравнения Бесселя, и его решение в интервале $t_{i-1} - t_i$ есть

$$T_k^{(i)} = \sqrt{a_k^{(i)} + b_k^{(i)} \bar{t}} Z_{\frac{1}{3}}(\bar{t}), \quad (4.59)$$

$$\text{где } \bar{t} = \frac{2 \sqrt{(a_k^{(i)} + b_k^{(i)})^3}}{3 b_k^{(i)}}; \quad Z_{\frac{1}{3}}(\bar{t}) = A J_{\frac{1}{3}}(\bar{t}) + B N_{\frac{1}{3}}(\bar{t});$$

$J_{\frac{1}{3}}$ — функция Бесселя первого рода; $N_{\frac{1}{3}}$ — функция Бесселя второго рода. Коэффициенты A и B , как обычно, определяются из начальных условий. Решение (4.59), требует большего числа операций, чем каноническое решение при $\omega_k = \text{const}$, но экономия за счет более редкого использования процедуры решения по Галеркину приводит к значительному понижению общего числа операций.

Если в речевом тракте существует какой-либо источник возбуждения, распределенный по его длине, то его уравнение становится неоднородным (в правой части находится функция возбуждения $F(x, t)$). Действуя аналогично методике, описанной в § 1.3, находим парциальные возбуждения $F_k(t)$ для каждой временной моды:

$$F_k(t) = -\frac{2}{\rho_0 l} \int_0^l F(x, t) \psi_k(x) dx. \quad (4.60)$$

Тогда уравнение (4.41) представляется в виде

$$T_k'' + \lambda_k^2 c_0^2 T_k = F_k(t).$$

В зависимости от распределения возбуждения по длине речевого тракта амплитуда k -й временной моды может быть больше или меньше, а в некоторых случаях эта мода вообще не возбуждается. Действительно, если возбуждение сосредоточено в одной точке x_0 , то можно записать $F(x, t) = f(t) \delta(x - x_0)$, где δ — функция Дирака. Тогда (4.60) принимает вид

$$F_k(t) = -\frac{2}{\rho_0 l} \psi_k(x_0) f(t).$$

Если x_0 совпадает с пучностью собственной функции ψ_k , то эффект воздействия F_k на k -ю моду будет максимальным. Если же x_0 приходится на нуль ψ_k , то и $F_k = 0$ и, следовательно, эта мода не возбуждается. Это явление не изменяется по существу и в том случае, если источник не сконцентрирован в одной точке, а распределен на некотором интервале.

Зависимость амплитуды возбуждения временных мод от положения источника в речевом тракте объясняет некоторые особенности спектральных характеристик фрикативных звуков и взрывного сегмента согласных звуков, в частности появление нулей в передаточной функции речевого тракта.

4.2. ИЗЛУЧЕНИЕ РЕЧЕВОЙ ВОЛНЫ

Излучение речевого сигнала в пространство происходит через ротовое отверстие и ноздри, а при сомкнутых губах и поднятой небной занавеске речевой сигнал излучается главным образом через губы и стенки речевого тракта в области гортани. Излучение через ткани речевого тракта составляет незначительную долю энергии сигнала, излучаемого через губное отверстие и ноздри, но оно становится основным для взрывных согласных с полной смычкой и определенным образом влияет на различие между глухими и звонкими взрывными.

Рассмотрим процесс излучения речевого сигнала через ротовое отверстие и губы. Обозначим проводимость ротового отверстия $Y_{\text{и}}$, а проводимость тканей губ $Y_{\text{г}}$, тогда проводимость излучения со стороны губ

$$Y_{\text{л}} = \gamma Y_{\text{и}} + (1 - \gamma) Y_{\text{г}},$$

где γ — коэффициент трансформации акустического сигнала; $\gamma = S_{\text{и}}/S_{\text{г.т}}$; $S_{\text{и}}$ — площадь ротового отверстия; $S_{\text{г.т}}$ — площадь речевого тракта непосредственно перед губами.

В [4.13] было получено выражение для импеданса излучения круглого поршня, радиус $a_{\text{п}}$ которого мал по сравнению с радиусом $a_{\text{э}}$ сферического экрана, в котором этот поршень находится:

$$\tilde{Z}_{\text{и}} = \pi a^2 \rho_0 c_0 (R_{\text{и}} - j X_{\text{и}}),$$

где a — радиус эквивалентного плоского поршня, равного по площади сферическому сегменту; $a = 2a_{\text{п}} \sin \frac{\vartheta_0}{2}$; ϑ_0 — максимальный угол, определяющий направление излучения по нормали к поверхности поршня. Средние значения активной и реактивной компонент импеданса излучения поршня представляются через полиномы Лежандра P_m и сферические функции Неймана n_m :

$$R_{\text{и}} = \frac{1}{2} \sum_{m=0}^{\infty} \frac{[P_{m-1}(\cos \vartheta_0) - P_{m+1}(\cos \vartheta_0)]^2}{v^2 (2m+1) D_m^2};$$

$$X_{\text{и}} = \frac{1}{2} \sum_{m=0}^{\infty} \frac{[P_{m-1}(\cos \vartheta_0) - P_{m+1}(\cos \vartheta_0)]^2}{(2m+1) D_m} [j_m(v) \sin \delta_m - n_m(v) \cos \delta_m],$$

где $v = 2\pi a/\lambda$; λ — длина волны; $j_m(v) = \sqrt{\frac{\pi}{2v}} J_{m+\frac{1}{2}}(v)$; $n_m(v) =$

$= \sqrt{\frac{\pi}{2v}} N_{m+\frac{1}{2}}(v)$; J — функция Бесселя первого рода; N — функции

Неймана; полином Лежандра

$$P_m(x) = \frac{1}{2^m m!} \frac{d^m}{dx^m} (x^2 - 1)^m.$$

При малых значениях Ka (K — волновое число), точнее при $Ka \ll m+1$, допустима аппроксимация параметров D_m и δ_m :

$$D_m \approx \frac{1 \cdot 3 \cdot 5 \dots (2m-1)(m+1)}{(Ka)^{m+2}}, \quad D_0 \approx \left(\frac{1}{Ka}\right)^2;$$

$$\delta_m \approx \frac{m(Ka)^{2m+1}}{12 \cdot 3^2 \cdot 5^2 \dots (2m-1)(2m+1)(m+1)}, \quad \delta_0 \approx \frac{(Ka)^3}{3}.$$

Тогда удельный безразмерный импеданс излучения

$$Z_{\text{и}} = \frac{\tilde{Z}_{\text{и}}}{\rho_0 c_0 S_{\text{и}}} = 1 - \frac{J_1(Ka)}{Ka} - j \frac{K_1(2Ka)}{2(Ka)^2},$$

где K_1 — функция Бесселя второго рода; $S_{\text{и}} = \pi a^2$. При малых значениях Ka можно ограничиться первыми членами разложения функций Бесселя J_1 и K_1 в степенные ряды. Если $Ka < 0,25$, то $R_{\text{и}} \approx (Ka)^2/2$, а при $Ka < 0,5$ $X_{\text{и}} \approx 8Ka/3\pi$. По данным [В.4] отношение эквивалентного радиуса ротового отверстия a к радиусу головы a_0 находится в диапазоне 0,06—0,14. При таких соотношениях погрешность приближенного представления $R_{\text{и}}$ и $X_{\text{и}}$ по сравнению с точным решением — менее 10%, поэтому обычно принимают импеданс излучения

$$Z_{\text{и}} = \frac{(Ka)^2}{2} - j \frac{8Ka}{3\pi}. \quad (4.61)$$

Активная компонента импеданса излучения учитывает потери энергии, уходящей из речевого тракта. Эта компонента пропорцио-

нальна площади ротового отверстия и квадрату частоты. Реактивная компонента учитывает присоединенную массу воздуха у источника излучения. Эта компонента создает так называемую концевую поправку на длину акустической системы. Физический смысл этой поправки в том, что система как бы удлиняется вследствие отражения акустической волны не от ее края, а на некотором расстоянии от него. В результате уменьшаются резонансные частоты трубы. Например, для однородной трубы длиной 17,5 см и площадью поперечного сечения $S = 5 \text{ см}^2$ при $S_{\text{и}} = S$ концевая поправка равна примерно 1 см, что составляет 6% исходной длины трубы.

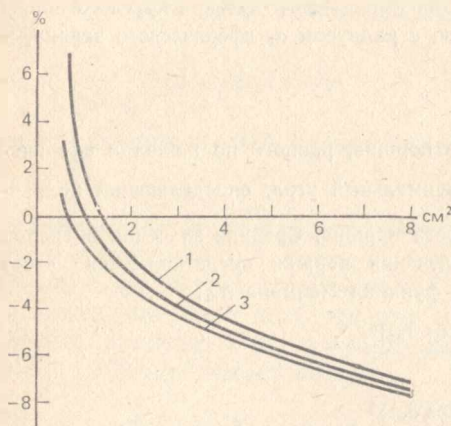


Рис. 4.7. Поправки на скорость звука в однородной трубе с прямоугольным сечением шириной 5 см в зависимости от площади сечения на частотах:

1 — 500 Гц; 2 — 1500 Гц; 3 — 2500 Гц

Пример поправок на собственные частоты однородной трубы с учетом излучения и податливости стенок показан на рис. 4.7.

Рассчитаем теперь сопротивление, оказываемое тканями губ, обозначив импеданс губ как $Z_{\text{г}} = 1/Y_{\text{г}}$. Реактивную составляющую импеданса $Z_{\text{г}}$ найдем по формуле для многослойных систем [4.7]

$$\tilde{Z}_{\text{г}} = j \rho_{\text{т}} c_{\text{т}} \frac{\rho_0 c_0 Z_{\text{и}} - j \rho_{\text{т}} c_{\text{т}} \operatorname{tg} K h_{\text{г}}}{\rho_0 c_0 Z_{\text{и}} \operatorname{tg} K h_{\text{г}} + j \rho_{\text{т}} c_{\text{т}}},$$

где ρ_t — плотность тканей губ; c_t — скорость звука в тканях; h_t — толщина губ; Z_n — импеданс нагрузки; K — волновое число. Нагрузкой для губ служит импеданс излучения, рассчитываемый аналогично (4.61), но с учетом площади губ. Произведение $Kh_t = \omega h_t / c_t$ значительно меньше единицы на частоте до 3 кГц, поэтому для низких частот справедлива замена $\operatorname{tg} Kh_t \approx Kh_t$. Значения импеданса излучения при низких частотах очень малы, и членами $\rho_0 c_0 Z_n$ и $\rho_0 c_0 Z_n \operatorname{tg} Kh_t$ можно пренебречь. При толщине губ $h_t = 1,5$ см коэффициент m_r в реактивной компоненте удельного безразмерного импеданса губ $Z_r = r_r - j\omega m_r$ равен примерно 0,04. Активная часть оказывается очень малой, поскольку не принимаются в расчет потери в тканях. Измерения же показывают, что при сомкнутых губах уровень излучаемого сигнала примерно на 20—40 дБ ниже уровня сигнала при раскрытых губах.

Оценим влияние импеданса излучения при раскрытых губах на резонансные частоты речевого тракта и затухание акустических колебаний. Для этого рассмотрим однородную трубу длиной l , закрытую на конце $x=0$ абсолютно жесткой стенкой и нагруженную на импеданс излучения Z_l на конце $x=l$. Для такой трубы можно записать уравнение относительно волновых чисел

$$(n-1)\pi - Kl = \operatorname{arctg}(jY_l), \quad n = 1, 2, \dots, \quad (4.62)$$

где Y_l — эффективная удельная безразмерная проводимость излучения [4.7]. Пусть $Kl = K_0 l + j\alpha + \beta$, где K_0 — волновые числа некоторой модели нулевого приближения (например, трубы с открытым концом при $x=l$ без учета импеданса излучения, т. е. при $Z_l=0$, или закрытой на этом конце абсолютно жесткой стенкой, т. е. $Y_l=0$); $|j\alpha + \beta| \ll 1$. Если площадь излучающего отверстия S_n не равна площади трубы S , то эффективный импеданс излучения $Z_l = \gamma Z_n$, где $\gamma = S/S_n$. Для «почти открытой» трубы, т. е. при $S_n \approx S$, имеем $K_0 l = (n-0,5)\pi$, поэтому из (4.62) получаем $Z_l = j \operatorname{tg}(j\alpha + \beta)$, или с учетом малости α и β $Z_l \approx j\beta - \alpha$. Пренебрегая величинами второго порядка малости, получаем простые выражения:

$$\alpha = -\gamma \frac{(K_0 a)^2}{2} \frac{1-\eta}{(1+\eta)^2}; \quad \beta = -\gamma \frac{8K_0 a}{3\pi(1+\eta)},$$

где $\eta = 8a\gamma/3\pi l$. При $l = 17,5$ см и $a = 1$ см $\eta \ll 1$ ($\gamma \approx 1$ по условию «почти открытой» трубы). В этом случае скорость изменения величины β почти пропорциональна скорости изменения радиуса излучающего отверстия, а коэффициент пропорциональности изменяется примерно от 0,1 на частоте 500 Гц до 1 на частоте 5 кГц. Это значит, что для почти полностью открытой трубы скорость изменения резонансных частот меньше скорости изменения размеров излучающего отверстия, а скорость изменения первого резонанса значительно меньше скорости изменения радиуса излучающего отверстия. Отметим, что скорости изменения разных резонансных частот оказываются различными.

Если труба «почти закрыта», т. е. $S_n \ll S$, то нулевое приближение соответствует закрытой трубе, для которой справедливо соот-

ношение $K_0 l = (n-1)\pi$, поэтому из (4.62) получаем $\gamma Z_l(j\beta - \alpha) \approx 1$ и поправки $\alpha = -9\pi^2/128\gamma$ и $\beta = 1,18/K_0 a \gamma$. Условие малости α и β , позволяющее совершить замену $\operatorname{tg}(j\alpha + \beta)$ на $j\alpha + \beta$ в этом случае выполняется лишь при очень малых раскрытиях излучающего отверстия ($S_{\text{и}}/S < 10^{-2}$), и полученные выражения для α и β представляют интерес лишь в асимптотическом смысле. Они показывают, что по мере закрывания излучающего отверстия потери все меньше зависят от частоты, тогда как поправки на собственные числа все сильнее зависят от частоты. Эта модель позволяет также обнаружить очень важное явление ускорения изменений собственных частот трубы. Поскольку в диапазоне частот речевого сигнала излучающее отверстие любой формы можно приближенно представить в виде круга, $S_{\text{и}} = \pi a^2$ и $\beta = 1,18\pi a/K_0 S$. Здесь коэффициент пропорциональности между скоростью изменения собственных чисел и скоростью изменения эффективного радиуса излучающего отверстия гораздо больше единицы, особенно для нижних частот. Например, для резонансной частоты $F_1 = 500$ Гц и средней площади трубы 4 см^2 скорость изменения F_1 почти в 10 раз больше скорости изменения a . Это означает, что и в процессах артикуляции скорость изменения резонансных частот речевого тракта может оказаться значительно выше скорости движения артикуляторных органов, если площадь поперечного сечения речевого тракта становится очень малой, как перед смычкой или после ее взрыва. Таким образом, для расчета речевой волны в этих случаях скорость изменения резонансных частот является существенным параметром, и ее необходимо учитывать хотя бы в форме, предлагаемой в предыдущем параграфе. В силу этого явления представление речевой волны в виде суперпозиции затухающих гармонических колебаний на участках с быстрым изменением резонансных частот приводит к большой погрешности, создающей искажения, характерные для формантных синтезаторов речи.

Пространственные характеристики звукового поля источника определяют параметры сигнала в различных точках пространства. Если линейные размеры излучающей поверхности (например, ротового отверстия) малы по сравнению с длиной излучаемой волны, то независимо от конкретной формы излучающей поверхности она может быть принята за плоский круглый поршень с той же площадью, поскольку в этом случае распределение скоростей частиц на поверхности излучения однородно и синфазно. Форма головы у человека близка к сферической, и, как показано в [4.13], ее поле хорошо описывается звуковым полем сферического источника.

Уравнение колебаний сферы относительно давления есть

$$\frac{1}{r^2} \frac{\partial}{\partial r} \left(r^2 \frac{\partial P}{\partial r} \right) = \frac{1}{c_0^2} \frac{\partial^2 P}{\partial t^2},$$

и его решение по Даламберу

$$P = \frac{f(r - c_0 t) - f(r + c_0 t)}{r}.$$

На поверхности шара радиусом a объемная скорость $Q(t) = 4\pi a^2 v_r(t)$, где $v_r(t)$ — радиальная скорость поверхности. Воспользовавшись связью между скоростью и давлением $\rho_0 \partial v_r / \partial t = -\partial P / \partial r$ и подставив для расходящейся волны $P = f(r - c_0 t) / r$, получим

$$\frac{1}{r} \frac{\partial f}{\partial r} - \frac{f}{r^2} = \frac{\rho_0}{4\pi a^2} \frac{\partial Q}{\partial t}. \quad (4.63)$$

Для гармонических колебаний, записанных в форме $v_r(t) = v_0 e^{-jKc_0 t}$, решение (4.63) есть

$$P(r) = v_0 \frac{\rho_0 a c_0}{r} \frac{jKa}{1 + jKa} e^{-jK(r-a)}. \quad (4.64)$$

Скорость частиц при $r \gg a$

$$v_r(r) = v_0 \frac{a}{r} \frac{jKa}{1 + jKa} \frac{1 + jKr}{jKr} e^{-jK(r-a)}. \quad (4.65)$$

При малых размерах шара относительно длины волны излучаемого звука ($Ka \ll 1$) источник излучения называется точечным или гипотетическим и его поле давления

$$P(r) \approx \frac{j\omega \rho_0 v_0 a^2}{r} e^{-jKr}. \quad (4.66)$$

Из (4.65) видно, что любой микрофон градиента давления, воспринимающий скорость звуковых колебаний, при приближении к источнику звука усиливает низкие частоты из-за jKr в знаменателе. В ближнем поле изменение расстояния до источника звука должно приводить к амплитудным модуляциям низкочастотных компонент сигнала. Член $jKa/(1 + jKa)$ в (4.64) и (4.65) соответствует операции так называемого реального дифференцирования, т. е. дифференцирования со сглаживанием, поэтому и сигнал на выходе речевого тракта подвергается дифференцированию.

Решая сферическое волновое уравнение относительно потенциала скорости

$$\frac{1}{r} \frac{\partial^2 (r\Phi)}{\partial r^2} = \frac{1}{c_0^2} \frac{\partial^2 \Phi}{\partial t^2}, \quad (4.67)$$

для расходящейся волны давление и скорость в звуковом поле можно записать так:

$$P = \frac{\rho_0}{r} \frac{\partial f_1(t - r/c_0)}{\partial t}; \quad (4.68)$$

$$v_r = \frac{1}{r^2} f_1(t - r/c_0) + \frac{1}{rc_0} \frac{\partial f_1(t - r/c_0)}{\partial t}, \quad (4.69)$$

где $f_1(t - r/c_0)$ — расходящаяся компонента решения Даламбера сферического волнового уравнения (4.67). В расходящейся волне P и v_r связаны соотношением

$$v_r = \frac{P}{\rho_0 c_0} + \frac{1}{\rho_0 r} \int_{-\infty}^t P dt. \quad (4.70)$$

Человеческое ухо чувствительно не только к звуковому давлению, но и к скорости колебания частиц воздуха, поэтому различия в кривизне волнового фронта могут ощущаться человеком как различия в удалении до источника. В 1938 г. Бекеши показал, что для ощущения расстояния важны низкочастотные составляющие сигнала. Это вытекает из возможности измерения различий между давлением и скоростью: на высоких частотах второй член в (4.69) значительно больше первого и скорость v_r становится неотличимой от давления P . У длительных установившихся звуков давление и скорость различаются только фазой, поскольку в (4.70) преобладает второй член. Таким образом, оценка расстояния может опираться только на переходные процессы и низкочастотные компоненты сигнала.

Для оценки характеристик давления и скорости важны понятия ближнего и дальнего полей. Зоны ближнего и дальнего полей различны для разных частот и определяются значением произведения волнового числа $K = \omega/c_0 = 2\pi/\lambda$ (λ — длина волны) на расстояние до источника r : $Kr > 1$ — дальнее поле; $Kr \ll 1$ — ближнее поле. Выражение для давления P сохраняет свой вид (4.66) или (4.69) на всех расстояниях — и в ближнем, и в дальнем полях. Для низких частот $\partial f_1/\partial t$ мало, и в ближнем поле

$$v_r \approx \frac{1}{r^2} f_1(t - r/c_0),$$

а в дальнем поле

$$v_r \approx \frac{1}{rc_0} \frac{\partial f_1(t - r/c_0)}{\partial t}.$$

Первое выражение соответствует (4.65), а второе выражение с учетом $Kr \ll 1$ дает

$$v_r(r) = \frac{v_0}{4\pi r^2} (1 - jKr) e^{-jKr}.$$

4.3. ГОЛОСОВАЯ ЩЕЛЬ И ПОДСВЯЗОЧНАЯ ОБЛАСТЬ

В большинстве случаев при артикуляции голосовые складки сближаются таким образом, что акустическое сопротивление голосовой щели оказывается весьма большим по сравнению с акустическим сопротивлением всего речевого тракта. Поэтому обычно верхняя поверхность голосовых складок считается началом речевого тракта, а голосовая щель и подвязочная область представляются в виде эквивалентного импеданса Z_0 , характеризующего граничные условия в начале координат $x=0$. Такое представление справедливо и в том случае, когда импеданс Z_0 сравним с полным импедансом речевого тракта.

Обозначим проводимость акустической системы при $x=0$ через $Y_0 = 1/Z_0$. Эта проводимость складывается из проводимости голосовой щели $Y_{г.щ}$ и проводимости голосовых складок $Y_{г.с}$:

$$Y_0 = \gamma Y_{г.щ} + (1 - \gamma) Y_{г.с},$$

где $\gamma = S_{г.щ}/S_0$; $S_{г.щ}$ — площадь голосовой щели; S_0 — площадь ре-

ческого тракта при $x=0$. Оценим каждую компоненту проводимости Y_0 . Безразмерный удельный импеданс голосовых складок аналогично импедансу губ определяется по формуле для многослойных систем

$$Z_{г.с} = j\xi \frac{Z_n - j\xi \operatorname{tg} K h_{г.с}}{Z_n \operatorname{tg} K h_{г.с} + j\xi},$$

где $\xi = \rho_t c_t / \rho_0 c_0$; $h_{г.с}$ — толщина голосовых складок; Z_n — импеданс, на который нагружены голосовые складки со стороны трахеи.

В подсвязочной области различаются следующие элементы: трахея, бронхи, бронхиолы и альвеолы. Последние два элемента образуют легкие. По данным [4.5] длина трахеи 9—15 см, а ее ширина 1,5—2,7 см; площадь трахеи равна примерно 3 см². Трахея делится на два бронха — правый (более широкий, длина 2—3 см, площадь около 2 см²) и левый (длина 4—5 см и площадь около 1,5 см²). Правый бронх делится на три ветви, а левый — на две. Эти ветви разделяются на бронхиолы диаметром около 0,1 см, а каждая бронхиола, в свою очередь, делится на 12—18 концевых бронхиол с выпячиваниями в их стенках (альвеолами) со средним диаметром 0,02—0,025 см.

Суммарная площадь бронхов, а также их ветвей близка к площади трахеи. В случае точного равенства площадей и длин ветвлений звуковая волна в месте разветвления не отражается и эквивалентную длину трахеи можно считать равной сумме длин самой трахеи, бронхов и их ветвей.

По данным [4.27] о распределении площади поперечного сечения дыхательной системы эквивалентная длина трахеи примерно 20 см (см. рис. 4.8). Поскольку импеданс трахеи как однородной трубы можно легко вычислить, переносный импеданс Z_n вдоль трахеи удобно выразить как

$$Z_n = \frac{Z_l - \operatorname{th} j K l_{тр}}{1 - Z_l \operatorname{th} j K l_{тр}}, \quad (4.71)$$

где $l_{тр}$ — эквивалентная длина трахеи; Z_l — импеданс легких.

Число дихотомий (делений бронхиол на более мелкие ветви) известно, поэтому, зная статистику размеров ветвей, можно было бы оценить импеданс легких Z_l посредством многократного применения формулы для переносного импеданса. Но этот путь довольно трудоемок, и мы применим более простой метод, представив легкие в виде полости с объемом V_l . Конечные ветви бронхиол и альвеолы имеют малые геометрические разме-

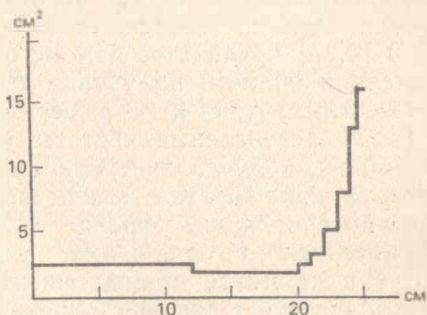


Рис. 4.8. Площадь поперечного сечения дыхательной системы.

По оси абсцисс — расстояние от голосовых складок

ры, поэтому в них преобладают потери на вязкое трение капиллярного типа, а безразмерный удельный импеданс легких

$$Z_{\text{л}} = R_{\text{л}} + \frac{j}{\omega X_{\text{л}}},$$

где $X_{\text{л}} = V_{\text{л}}/\rho_0 c^2_0 S_{\text{тр}}$; $S_{\text{тр}}$ — площадь трахен.

Рассмотрим подглоточную область как однородную трубу с жесткими стенками длиной $l_{\text{тр}}$, нагруженную, с одной стороны, на импеданс легких $Z_{\text{л}}$, а с другой — на импеданс излучения $Z_{\text{и}}$ из подсвязочной области в речевой тракт. Для такой системы по [4.7] существует следующее соотношение:

$$(n-1)\pi - K l_{\text{тр}} = \arctg(j Y_{\text{л}}) + \arctg(j Y_{\text{и}}), \quad (4.72)$$

где $Y_{\text{л}}$ и $Y_{\text{и}}$ — удельные безразмерные проводимости легких и голосовой щели: $Y_{\text{л}} = 1/Z_{\text{л}}$; $Y_{\text{и}} = 1/Z_{\text{и}}$; K — волновое число; $n = 1, 2, 3 \dots$

Представим снова $Kl = K_0 l_{\text{тр}} + j\alpha + \beta$, где K_0 — волновое число трубы с известными импедансами $Z_{\text{л}}$ и $Z_{\text{и}}$; $|j\alpha + \beta| \ll 1$.

Пренебрегая импедансом излучения $Z_{\text{и}}$ и принимая $K_0 l_{\text{тр}} = (n-0,5)\pi$, можно записать $jZ_{\text{л}} \tg(j\alpha + \beta) = 1$, или учитывая малость α и β ,

$$Z_{\text{л}}(j\beta - \alpha) \approx 1. \quad (4.73)$$

Разделяя в (4.73) мнимую и действительную части и решая полученную систему уравнений, имеем

$$\alpha = \frac{R_{\text{л}}}{R_{\text{л}}^2 + \left(\frac{1}{\omega X_{\text{л}}}\right)^2}.$$

Для частот выше 500 Гц реактивным членом при оценке α можно пренебречь, и тогда $\alpha \approx 1/R_{\text{л}}$. Представим $R_{\text{л}} = r_1 + r_2 \sqrt{f} + r_3 f^2$, где f — частота. Физический смысл первого члена соответствует потерям на капиллярное трение в альвеолах легких. Второй член отражает потери на некапиллярное трение в бронхиолах. Третий член появляется, если учесть излучение стенок легких.

В [4.27] приводятся данные о ширине резонансных частот подглоточной области на уровне дыхательного отверстия. Имея три неизвестных: r_1 , r_2 и r_3 , выберем в качестве опорных измеренные в [4.27] резонансные частоты $F_1 = 640$ Гц; $F_2 = 1400$ Гц и $F_3 = 2850$ Гц, получив линейную систему уравнений, задача решения которой эквивалентна задаче интерполяции. Как и во всякой задаче интерполяции, решения относительно r_i справедливы только на интервале интерполяции (в данном случае 640—2850 Гц) и могут и не иметь прямой физической интерпретации в силу неоднозначности решения. Решения этой системы дают следующие значения: $r_1 = 1,097$; $r_2 = 1,526 \cdot 10^{-2}$; $r_3 = 1,012 \cdot 10^{-7}$. Зависимость ширины полосы резонансов от частоты при таких коэффициентах r_i показана на рис. 4.9.

Потери в тканях удобно представить через комплексную составляющую скорости звука: $c_T = c^{(0)}_T (1 - j\xi_T)$. Имеется ряд работ, описывающих результаты измерений импеданса тканей человека, однако, для того чтобы оценить потери в тканях, необходимо сделать некоторые предположения об условиях измерения импеданса. Обычно в измеренных импедансах в реактивном члене преобладает инерционная составляющая, поэтому удобно допустить, что в (4.71) импеданс слоев, на который опирается измеряемый участок ткани, $Z_H = 0$. Тогда, сопоставляя активную и реактивную части измеренного импеданса $Z_T = R_T - j\omega M_T$ с импедансом $Z = \xi(\xi_T - j) \operatorname{tg} Kh$, получаем $\xi_T = R_T / \omega M_T$.

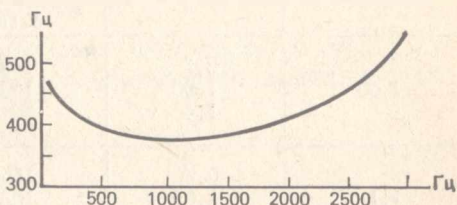


Рис. 4.9. Ширина полосы резонансов в дыхательной системе

Оценим теперь импеданс голосовой щели. Податливые стенки в голосовой щели изменяют ее частотные характеристики и вносят определенный вклад в потери, особенно на низких частотах. Как следует из (4.16), для трубки с податливыми стенками волновые числа можно записать в виде

$$K^2 = K_0^2 \left(1 - \frac{2Y_T}{jK_0 a} \right),$$

где Y_T — удельная безразмерная проводимость стенок; K_0 — волновое число в трубе с жесткими стенками; a — радиус трубки или наименьшая сторона прямоугольного сечения. Импеданс стенок внутри голосовой щели $Z_T = 1/Y_T$ рассчитаем как импеданс системы, состоящей из сосредоточенной массы и сосредоточенной упругости, опертой на жесткую стенку:

$$Z_T = R_T + j \left(\frac{C_{г.с}}{\omega} - \omega M_T \right),$$

где $C_{г.с}$ — упругость голосовых складок; $C_{г.с} = E/\delta_{г.с}$; M_T — поверхностная плотность тканей: $M_T = \rho_T \delta_{г.с}$; $\delta_{г.с}$ — толщина голосовых связок в направлении наименьшего размера голосовой щели. Потери на колебания стенок R_c найдем из мнимой части волнового числа K с учетом Z_T .

Воспользовавшись данными, приведенными в гл. 1, относительно модуля упругости расслабленных и напряженных мышечных тканей голосовых складок, оценим частотные параметры голосовых складок. Изменения модуля упругости в 60 раз совместно с изменением толщины стенок в пределах 0,7—1,4 см и плотности тканей в пределах 1,05—1,15 г/см³ дают резонансные частоты голосовых складок в диапазоне 76—390 Гц (табл. 4.1), что примерно соответствует диапазону наблюдаемых частот колебаний голосовых складок людей. Как видно, эта весьма упрощенная по сравнению с рассматривавшейся в гл. 2 модель голосовых складок об-

Таблица 4.1

Плотность тканей, г/см ³	Толщина тканей, см	Частота колебаний голосовых складок, Гц, при модуле упругости, Па		
		5·10 ⁴	1,5·10 ⁵	3·10 ⁵
1,05	0,7	158	274	388
	1,05	106	183	259
	1,4	79	137	194
1,1	0,7	155	269	379
	1,05	103	179	253
	1,4	77	134	190

ладает удовлетворительными характеристиками, позволяющими использовать ее для расчета акустических процессов в речевом тракте.

Кроме потерь на податливость стенок в голосовой щели необходимо учесть потери на трение в пограничном слое и кинетическое сопротивление, возникающее в результате эффекта Бернулли. Если раскрытие голосовой щели меньше некоторого значения $a_{г.щ} \leq \Delta \sqrt{2}$, то потери на трение описываются законом Пуазейля для капиллярных трубок [2.5]

$$R_r = 12 \mu h_{г.щ} / a_{г.щ}^2,$$

где $h_{г.щ}$ — глубина голосовой щели; μ — коэффициент вязкости воздуха. Так называемый пограничный слой Прандтля Δ оценивается как $\Delta = \sqrt{2\mu/\rho_0\omega}$, где $\omega = 2\pi f$ — круговая частота. Как упоминалось в § 2.4, даже на частоте $f = 100$ Гц пограничный слой $\Delta \approx 0,02$ см, так что капиллярное трение преобладает лишь на начальной стадии раскрытия голосовой щели.

Если $a_{г.щ} \gg \Delta \sqrt{2}$, то действует другой закон трения:

$$R_r = \frac{h_{г.щ} L_{г.щ}}{S_{г.щ}} \sqrt{\frac{\rho_0 \omega \mu}{2}}, \quad (4.74)$$

где $L_{г.щ}$ и $S_{г.щ}$ — соответственно периметр и площадь голосовой щели.

В промежуточном случае $\Delta \sqrt{2} < a_{г.щ} < 10 \Delta \sqrt{2}$ закон трения весьма сложен, однако и тогда (4.74) может служить удовлетворительным приближением.

Кинетическое сопротивление оценим по [В.3] как $R_b = 0,875 v / c_0$, где v — линейная скорость воздушного потока через голосовую щель.

Потери на трение и колебание стенок в голосовой щели при $a_{г.щ} = 0,1$ см, $l_{г.щ} = 2$ см и $v = 3,5 \cdot 10^3$ см/с сопоставляются на рис. 4.10.

Реактивная часть импеданса голосовой щели должна учитывать и концевые эффекты [В.3], так что окончательная форма выражения импеданса голосовой щели принимается в следующем виде:

$$Z_{г.щ}^{(0)} = R_T + R_B - j 1,2 h_{г.щ} \sqrt{1 - \frac{2 Y_T}{j K_0 a_{г.щ}}}.$$

Эквивалентный импеданс голосовой щели состоит из суммы собственного импеданса щели $Z_{г.щ}^{(0)}$ и переносного импеданса легких:

$$Z_{г.щ} = Z_{г.щ}^{(0)} + \frac{j Z_L - \operatorname{th} j K l_{TP}}{1 - Z_L \operatorname{th} j K l_{TP}}. \quad (4.75)$$

Рассматривая в общем виде однородную трубу, нагруженную на одном конце на импеданс Z_0 , а на другом — на импеданс Z_l , можно описать ее частотные характеристики через (4.72), преобразовав это уравнение к виду

$$G(K) = \operatorname{tg} K l + j \frac{Y_0 + Y_l}{1 + Y_0 Y_l} = 0, \quad (4.76)$$

где Y_0 и Y_l — проводимости переднего и заднего концов трубы.

Проводимость Y_0 рассматривалась выше, а проводимость излучающего конца

$$Y_l = Y_{из} + (1 - \eta) Y_{п.с},$$

где $Y_{из}$ — проводимость излучающего отверстий; $Y_{п.с}$ — проводимость передней стенки; $\eta = S_{из}/S_{г.т}$; $S_{из}$ — площадь излучающего отверстия; $S_{г.т}$ — площадь голосовой щели.

Предположим для простоты, что проводимость передней стенки равна нулю, т. е. $Y_{п.с} = 0$, а импеданс излучения $Z_{из}$ принимаем в простейшем виде как (4.61). Уравнение (4.76) является нелинейным и, очевидно, может быть решено только численными методами. Рассмотрим сначала свойства (4.76) без учета влияния потерь. Результаты решения приведены в табл. 4.2 при следующих параметрах модели: $l_{г.т} = 17,5$ см; $l_{TP} = 20$ см; $S_{г.т} = S_{TP} = 3$ см². Как видно из этой таблицы,

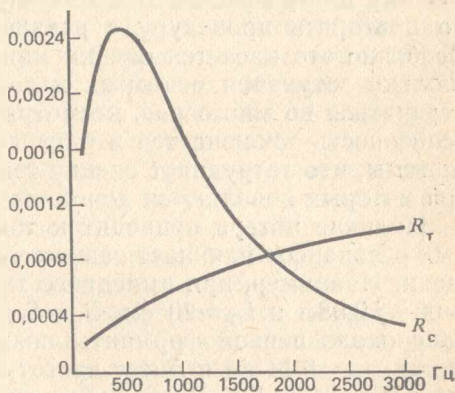


Рис. 4.10. Потери на трение R_T и колебания стенок $R_с$ в голосовой щели

при открытой голосовой щели появляются дополнительные резонансы, частоты которых почти не зависят от степени раскрытия голосовой щели. Распределение частот дополнительных формант на-

Ширина сужения, см	Резонансная частота, Гц									
0	3	479		1434		2391		3355		
0,01	3	29	479	855	1432	1708	2387	2564	3340	3426
0,1	3	28	476	859	1425	1718	2372	2582	3311	3460

ходится в удовлетворительном совпадении с результатами расчетов, проведенных в [4.27].

Учет потерь усложняет процедуру поиска решений (4.76) и заставляет пользоваться другим методом, который состоит в разделении (4.76) на мнимую и действительную части: $\text{Im } G(K)=0$; $\text{Re } G(K)=0$. Эта система нелинейных уравнений решается методом Ньютона — Рапсона относительно мнимой и действительной частей волнового числа $K=X-jR$. Значение K на $(i+1)$ -й итерации определяется как $K_{i+1}=K_i+\delta K_i$, где δK_i — вектор, являющийся решением линейной системы $\mathbf{H}(K_i)\delta K_i=-G(K_i)$, а $\mathbf{H}(K_i)$ есть матрица частных производных

$$\begin{vmatrix} \frac{\partial}{\partial X} \text{Re } G(K_i) & \frac{\partial}{\partial R} \text{Re } G(K_i) \\ \frac{\partial}{\partial X} \text{Im } G(K_i) & \frac{\partial}{\partial R} \text{Im } G(K_i) \end{vmatrix}.$$

Достоинством этого метода является быстрая сходимость, однако он весьма чувствителен к выбору начальных значений $X^{(0)}$ и $R^{(0)}$, поэтому для уверенности в том, что все корни найдены, необходимо повторять процедуру с различными начальными условиями. Особенно это касается оценки мнимой части волнового числа, поскольку затухания основных и дополнительных резонансов могут отличаться во много раз, несмотря на близость их частот. Другая особенность заключается в близости резонансов и антирезонансов системы, что затрудняет оценку тех или иных параметров системы, при которых появляются дополнительные резонансы.

Наличие потерь приводит к тому, что появление дополнительных резонансов начинает зависеть от степени раскрытия голосовой щели. Например, при импедансе тканей, измеренном в [4.30], $Z_T = 60 - j 0,08 \omega$ и $l_{TP} = 20$ см, $S_H = S_{HT} = 3$ см², дополнительный резонанс около первой форманты появляется при ширине голосовой щели $a_{г.щ} = 0,14$ см и имеет частоту $F^{(1)}_1 = 450$ Гц при ширине полосы $\Delta F^{(1)}_1 = 430$ Гц. Второй дополнительный резонанс появляется уже при $a_{г.щ} = 0,2$ см и имеет частоту $F^{(1)}_2 = 1258$ Гц при ширине $\Delta F^{(1)}_2 = 416$ Гц. Третий и четвертый дополнительные резонансы не появляются до максимально исследованного раскрытия голосовой щели, равного $a_{г.щ} = 0,25$ см. Изменения частот основной первой форманты $F^{(0)}_1$ и дополнительной форманты $F^{(1)}_1$ невелики: они составляют соответственно примерно 2,3 и 1,5% при $a_{г.щ} = 0,2$ см.

При укорочении трахеи до $l_{\text{тр}}=15$ см дополнительные резонансы не появляются на всем исследованном диапазоне раскрытой голосовой щели. Девиация частоты $F^{(0)}_1$ при этом меняет знак, оставаясь по величине той же самой: $-2,3\%$ при $a_{\text{г.щ}}=0,2$ см. Увеличение площади голосового тракта $S_{\text{г.т}}$ до 6 см^2 при полном раскрытии излучающего отверстия $S_{\text{и}}=S_{\text{г.т}}$ также уничтожает дополнительные резонансы.

При этих вычислениях выяснилась большая роль импеданса тканей $Z_{\text{т}}$. Так, для импеданса, соответствующего данным [4.33], $Z_{\text{т}}=200-j0,05\omega$, влияние подглоточной части значительно сильнее: первый дополнительный резонанс $F^{(1)}_1$ появляется уже при $a_{\text{г.щ}}=0,04$ см, резонанс $F^{(1)}_4$ — при $a_{\text{г.щ}}=0,1$ см, а резонанс $F^{(1)}_2$ — при $a_{\text{г.щ}}=0,14$ см (рис. 4.11). Изменения частот $F^{(0)}_1$ и $F^{(1)}_1$ в процессе раскрытия голосовой щели также значительно больше: они составляют для $F^{(0)}_1$ примерно $9,2\%$, а для $F^{(1)}_1$ около $8,4\%$. Тем не менее при укорочении трахеи до $l_{\text{тр}}=15$ см все резонансы вновь исчезают, а девиация частоты $F^{(0)}_1$ меняет знак. Увеличение площади голосового тракта до $S_{\text{г.т}}=6 \text{ см}^2$ ($S_{\text{и}}=S_{\text{г.т}}$) при этом импедансе стенок приводит к тому, что дополнительные резонансы возникают при большем раскрытии голосовой щели.

Следует отметить, что в ряде случаев измерения импеданса тканей указывают на еще больший коэффициент затухания $\xi_{\text{т}}$, чем те, которые использовались в наших вычислениях. Это говорит о том, что при определенных условиях можно ожидать появления дополнительных резонансов на еще более ранних стадиях раскрытия голосовой щели. При этом и девиация основной и дополнительной частот первой форманты будут еще значительней.

Итак, в результате анализа описанной довольно простой модели можно заключить, что при определенных соотношениях геометрических параметров и потерь на интервале открытой голосовой щели в акустической системе могут возникать дополнительные резонансы, порождаемые подглоточной областью. Момент появления каждого дополнительного резонанса зависит от степени раскрытия голосовой щели, потерь и соотношения между частотами голосового тракта и подглоточной области. Затухание дополнительных резонансов значительно больше, чем затухание основных резонансов, и уменьшается по мере раскрытия голосовой щели в противоположность затуханию основных резонансов, которое увеличивается при расширении голосовой щели. Мы видели также, что потери в тканях голосовых складок оказывают большое влияние как на момент появления дополнительных резонансов, так и на девиацию частот $F^{(0)}_1$ и $F^{(1)}_1$ в процессе раскрытия голосовой щели, которая может достичь $\pm 10\%$ при $a_{\text{г.щ}}=0,2$ см и даже больше. Эта величина больше чем на порядок превышает оценки девиации частоты первой форманты при раскрытии голосовой щели, приводимые в [В.3, В.4]. Дополнительные резонансы наблюдаются в спектре речевого сигнала в виде максимумов, называемых иногда «ложными формантами».

Характер влияния голосовой щели на частотные параметры

речевого тракта был описан в [4.18]. Аналогичные оценки девиации формантных частот при раскрытии голосовой щели для более адекватной модели речевого тракта в виде трубы с переменным сечением получены в [4.31]. Ряд измерений формантных частот реальных речевых сигналов подтверждает вывод о существовании заметной частотной модуляции на интервале открытой голосовой щели.

При закрытой голосовой щели акустические колебания в речевом тракте замыкаются на импеданс голосовых складок $Z_{г.с.}$, рассчитываемый аналогично (4.17). Импедансом нагрузки Z_H можно пренебречь, в результате чего получим оценку реактивной компоненты импеданса $Z_{г.с.} = R_{г.с.} - j\omega M_{г.с.}$, где $M_{г.с.} \approx 0,008 - 0,016$ при толщине голосовых складок $h_{г.с.} \approx 0,3 - 0,6$ см. Воспользовавшись этой оценкой и данными о потерях в тканях, можно определить влияние импеданса $Z_{г.с.}$ на сдвиг частот однородной трубы, имеющей длину и площадь поперечного сечения такие же, как среднестатистическая длина и площадь поперечного сечения речевого тракта. При этом оказывается, что сдвиг резонансных частот измеряется долями процента, так что в первом приближении податливостью сомкнутых голосовых складок при расчетах частотных характеристик речевого тракта можно пренебречь.

Оценим теперь влияние параметров подсвязочной области на характер граничных условий в области голосовой щели. Средняя величина эквивалентной длины трахеи $l_{тр}$ оказывается близкой к средней длине голосового тракта $l_{г.т} = 17,5$ см, поэтому в расчетах удобно принять такую форму записи: $l_{тр}/l_{г.т} = 1 \pm \epsilon$, $|\epsilon| < 1$. Пренебрегая потерями в выражении (4.75) для эквивалентного импеданса голосовой щели, записываем

$$\bar{Z}_{г.щ} = Z_{г.щ} + \frac{Z_L - j \operatorname{tg} K l_{тр}}{1 - j Z_L \operatorname{tg} K l_{тр}} = Z_{г.щ} + \frac{j - Z_L \operatorname{ctg} K l_{тр}}{j Z_L - \operatorname{ctg} K l_{тр}}.$$

Для открытой трубы $K = (n - 0,5)\pi/l_{г.т}$ и $\operatorname{tg} K l_{тр} = \mp \operatorname{ctg} \epsilon \pi (n - 0,5)$. Если $\epsilon = 0$, то $\bar{Z}_{г.щ} = Z_{г.щ} + 1/Z_L$. Для закрытой трубы $K = (n - 1)\pi/l_{г.т}$ и при $\epsilon = 0$ $\bar{Z}_{г.щ} = Z_{г.щ} + Z_L$. Иными словами, при любом соотношении длины голосового тракта и трахеи изменение выходного сопротивления трубы при ее закрывании меняет вид эквивалентного импеданса голосовой щели. При этом изменяется не только величина, но и характер добавки к $Z_{г.щ}$: вместо инерционной она может стать упругой или наоборот. Изменение длины трахеи $l_{тр}$ также меняет вид эквивалентного импеданса голосовой щели, причем характер изменений зависит и от порядкового номера резонанса.

Из анализа свойств эквивалентного импеданса голосовой щели можно установить несколько видов изменчивости параметров речевого тракта, связанной с импедансом легких. Во-первых, это зависимость эквивалентного импеданса голосовой щели от импеданса излучения. В этом случае поправки на частоту для каждого резонанса изменяются как по величине, так и (при определенных соотношениях между $Z_{г.щ}$ и Z_L) по знаку. Во-вторых, различия в длине трахеи у разных людей также приводят к вариации величин и знаков поправок для каждого резонанса. И наконец, разница в объеме легких, плотности легочной ткани и т. д. создает разброс характеристик самого импеданса Z_L . В результа-

те совместного действия этих факторов может наблюдаться значительное разнообразие в поведении параметров голосового тракта на фазе открытой голосовой щели.

Итак, колебания голосовых складок приводят к изменению потерь и собственных частот речевого тракта. Это явление можно формально изобразить в виде уравнения для временных мод, видоизменив (4.41) с учетом упомянутых переменных параметров:

$$T_k' + g_k(t) [1 + \delta_k(t)] T_k' + c_0^2 \lambda_k^2(t) [1 + \eta_k(t)] T_k = 0. \quad (4.77)$$

Здесь штрих означает, как и раньше, производную по времени $g_k(t)$ — переменное затухание; $\lambda_k(t)$ — медленно меняющиеся собственные числа речевого тракта; $\delta_k(t)$ и $\eta_k(t)$ — изменения затухания и собственных чисел, связанные с раскрытием голосовой щели. Как δ_k , так и η_k являются функциями площади поперечного сечения голосовой щели, т. е. $\delta_k = \delta_k(S_{г.щ}, t)$ и $\eta_k = \eta_k(S_{г.щ}, t)$. Это значит, что δ_k и η_k изменяются синхронно с колебаниями голосовых складок, причем максимальное значение δ_k сравнимо с 1, а максимум η_k может достигать 0,1—0,2. В результате такой синхронности параметры δ_k и η_k в (4.77) заметно изменяются в интервале времени, соответствующем периоду резонансных колебаний речевого тракта (особенно для первой форманты), т. е. по существующей терминологии (4.77) является уравнением с быстро изменяющимися параметрами. Вследствие этого решения (4.77) уже не могут быть представлены в виде суперпозиции гармонических затухающих колебаний. Поскольку параметры δ_k и η_k изменяются почти периодически во времени, (4.77) оказывается родственным уравнению Хилла, свойства которого рассматривались в § 1.4 применительно к описанию переменного натяжения голосовых складок. Особенностью неоднородного уравнения является синхронность изменения параметров δ_k и η_k с возбуждением переменной объемной скоростью воздушного потока, протекающего через голосовую щель.

Периодическое изменение собственных чисел λ_k может привести к возникновению так называемых параметрических резонансов. Возникновение параметрического резонанса наиболее вероятно на частоте, равной половине частоты изменения параметра λ_k , но при соответствующей глубине модуляции η_k могут появиться резонансные колебания и на частоте, равной частоте изменения λ_k , и даже на более высоких частотах. Условия вхождения в параметрический резонанс определяются коэффициентом затухания g_k , при увеличении которого для возникновения резонанса требуется большая глубина модуляции η_k , причем g_k увеличивается синхронно с λ_k .

Эффекты, порождаемые таким своеобразием условий возбуждения акустических колебаний в речевом тракте, пока не изучены, и их влияние на восприятие речи представляет значительный интерес. Вводя девиацию формантных частот и их затуханий синхронно с колебаниями голосовых складок на артикуляторном синтезаторе, мы отмечали некоторое улучшение натуральности звуча-

ния синтетической речи, состоящее в увеличении «живости» синтезированного голоса. В [4.45] переменный импеданс голосовой щели был непосредственно включен в уравнение речевого тракта, в результате чего синтезированные гласные отличались очень высокими натуральностью и разборчивостью.

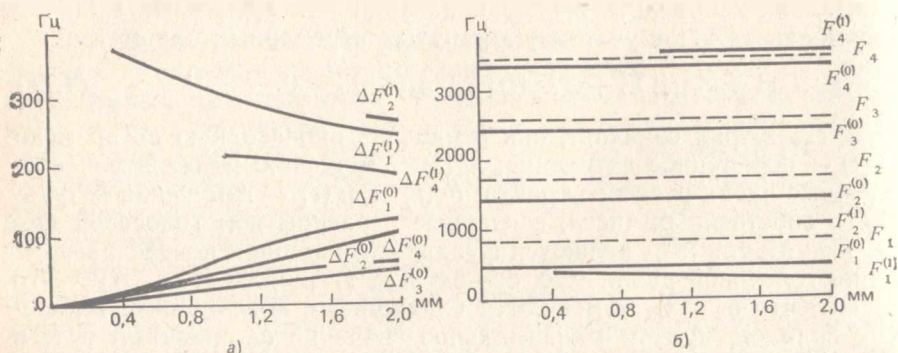


Рис. 4.11. Зависимость ширины резонансов (а) и частоты резонансов и антирезонансов (б) (— — —) от ширины голосовой щели

Рассмотрим теперь влияние воздушного потока, протекающего через голосовую щель, и вертикальных колебаний голосовых складок на процессы возбуждения акустических колебаний в речевом тракте. Акустические возмущения, связанные с движениями голосовых складок, действуют на нижнем конце речевого тракта, поэтому их удобно представить в виде граничных условий, не принимая пока во внимание влияние подсвязочной области, которое рассматривалось выше. Пусть начальная точка системы координат для речевого тракта находится на верхней поверхности голосовых складок: $x=0$. Согласно уравнению Эйлера на границе $x=0$ справедливо соотношение

$$\left. \frac{\partial P}{\partial x} \right|_{x=0} = -\rho_0 \left. \frac{\partial v}{\partial t} \right|_{x=0},$$

где P — давление; v — линейная скорость частиц воздуха; ρ_0 — плотность воздуха; t — время. Если обозначить вертикальное смещение верхней поверхности складок через $h(t)$, то на границе «поверхность складок — воздух» имеем для поршневого источника

$$\left. \frac{\partial P}{\partial x} \right|_{x=0} = -\rho_0 \frac{\partial}{\partial t} \left(\frac{\partial h}{\partial t} \right).$$

Для источника объемной скорости производительностью w объемных единиц в секунду из $w(x, t) = v(x, t)S(x, t)$ и уравнения Эйлера получим

$$\left. \frac{\partial P}{\partial x} \right|_{x=0} = -\rho_0 \frac{\partial}{\partial t} \left(\frac{w_0}{S_0} \right),$$

где w_0 — объемная скорость воздушного потока при выходе из голосовой щели; S_0 — площадь речевого тракта при $x=0$. Таким образом, при $x=0$ граничное условие есть

$$\left. \frac{\partial P}{\partial x} \right|_{x=0} = -\rho_0 \frac{\partial}{\partial t} \left(\frac{w_0}{S_0} + \frac{\partial h}{\partial t} \right) = F(t). \quad (4.78)$$

Граничное условие на верхнем конце речевого тракта при $x=l$ (где l — длина речевого тракта) примем нулевым: $P(l, t)=0$, — пренебрегая тем самым импедансом излучения.

При голосовом возбуждении граничное условие (4.78) на нижнем конце речевого тракта неоднородное. Для того чтобы избавиться от этой неоднородности, перейдем к новой переменной $\tilde{P}(x, t) = P(x, t) - R(x, t)$, где вспомогательная функция $R(x, t) = F(t) \times (x-l)^2/2l$. Тогда после несложных преобразований волновое уравнение (4.29) представляется в виде

$$\frac{1}{S} \frac{\partial}{\partial x} \left(S \frac{\partial \tilde{P}}{\partial x} \right) - \frac{1}{c_0^2} \frac{\partial^2 \tilde{P}}{\partial t^2} \approx \tilde{F}(x, t), \quad (4.79)$$

$$\text{где } \tilde{F}(x, t) = \left(\frac{1}{l} + \frac{1}{S} \frac{\partial S}{\partial x} \frac{x-l}{l} \right) F(t) - \frac{1}{c_0^2} \frac{(x-l)^2}{2l} \frac{\partial^2 F}{\partial t^2}. \quad (4.80)$$

В результате замены переменной для (4.80) изменяются не только граничные условия, которые становятся однородными, но и начальные условия:

$$\begin{aligned} \tilde{P}(x, 0) &= f_1(x) - \frac{(x-l)^2}{2l} F(0); \quad \left. \frac{\partial \tilde{P}(x, t)}{\partial t} \right|_{t=0} = \\ &= f_2(x) - \frac{(x-l)^2}{2l} \frac{\partial F(0)}{\partial t}, \end{aligned}$$

где $f_1(x)$ и $f_2(x)$ — начальные условия для (4.29). После решения (4.79) возвращаемся к старой переменной $P(x, t)$.

Из (4.80) видно, что импульс возбуждения речевого тракта содержит три компоненты, тогда как до сих пор считалось, что акустические колебания в речевом тракте возбуждаются лишь сигналом, пропорциональным производной от объемной скорости потока через голосовую щель по времени, что соответствует первому члену в (4.80). Второй член уравнения (4.80), не изменяя характера возбуждения во времени, влияет на амплитудные соотношения резонансных колебаний в соответствии со степенью гладкости функции площади поперечного сечения речевого тракта. Третий член вносит наибольший вклад в возбуждение в моменты смыкания и размыкания голосовых складок. Возможно, что именно этот член является причиной дополнительного импульса возбуждения, наблюдаемого иногда в моменты, предшествующие схлопыванию складок.

Точки верхней поверхности голосовых складок совершают несинфазное движение. Учитывая отношение максимального поперечного размера речевого тракта к его длине, это несинфазное дви-

жение можно заменить эквивалентным движением плоского поршня вплоть до частот порядка 5 кГц [4.13]. Такая замена существенно упрощает анализ количественных соотношений между вкладом источника объемной скорости (связанного с попаданием некоторой массы воздуха из легких в речевой тракт) и поршневого источника (вертикальных движений голосовых складок) в возбуждение акустических колебаний в речевом тракте.

Оценка роли поршневого источника возбуждения производилась путем моделирования на артикуляторном синтезаторе ИППИ АН СССР [4.3]. Аудиторам предъявлялись для сравнительной оценки односложные сочетания гласных и звукосочетания с назальными согласными типа *МАМА*, *НИНА* и т. д. Аудиторская оценка определенно показывает, что подключение поршневого источника изменяет тембр звучания синтетической речи и улучшает его натуральность. Соотношение объемной скорости на фазах открытой и закрытой голосовых щелей в этой модели получилось близким к тому, которое можно усмотреть на натуральной речи при использовании метода обратной фильтрации импульсов возбуждения. Это означает, что для акустических колебаний в речевом тракте может не существовать режима свободных затухающих колебаний, в расчете на который строятся многие алгоритмы анализа речевых сиг-

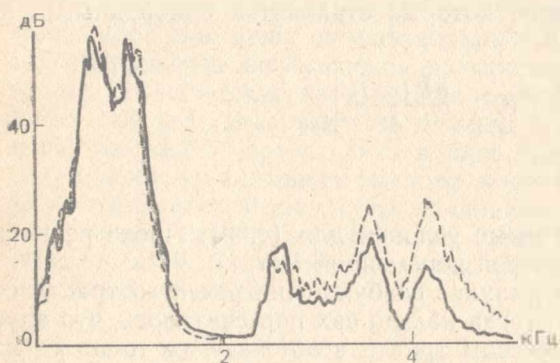


Рис. 4.12. Спектр гласного *А*: — источник объемной скорости; --- источник объемной скорости и поршневой источник

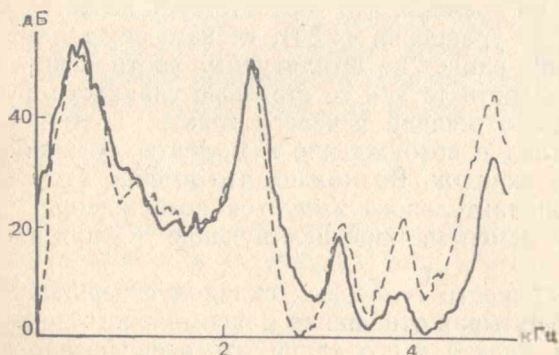


Рис. 4.13. Спектр гласного *И*: — источник объемной скорости; --- источник объемной скорости и поршневой источник

налов. В частотной области влияние поршневого источника проявляется в подъеме верхних частот и изменении амплитудных соотношений между формантами (рис. 4.12 и 4.13).

Скорость потока, создаваемая поршневым источником, может превысить пороговое значение числа Рейнольдса, вызывая возникновение турбулентных шумов, что соответствует наблюдениям [2.82] о шумах на интервале закрытой голосовой щели. При увеличении амплитуды вертикальных движений голосовых складок шумовая составляющая также увеличивается, придавая голосу хриловатый, резкий оттенок.

Роль голосового источника не ограничивается возбуждением акустических колебаний в речевом тракте. Масса воздуха, поступающая в речевой тракт из легких на интервале открытой голосовой щели, должна его покинуть. Объемная скорость этой массы воздуха у губ та же, что и в голосовой щели, вследствие несжимаемости газа: $w_l = w_0$. Отсюда линейная скорость потока у губ $v_l = w_l / S_l$, где S_l — площадь губного отверстия. Давление, создаваемое этим потоком, $P_l = c_x \frac{\rho_0 v_l^2}{2}$ порождает аддитивную низкочастотную компоненту в речевом сигнале, форма которой изменяется во времени синхронно с площадью голосовой щели. Аналогично смещение голосовых складок вдоль оси речевого тракта приводит к вытеснению некоторого объема воздуха из речевого тракта и создает поток через губное отверстие. В спектре речевого сигнала это явление отображается подъемом низкочастотной области и никак не сказывается на высоких частотах (рис. 4.14). Добавление аддитивной компоненты в синтезированный речевой сигнал улучшает натуральность его звучания, делая его более сочным и глубоким.

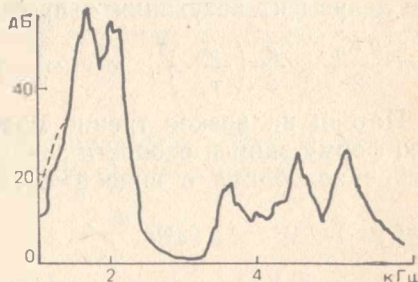


Рис. 4.14. Спектр гласного А без учета аэродинамического потока (—) и с его учетом (---)

4.4. ПОТЕРИ В РЕЧЕВОМ ТРАКТЕ

Распространение звука в воздушной среде сопровождается потерями, вызванными различными причинами: вязкостью, теплопроводностью, диссипативностью и т. д. Вязкое трение порождается различиями в скоростях слоев газа и пропорционально скорости сдвиговой деформации частиц. Обозначая смещение частиц воздуха через ξ , записываем уравнение движения в плоской волне

$$\rho_0 \frac{\partial^2 \xi}{\partial t^2} = - \frac{\partial P}{\partial x},$$

где P — суммарное давление: $P = P_0 - \kappa \frac{\partial \xi}{\partial x} - r \frac{\partial}{\partial t} \left(\frac{\partial \xi}{\partial x} \right)$; P_0 — статическое атмосферное давление; κ — модуль объемной упругости газа; r — коэффициент потерь; ρ_0 — плотность воздуха.

Член $\kappa \frac{\partial \xi}{\partial x}$ представляет собой изменение давления вследствие распространения звуковой волны, а член $r \frac{\partial}{\partial t} \left(\frac{\partial \xi}{\partial x} \right)$ — дополнительное давление, возникающее при сдвиге слоев воздуха. В результате волновое уравнение для смещений принимает вид

$$\frac{\partial^2 \xi}{\partial t^2} = c_0^2 \frac{\partial^2 \xi}{\partial x^2} + g \frac{\partial^3 \xi}{\partial x^2 \partial t}, \quad (4.81)$$

где скорость звука $c_0 = \sqrt{\kappa/\rho_0}$; $g = r/\rho_0$. Принимая свободные колебания в виде $\xi = A(t) \cos K_0 x$, вместо (4.81) получаем уравнение

$$\frac{d^2 A}{dt^2} + \alpha \frac{dA}{dt} + (K_0 c_0)^2 A = 0,$$

где $K_0 = \omega_0/c_0$ — волновое число; ω_0 — круговая частота. Решение этого уравнения есть

$$A(t) = C_A e^{-\alpha t} \cos(\omega t + \theta),$$

где декремент затухания α пропорционален квадрату частоты:

$$\alpha = \frac{g K_0^2}{2} = \frac{g}{2} \left(\frac{\omega_0}{c_0} \right)^2, \quad \omega^2 = \omega_0^2 - g^2.$$

Потери на вязкое трение можно представить через комплексную форму записи скорости звука. Для этого представим гармонические колебания в виде $\xi = X(x) e^{-j\omega t}$, подставим в (4.81) и из

$$-\omega^2 \rho_0 X = (\kappa - j g \rho_0 \omega) \frac{d^2 X}{dx^2}$$

$$\text{получим } \frac{d^2 X}{dx^2} + \frac{\omega^2}{c_0^2 \left(1 - j \frac{g \omega}{c_0^2} \right)} X = 0,$$

где комплексная скорость звука

$$c^* = c_0 \sqrt{1 - j \frac{g \omega}{c_0^2}} \approx c_0 \left(1 - j \frac{g \omega}{2 c_0^2} \right).$$

Рассматривая только сдвиговые напряжения, Максвелл принял коэффициент потерь $g = \frac{4}{3} \mu/\rho_0$, где μ — коэффициент вязкого сопротивления при сдвиге ($\mu \approx 1,86 \cdot 10^{-4}$ г/см·с при 20° С, 760 мм·рт. ст.); $\rho_0 = 1,14 \cdot 10^{-3}$ г/см³ при 37° С) [4.9]. Отношение $\mu/\rho_0 = \mu'$ называется кинематической вязкостью ($\mu' \approx 0,16$). Учет всестороннего сжатия среды увеличивает затухание на коэффициент объемной вязкости ξ : $g_1 = \frac{4}{3} \frac{\mu}{\rho_0} + \xi$.

Потери на теплопроводность возникают из-за выравнивания температур в зонах с повышенным и пониженным давлением. Эти потери пропорциональны звуковому давлению, их зависимость от частоты, плотности и скорости звука такая же, как и у вязких потерь, поэтому часто суммарный коэффициент потерь рассматривают в виде

$$g_2 = \frac{4}{3} \frac{\mu}{\rho_0} + \xi + \xi_t \left(\frac{1}{c_V} - \frac{1}{c_P} \right),$$

где ξ_t — коэффициент теплопроводности; c_V — теплоемкость воздуха при постоянном объеме: $c_V \approx c_P/1,4$; c_P — теплоемкость воздуха при постоянном давлении ($c_P = 1005$ Дж/кг·град при 0°C и 1 атм) [4.7]. Потери на теплопроводность того же порядка, что и потери на вязкость, и в диапазоне 300—3000 Гц декремент затухания находится в пределах 0,001—0,1. Эти потери весьма малы, так что в помещениях потери на стенках преобладают над потерями в среде до частот порядка 2000 Гц [4.7].

В узких трубах потери на вязкое трение значительно выше, чем аналогичные потери в свободном пространстве, вследствие большего градиента скоростей между слоями воздуха. Этот градиент создается торможением частиц воздуха у стенок из-за молекулярного взаимодействия. Обратим задачу и рассмотрим не движение воздуха относительно стенок, а движение стенок в покоящейся среде. Уравнение диффузии вдоль нормали к стенке y есть

$$\mu \frac{\partial}{\partial y} \left(\frac{\partial v}{\partial y} \right) = \rho_0 \frac{\partial v}{\partial t}.$$

Это уравнение было впервые решено Кирхгофом. Для гармонических колебаний стенки вдоль оси x $v = V(y)e^{-j\omega t}$ имеем

$$\frac{d^2 V}{dy^2} = -\frac{j\omega\rho_0}{\mu} V = (1-j)^2 \frac{\omega\rho_0}{2\mu} V$$

и решение этого уравнения следующее:

$$V(y) = V_m e^{-\sqrt{\frac{\omega\rho_0}{2\mu}} y} e^{j\sqrt{\frac{\omega\rho_0}{2\mu}} y},$$

где V_m — максимальная скорость вдали от стенок.

Расстояние Δ , на котором скорость частиц V уменьшается в e раз по сравнению со скоростью стенки, называется толщиной пограничного слоя или слоем Прандтля, о котором уже упоминалось в предыдущих разделах.

Значения Δ для некоторых частот приведены в табл. 4.3.

Таблица 4.3

Частота, Гц	100	500	1000	1500	2500	3500
Δ , мм	0,22	0,1	0,07	0,06	0,045	0,04

Рэлеи получил уравнение, в котором потери на стенках выражаются как функция от отношения периметра стенок к площади поперечного сечения:

$$\frac{\partial^2 \omega}{\partial t^2} \left(1 + \frac{L}{S} \sqrt{\frac{\mu}{2 \omega \rho_0}} \right) + \frac{L}{S} \sqrt{\frac{\omega \mu}{\rho_0}} \frac{\partial \omega}{\partial t} = c_0^2 \frac{\partial^2 \omega}{\partial x^2},$$

где ω — полный поток через сечение; L — периметр трубы.

Усредненный по сечению импеданс круглой трубы с радиусом a_0 по [2.5]

$$Z = R - jX = \frac{j \omega \rho_0 l}{1 - \frac{2J_1(K' a_0)}{K' a_0 J_0(K' a_0)}},$$

где J_0, J_1 — функции Бесселя; $K' = \sqrt{\omega \rho_0 / j \mu}$.

Если соблюдается условие $|K' a_0| < 2$, т. е. $a_0 < 2\Delta$, то удельный импеданс

$$\bar{Z} \approx \frac{8 \mu l}{a_0^2} - j \frac{4}{3} \rho_0 \omega l.$$

Активное удельное сопротивление в этом случае описывается законом Пуазейля для капиллярных трубок $\bar{R}_\pi = 8 \mu l / a_0^2$. Для капиллярных трубок эллиптического сечения $\bar{R}_{эл} = 4 \mu' (a^2 + b^2) / a^2 b^2$, где μ' — динамический коэффициент вязкости; a и b — полуоси эллипса. Полное сопротивление для капиллярных круглых трубок, таким образом, не зависит ни от частоты, ни от площади:

$$R_\pi = \bar{R}_\pi S = 8 \pi \mu l.$$

Для больших a_0 и высоких частот ($|K' a_0| \geq 10$, т. е. $a_0 > 10\Delta$) можно приближенно считать $J_1(x \sqrt{-j}) / J_0(x \sqrt{-j}) \approx -j$ и

$$\bar{Z} \approx l \left[\frac{\sqrt{2 \rho_0 \mu \omega}}{a_0} - j \omega \rho_0 \left(1 + \frac{\Delta}{a} \right) \right]. \quad (4.82)$$

Активное сопротивление можно представить и в других формах:

$$R = \bar{R} S = \frac{\pi l a_0^2 \sqrt{2 \rho_0 \mu \omega}}{a_0} = \pi l a_0 \sqrt{2 \rho_0 \mu \omega} = L l \sqrt{\frac{\rho_0 \mu \omega}{2}},$$

где L — периметр кругового сечения; l — длина трубки.

Последнее решение получено Гельмгольцем в 1863 г. Вводя так называемый коэффициент формы $\kappa_\Phi = L / 2 \sqrt{\pi S}$, записываем удельное сопротивление

$$\bar{R} = \frac{\kappa_\Phi}{\sqrt{S}} \sqrt{2 \pi \rho_0 \mu \omega} = \frac{2 \pi \kappa_\Phi}{\sqrt{S}} \sqrt{\rho_0 \mu f}, \quad \omega = 2 \pi f.$$

Для прямоугольного сечения со сторонами a и b удобнее пользоваться отношением L/S , так как при больших a/b $L/S \approx \text{const}$ и $\bar{R} = L \sqrt{\pi \rho_0 \mu f} / S$. Для узких прямоугольных щелей, у которых длина гораздо больше ширины, импеданс

$$\bar{Z} = \frac{j \omega \rho_0 l}{1 - \operatorname{tg} \frac{K'_n b}{2} \left/ \frac{K'_n b}{2} \right.}, \quad (4.83)$$

где $K'_n = \sqrt{j \omega \rho_0 / \mu}$. При $K'_n b < 1$, т. е. $b < \Delta \sqrt{2}$, получаем импеданс для капиллярной трубки с прямоугольным сечением

$$\bar{Z} = \frac{12 \mu l}{b^2} - j \frac{6}{5} \omega \rho_0 l.$$

Отношение активного сопротивления при $a_0 > 10\Delta$ и $b > 20\Delta$ к сопротивлению капиллярной трубки (при $a_0 < 2\Delta$, $b < \Delta$) для круглого и прямоугольного сечений было рассчитано в [2.5]. В диапазоне 100—4000 Гц оно изменяется в несколько раз. Форма поперечного сечения речевого тракта изменяется примерно от круглой и до прямоугольной, и в большинстве артикуляционных позиций выполняются условия $a_0 > 10\Delta$, $b > 20\Delta$. Следовательно, потери на трение у стенок в речевом тракте находятся в диапазоне значений, вычисляемых по (4.82) и (4.83).

Потери энергии в трубах, связанные с ненулевой теплопроводностью стенок, также значительно превышают потери на теплопроводность в свободной среде. Уравнение для распределения температуры по нормали к стенке аналогично уравнению для распределения нормальных скоростей:

$$\frac{\partial^2 T}{\partial y^2} = \frac{c_P \rho_0}{\kappa_{т.п}} \frac{\partial T}{\partial t},$$

где T — температура; $\kappa_{т.п}$ — коэффициент теплопроводности, ($\kappa_{т.п} = 23,03 \cdot 10^{-3}$ Вт/м·град при 0°C). Определив глубину теплового пограничного слоя как $\Delta_{т.п} = \sqrt{2\kappa_{т.п}/\rho_0 c_P \omega}$, можно представить нормальную составляющую скорости потока тепла у стенки в виде

$$v_{y,т.п} = P \frac{\omega}{c_0} \frac{\eta - 1}{\rho_0 c_0} \frac{j}{1 + j} \Delta_{т.п},$$

а эквивалентную проводимость на единицу длины трубы, обусловленную теплопроводностью, как

$$Y_{т.п} = L \frac{\eta - 1}{\rho_0 c_0^2} \sqrt{\frac{\kappa_{т.п} \omega}{2 \rho_0 c_P}},$$

где $\eta = c_P / c_V = 1,4$, L — периметр трубы [В.4]. Коэффициент потерь на теплопроводность $R_{т.п}$ имеет такую же зависимость от частоты, что и коэффициент потерь на вязкость R , причем по [В.3] $R_{т.п} \approx 0,45R$, поэтому потери на теплопроводность обычно учитывают как кажущимся увеличением вязкости.

Еще один механизм потерь, пропорциональных скорости смещения воздушных частиц, связан с испарением жидкости с влажных стенок речевого тракта, так что в действительности эквивалентный коэффициент вязкого трения оказывается еще большим.

В трубах с открытыми концами происходят потери колеба-

тельной энергии также вследствие излучения сигнала. Как было показано в § 4.2, потери на излучение из ротового отверстия примерно пропорциональны квадрату частоты и площади излучения. Потери на излучение в подвязочную область на интервале открытой голосовой щели рассматривались в § 4.3, причем выяснилось, что потери на сомкнутых голосовых складках пренебрежимо малы. Измерения амплитудно-частотных характеристик речевого тракта с использованием внешнего генератора скользящей частоты показали, что в низкочастотной области ширина формант, а следовательно, и потери увеличиваются с уменьшением частоты [4.35]. Это явление в дальнейшем было связано с потерями энергии на колебания стенок речевого тракта [4.29]. Как было показано в § 4.1, податливость стенок приводит к потерям, которые могут быть представлены в виде комплексной компоненты скорости звука [см. (4.22)]. Подставляя в это выражение различные значения измеренных импедансов тканей, можно оценить потери на колебания

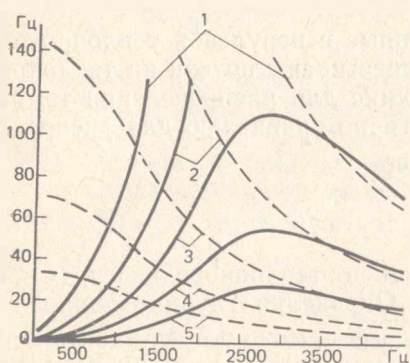


Рис. 4.15. Ширина полосы резонансов в однородной трубе с прямоугольным сечением для упруго-инерционного импеданса (—) и инерционного импеданса (---). Площадь трубы (см^2): 1—0,5; 2—1; 3—2; 4—4; 5—8

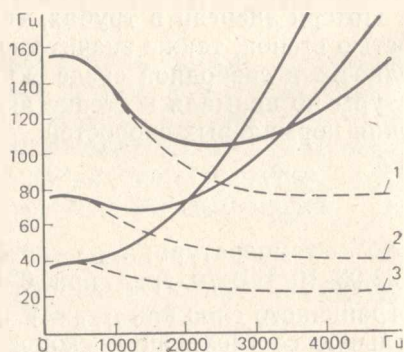


Рис. 4.17. Ширина полосы резонансов для открытой (—) и закрытой (---) однородных труб с прямоугольным сечением шириной 5 см, площадь 5 см^2 . Площадь излучения: 1—1 см^2 ; 2—2 см^2 ; 3—4 см^2

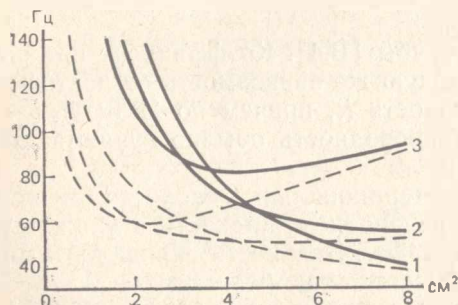


Рис. 4.16. Ширина полосы резонансов для круглого (---) и прямоугольного (—) сечений однородной трубы на частотах: 1—500 Гц; 2—1500 Гц; 3—2500 Гц

мягких и жестких участков стенок речевого тракта. Форма поперечного сечения трубы также влияет на характеристики податливости стенок [4.17]. На рис. 4.15 и 4.16 показано, как на характер зависимости ширины полосы резонансов ΔF от частоты влияют тип импеданса стенок и форма поперечного сечения. Максимум потерь приходится на разные частоты для импедансов разного типа. Частотная зависимость суммарных потерь на излучение, податливость стенок, вязкое трение и теплопроводность показана на рис. 4.17.

Как мы видели, все виды потерь в речевом тракте, включая потери на трение и податливость стенок, могут быть описаны с помощью комплексной скорости звука, форма волнового уравнения при этом сохраняется в неизменном виде:

$$\frac{\partial^2 \xi}{dx^2} = \frac{1}{c_0^2 (1 - j\alpha)} \frac{\partial^2 \xi}{\partial t^2}.$$

Решение такого уравнения, как известно, содержит экспоненциально затухающие колебания.

Это решение, однако, справедливо лишь для однородной трубы, в которой коэффициент потерь не зависит от пространственной координаты. В речевом же тракте неоднородность механических свойств стенок и переменная площадь поперечного сечения приводят к появлению переменного коэффициента потерь в уравнении речевого тракта:

$$\kappa \frac{\partial}{\partial x} \left(S \frac{\partial \xi}{\partial x} \right) = \rho_0 \frac{\partial}{\partial t} \left(S \frac{\partial \xi}{\partial t} \right) + Q(x, t) \frac{\partial \xi}{\partial t}. \quad (4.84)$$

Этот член с коэффициентом $Q(x, t)$ затрудняет решение уравнения (4.84), поэтому был разработан метод вычисления такого эквивалентного постоянного коэффициента затухания, который позволяет по решению уравнения Вебстера без потерь (4.36) найти решение (4.84) в виде

$$\xi(x, t) = \sum_{k=1}^{\infty} e^{-\delta_k t} \psi_k(x) \varphi_k(t), \quad (4.85)$$

где $\psi_k(x)$ и $\varphi_k(t)$ — пространственные и временные моды уравнения Вебстера.

В [4.17] эквивалентный постоянный коэффициент затухания определяется из условия, что на периоде собственных колебаний T_k рассеивается та же энергия, что и для распределенных потерь $Q(x, t)$. Работа сил сопротивления на периоде T_k равна изменению кинетической энергии $W_{\text{кин}} = \Delta E_{\text{кин}} = E_{\text{кин}}(t) - E_{\text{кин}}(t - T_k)$, где кинетическая энергия воздушного слоя в поперечном сечении для уравнения Вебстера

$$E_{\text{кин}} = \frac{\rho_0}{2} \int_0^l S \left(\frac{\partial \xi}{\partial t} \right)^2 dx.$$

Представим временные моды в (4.85) как $\overline{\varphi}_k(t) = \varphi_k(t) e^{-\delta_k t} = A_k \cos(\omega_k t + \theta_k) e^{-\delta_k t}$. Допустим, что наличие затухания существ-

венно не изменяет частоты собственных колебаний, и определим изменение кинетической энергии для k -го собственного колебания между ее максимальными значениями. Тогда из условия $\overline{\varphi'_k} = \max$ найдем фазу θ_k , которая при $t=0$ равна $\pi/2$. Отсюда имеем:

$$= 0; \quad \overline{\varphi'_k}^2 = \omega_k^2; \quad t = T_k; \quad \overline{\varphi'_k}^2 = e^{-2\delta_k T_k} \omega_k^2 = e^{-\frac{4\pi\delta_k}{\omega_k}} \omega_k^2.$$

Таким образом, изменение кинетической энергии для k -го собственного колебания

$$\Delta E_{\text{кин}} = \frac{\rho_0 A_k \omega_k^2}{2} \left(1 - e^{-\frac{4\pi\delta_k}{\omega_k}} \right) \int_0^l S \psi_k^2 dx.$$

Сила сопротивления определяется как $F_c = QS \partial \xi / \partial t$, а ее работа в периоде T_k

$$W_{\text{кин}} = \int_0^l \int_0^{T_k} F_c d\xi dx = \int_0^l \int_0^{T_k} QS \left(\frac{\partial \xi}{\partial t} \right)^2 dt dx.$$

Вследствие медленного изменения конфигурации речевого тракта можно принять, что за период T_k его площадь S остается постоянной. Тогда для k -го собственного колебания работа

$$W_{\text{кин}} \approx \int_0^l QS \psi_k^2 dx \int_0^{T_k} \overline{\varphi'_k}^2 dt.$$

Проведя необходимые преобразования и приравняв работу $W_{\text{кин}}$ изменению кинетической энергии $\Delta E_{\text{кин}}$, получим

$$\omega_k \frac{4 + \left(\frac{2\delta_k}{\omega_k} \right)^2}{\delta_k - \frac{\omega_k}{\delta_k} - 1} = \frac{\int_0^l QS \psi_k^2 dx}{\frac{\rho_0}{2} \int_0^l S \psi_k^2 dx}.$$

Если $(\delta_k/\omega_k)^2 \approx 0$, а отношением δ_k/ω_k пренебречь нельзя,

$$k = \frac{\omega_k I_k}{4\omega_k + I_k},$$

$$\text{где } I_k = \frac{\int_0^l QS \psi_k^2 dx}{\frac{\rho_0}{2} \int_0^l S \psi_k^2 dx}.$$

Если же $\delta_k/\omega_k \approx 0$, то

$$\delta_k = \frac{\int_0^l QS \psi_k^2 dx}{2\rho_0 \int_0^l S \psi_k^2 dx}. \quad (4.86)$$

Таким образом, мы получили изменение показателя затухания во времени как функцию артикуляторных движений, но избавились от пространственной зависимости. При $S = \text{const}$ выражение (4.86) переходит в известную формулу для потерь в однородной трубе $\delta_k = Q/2\rho_0$. Выражение (4.86) позволяет избежать трудностей в решении уравнения речевого тракта (4.84), заменяя его решением уравнения Вебстера без потерь и внося соответствующие поправки на затухание в виде экспоненты с показателем δ_k . Кроме того, (4.86) поясняет зависимость между затуханием и формой собственных функций речевого тракта, которая была наглядно продемонстрирована в [4.51].

4.5. НАЗАЛИЗАЦИЯ

Как было показано в § 3.3, существует довольно сложная временная организация процессов артикуляции назальных звуков, роль которой в формировании акустических характеристик речевого сигнала и влияние на субъективное восприятие эффекта назализации до сих пор почти не исследованы, в отличие от влияния постоянно открытого прохода в носовую полость на характеристики согласных M , N (как, например, для русского языка [В.3] и французских назальных гласных [4.49]). Опускание небной занавески приводит к разветвлению акустической системы речевого тракта, изменяя его амплитудно-частотные характеристики. Моделирование процессов назализации с учетом динамики артикуляции было описано в [4.1].

Рассмотрим сначала разветвленную акустическую систему (рис. 4.18), состоящую из однородных труб с длинами l_0, l_1, l_2 и площадями поперечного сечения S_0, S_1, S_2 соответственно. Площади излучающих отверстий обозначим как S_{1l} и S_{2l} .

Введем независимые координаты x_0, x_1, x_2 для каждой трубы и запишем систему волновых уравнений относительно давления:

$$\frac{\partial^2 P_0}{\partial x_0^2} - \frac{1}{c_0^2} \frac{\partial^2 P_0}{\partial t^2} = 0, \quad 0 \leq x_0 \leq l_0; \quad (4.87)$$

$$\frac{\partial^2 P_1}{\partial x_1^2} - \frac{1}{c_0^2} \frac{\partial^2 P_1}{\partial t^2} = 0, \quad 0 \leq x_1 \leq l_1; \quad (4.88)$$

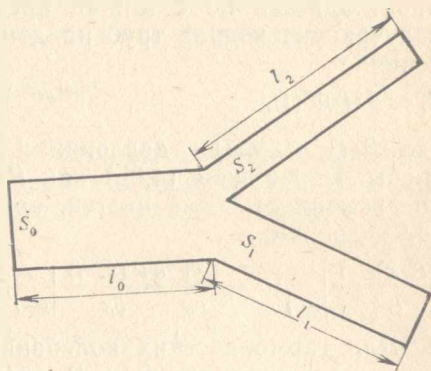


Рис. 4.18. Модель нулевого приближения для разветвленного речевого тракта

$$\frac{\partial^2 P_2}{\partial x_2^2} - \frac{1}{c_0^2} \frac{\partial^2 P_2}{\partial t^2} = 0, \quad 0 \leq x_2 \leq l_2, \quad (4.89)$$

где P_0 , P_1 и P_2 — акустические давления соответственно в каждой трубе; c_0 — скорость звука в воздухе. Примем граничные условия на задней стенке при $x=0$ как абсолютно жесткие, т. е.

$$\partial P_0(x_0, t)/\partial x_0|_{x_0=0} = 0. \quad (4.90)$$

Граничные условия в точке разветвления найдем из условия равенства давлений и неразрывности потока:

$$P_0(l_0, t) = P_1(0, t) = P_2(0, t); \quad V_0(l_0, t) S_0 = V_1(0, t) S_1 + V_2(0, t) S_2, \quad (4.91); \quad (4.92)$$

где V_0 , V_1 и V_2 — линейные скорости акустических колебаний. Воспользовавшись уравнением Эйлера

$$\frac{P}{\partial x} = -\rho_0 \frac{\partial V}{\partial t}, \quad (4.93)$$

где ρ_0 — плотность воздуха, перейдем от (4.92) к

$$\frac{\partial P_0(x_0, t)}{\partial x_0} S_0 \Big|_{x_0=l_0} = \frac{\partial P_1(x_1, t)}{\partial x_1} S_1 \Big|_{x_1=0} + \frac{\partial P_2(x_2, t)}{\partial x_2} S_2 \Big|_{x_2=0}, \quad (4.94)$$

предполагая, что S_0 , S_1 и S_2 медленно меняются во времени и их производными по t можно пренебречь. Граничные условия на излучающих концах труб найдем из определения импеданса излучения:

$$Z_l = P(l)/V(l), \quad (4.95)$$

где $P(l)$ и $V(l)$ — давление и скорость на излучающем конце трубы. Перепишем (4.95) как $P(l) = Z_l V(l)$, возьмем производные по времени от обеих частей, примем $\partial Z_l / \partial t \approx 0$ и, снова учитывая (4.93), получим

$$\frac{\partial P(x, t)}{\partial x} \Big|_{x=l} = -\frac{Z_l}{\rho_0} \frac{\partial P(x, t)}{\partial x} \Big|_{x=l}. \quad (4.96)$$

Для гармонических колебаний $e^{-j\omega t}$ пространственная компонента $p(x)$ давления $P(x, t) = p(x) e^{-j\omega t}$ на излучающем конце может быть записана как

$$Z_l \frac{\partial p(x)}{\partial x} \Big|_{x=l} - j\omega \rho_0 p(l) = 0, \quad (4.97)$$

где согласно § 4.2 импеданс излучения

$$Z_l = \rho_0 c_0 \frac{S}{S_l} \left[\frac{(\lambda a)^2}{2} - j \frac{8 \lambda a}{3\pi} \right],$$

где $a = \sqrt{S_l/\pi}$; $\lambda = \omega/c_0$; ω — круговая частота. Принимая во внимание только потери на излучение и пренебрегая влиянием импеданса излучения на резонансные частоты, получаем граничные условия для (4.88) и (4.89):

$$\left. \frac{\partial P_1(x_1, t)}{\partial x_1} \right|_{x_1=l_1} + A_1 P_1(l_1, t) = 0; \quad (4.98)$$

$$\left. \frac{\partial P_2(x_2, t)}{\partial x_2} \right|_{x_2=l_2} + A_2 P_2(l_2, t) = 0, \quad (4.99)$$

$$\text{где } A_1 = \frac{3\pi^2 a_1}{8S_1}; \quad A_2 = \frac{3\pi^2 a_2}{8S_2}; \quad a_1 = \sqrt{\frac{S_{1l}}{\pi}}; \quad a_2 = \sqrt{\frac{S_{2l}}{\pi}}.$$

Применяя метод разделения переменных и представляя $P(x, t) = p(x)T(t)$, получаем из (4.87) — (4.89) с учетом граничных условий (4.90), (4.91), (4.94), (4.98) и (4.99) следующую систему решений:

$$\psi_{0k}(x_0) = C_k \cos \lambda_k x_0; \quad (4.100)$$

$$\psi_{1k}(x_1) = C_k \cos \lambda_k l_0 \left(\cos \lambda_k x_1 + \frac{\lambda_k \cos \lambda_k l_1 - A_1}{A_1 \cos \lambda_k l_1 + \lambda_k} \sin \lambda_k x_1 \right); \quad (4.101)$$

$$\psi_{2k}(x_2) = C_k \cos \lambda_k l_0 \left(\cos \lambda_k x_2 + \frac{\lambda_k \cos \lambda_k l_2 - A_2}{A_2 \cos \lambda_k l_2 + \lambda_k} \sin \lambda_k x_2 \right), \quad (4.102)$$

где ψ_{0k} , ψ_{1k} и ψ_{2k} — собственные функции для каждой трубы, а λ_k — собственные числа разветвленной акустической системы, которые определяются из трансцендентного уравнения ($k=1, 2, \dots$)

$$S_0 \operatorname{tg} \lambda_k l_0 + S_1 \frac{\lambda_k \operatorname{tg} \lambda_k l_1 - A_1}{A_1 \operatorname{tg} \lambda_k l_1 + \lambda_k} + S_2 \frac{\lambda_k \operatorname{tg} \lambda_k l_2 - A_2}{A_2 \operatorname{tg} \lambda_k l_2 + \lambda_k} = 0. \quad (4.103)$$

Нормирующие коэффициенты C_k находим из условия

$$\int_0^{l_0} \psi_{0k}^2(x_0) dx_0 + \int_0^{l_1} \psi_{0k}^2(x_1) dx_1 + \int_0^{l_2} \psi_{0k}^2(x_2) dx_2 = 1. \quad (4.104)$$

Обращаясь к речевому тракту, в котором площадь поперечного сечения зависит от пространственной координаты, запишем для случая подсоединенной носовой полости уравнения Вебстера:

$$\frac{1}{S^{(0)}} \frac{\partial}{\partial x_0} \left[S^{(0)} \frac{\partial P_0}{\partial x_0} \right] - \frac{1}{c_0^2} \frac{\partial^2 P_0}{\partial t^2} = 0, \quad 0 \leq x_0 \leq l_0; \quad (4.105)$$

$$\frac{1}{S^{(1)}} \frac{\partial}{\partial x_1} \left[S^{(1)} \frac{\partial P_1}{\partial x_1} \right] - \frac{1}{c_0^2} \frac{\partial^2 P_1}{\partial t^2} = 0, \quad 0 \leq x_1 \leq l_1; \quad (4.106)$$

$$\frac{1}{S^{(2)}} \frac{\partial}{\partial x_2} \left[S^{(2)} \frac{\partial P_2}{\partial x_2} \right] - \frac{1}{c_0^2} \frac{\partial^2 P_2}{\partial t^2} = 0, \quad 0 \leq x_2 \leq l_2. \quad (4.107)$$

Здесь l_0 — расстояние от голосовых складок до небной занавески, измеренное по средней линии речевого тракта; l_1 — расстояние от небной занавески до губ; l_2 — длина носовой полости от небной занавески до ноздрей. Площадь поперечного сечения носовой полости $S^{(2)}(x_2)$ считается равной удвоенной площади одного из носовых каналов, т. е., как обычно, предполагается их полная идентичность.

Потери в носовой полости, как и в речевом тракте, распределены по его длине, однако, воспользовавшись методом, описан-

ным в предыдущем параграфе, можно перейти к эквивалентному постоянному показателю затухания и, произведя замену переменных, свести исходные волновые уравнения к виду (4.105) — (4.107). Для расчета параметров системы нулевого приближения, состоящей из однородных труб, проще всего принять

$$S_i = \frac{1}{l_i} \int_0^{l_i} S^{(i)}(x_i) dx_i \quad i = 0, 1, 2.$$

Разделим переменные в (4.105) — (4.107) и запишем уравнения относительно пространственных компонент:

$$(S^{(0)} P_0')' + \lambda^2 S^{(0)} P_0 = 0; \quad (4.108)$$

$$(S^{(1)} P_1')' + \lambda^2 S^{(1)} P_1 = 0; \quad (4.109)$$

$$(S^{(2)} P_2')' + \lambda^2 S^{(2)} P_2 = 0, \quad (4.110)$$

где штрих означает производную по пространственной координате. Следуя методу Галеркина, представляем пространственную компоненту акустического давления на каждом из участков речевого тракта как взвешенную сумму собственных функций ψ_{ik} системы нулевого приближения (4.100) — (4.102):

$$P_0(x_0) = \sum_{k=1}^N a_k \psi_{0k}(x_0); \quad (4.111)$$

$$P_1(x_1) = \sum_{k=1}^N a_k \psi_{1k}(x_1); \quad (4.112)$$

$$P_2(x_2) = \sum_{k=1}^N a_k \psi_{2k}(x_2), \quad (4.113)$$

где N — порядок приближения. Потребуем затем ортогональности каждого уравнения (4.108) — (4.110) к системе соответствующих функций нулевого приближения $\{\psi_0\}$, $\{\psi_1\}$ и $\{\psi_2\}$, для чего умножим (4.108) на ψ_{0m} , (4.109) на ψ_{1m} , (4.110) на ψ_{2m} и возьмем интегралы вдоль каждого участка, получив при этом

$$\begin{aligned} & \int_0^{l_0} [(S^{(0)} P_0')' + \lambda^2 S^{(0)} P_0] \psi_{0m} dx_0 + \int_0^{l_1} [(S^{(1)} P_1')' + \lambda^2 S^{(1)} P_1] \psi_{1m} dx_1 + \\ & + \int_0^{l_2} [(S^{(2)} P_2')' + \lambda^2 S^{(2)} P_2] \psi_{2m} dx_2 = 0. \end{aligned} \quad (4.114)$$

Затем подставим (4.111) — (4.113) в (4.114), что приведет к системе из N линейных относительно a_k уравнений:

$$\sum_{k=1}^N [B_{ki}^{(0)} + B_{ki}^{(1)} + B_{ki}^{(2)} + \lambda^2 (D_{ki}^{(0)} + D_{ki}^{(1)} + D_{ki}^{(2)})] = 0; \quad (4.115)$$

$$B_{ki}^{(j)} = a_{ki} \int_0^{l_j} (S^{(j)} \psi_{jk}')' \psi_{ji} dx_j; \quad D_{ki}^{(j)} = a_{ki} \int_0^{l_j} S^{(j)} \psi_{jk} \psi_{ji} dx_j,$$

причем j есть номер участка речевого тракта: $j=0$ — фарингиальная область, $j=1$ — переднеязычная область, $j=2$ — носовая полость. Как известно, система (4.115) имеет нетривиальное решение только в том случае, если ее определитель равен нулю. Отсюда найдем собственные числа λ_k и коэффициенты a_k , присоединяя недостающее уравнение, полученное из условия нормировки аналогично (4.104), где вместо ψ_{jk} подставляются новые собственные функции: $\varphi_{jm}(x_j) = \sum_{k=1}^N a_{km} \psi_{jk}(x_j)$, $j=0, 1, 2$. Таким образом, система волновых уравнений (4.105) — (4.107) оказывается решенной.

В экспериментах по моделированию эффекта назализации и вычислению частотных характеристик речевого тракта использовались функции площади поперечного сечения, заимствованные из [В.3]. Порядок приближения N выбирался из условия минимума отклонений от нуля недиагональных элементов матрицы скалярных произведений нормированных собственных функций φ_{jm} . При $N=10$ наибольшее отклонение в этой матрице примерно 0,04—0,06.

Синтез назальных звуков проводился на артикуляторном синтезаторе, в котором по функциям площади речевого тракта вычисляются частоты и затухания резонансных колебаний, учитываются характеристики излучения и влияние подсвязочной области на интервале открытой голосовой щели. Данные о площадях поперечного сечения дополнялись сведениями, полученными из кинорентгенографических измерений. Разветвление речевого тракта при подсоединении носовой полости создает сток энергии, который приводит к появлению антирезонансов. При этом возникают и новые резонансы, а потери увеличиваются вследствие излучения из ноздрей и дополнительного трения, теплопроводности и податливости стенок носовой полости. На рис. 4.19 показаны расчетные амплитудно-частотные характеристики речевого тракта для гласного звука A при закрытом и открытом проходах в носовую полость. Из рисунка видно, что в результате увеличения потерь и числа формант спектр становится более гладким, маскируя возрастание числа формант с 5 до 8. Аналогичные расчеты спектров

во время губной смычки показывают, что для опущенной небной занавески (соответствующей артикуляции звука M) уровень первой форманты понижается по сравнению с ее уровнем для звука B . Этот эффект обычно принимается за основной признак на-

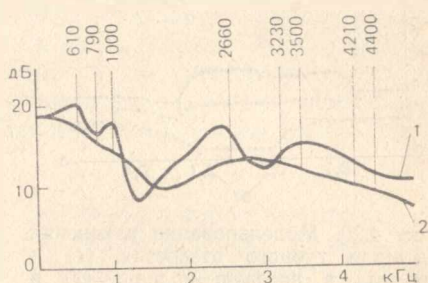


Рис. 4.19. Амплитудно-частотные характеристики неназализованного (1) и назализованного (2) гласного A

зализации, отличающий звонкие взрывные *Б, Д* от назальных *М, Н*.

В процессе моделирования динамики назализации изменялось время сдвига начала и конца движений небной занавески относительно движений губ (для *М*) или языка (для *Н*). Фонетическое качество синтезированного звуко сочетания оценивалось субъективно на слух. При этом выяснилось, что относительные сдвиги движений артикуляторных органов влияют на восприятие назальных звуков. Наилучшее качество достигается в том случае, когда небная занавеска начинает опускаться за 20—30 мс до начала смыкания губ (или подъема кончика языка), а к моменту взрыва смычки небная занавеска поднимается таким образом, что площадь прохода в носовую полость составляет около $0,03 \text{ см}^2$ (при площади полного прохода примерно 3 см^2). В момент завершения переходного процесса губ или языка небная занавеска должна полностью перекрывать проход в носовую полость. Соответствующие этим движениям траектории формант для звуко сочетания *АМА* показаны на рис. 4.20.

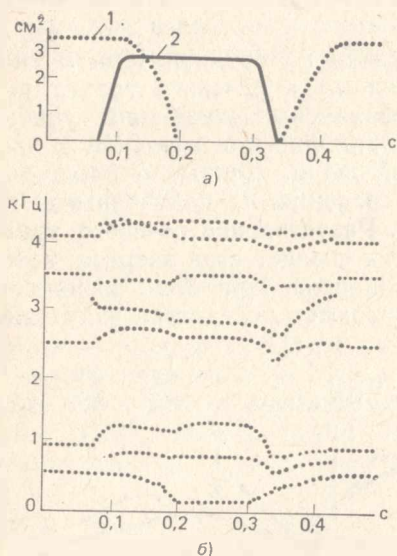


Рис. 4.20. Моделирование изменения площади губного отверстия (1) и прохода в носовую полость (2) в слове *АМА* (вверху) и вычисленные резонансные частоты (внизу)

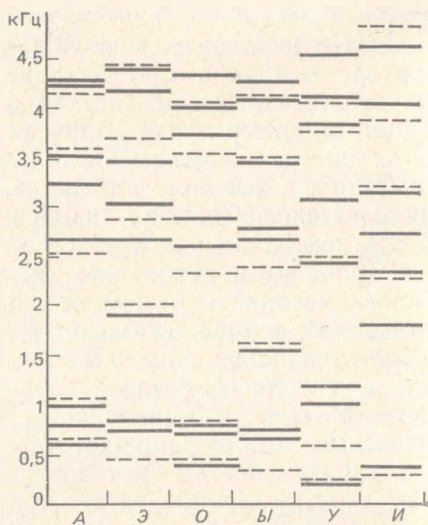


Рис. 4.21. Расчетные резонансные частоты неназализованных (---) и назализованных (—) гласных, упорядоченных по степени подъема небной занавески

Увеличение числа формант при подключении носовой полости происходит не только во время смычки, но и в ее окрестностях. При спектральном анализе близко расположенные резонансные частоты маскируют появление дополнительных формант, создавая впечатление значительного их расширения. Наряду с действитель-

ным увеличением потерь необходимо принимать во внимание и увеличение числа формант. Другой существенный элемент эффекта назализации состоит в несимметричности артикуляции и соответственно формантных траекторий слева и справа от смычки.

Следует отметить, что сдвиг частот основных резонансов речевого тракта и частоты дополнительных формант, возникающих при опускании небной занавески, слабо зависит от степени этого опускания. На рис. 4.21 сопоставляются статические значения резонансных частот неназализованных и назализованных гласных звуков русской речи.

Относительная энергия акустических колебаний, излучаемых через рот и через ноздри, зависит от величины прохода в носовую полость. Как известно, амплитуда резонансных колебаний пропорциональна значению соответствующих собственных функций на излучающем конце акустической системы. Например, значение собственных функций на губах для назализованного гласного *А* относительно больше значений собственных функций на ноздрах, тогда как для назализованного гласного *И* соотношение обратное.

4.6. ТУРБУЛЕНТНЫЕ ПРОЦЕССЫ

При артикуляции фрикативных звуков *С, Ш, Ф, Х, З, Ж, В* и аффрикатов *Ц, Ч* в речевом тракте создаются условия для возникновения турбулентности протекающего воздушного потока. В результате этого генерируется шум, имеющий характерный спектр для каждого такого звука. В настоящее время существует лишь так называемая полуэмпирическая теория турбулентных процессов в газовых и жидких средах, т. е. такая теория, в которую входит ряд параметров и зависимостей, устанавливаемых непосредственно опытным путем для конкретного класса систем. В приложении к процессам речеобразования вопросы теории шипящих звуков рассматривались в [В.3, В.4, 4.6, 4.19, 4.52, 4.54].

Для труб с постоянным сечением было установлено, что процесс турбулизации потока начинается, когда число Рейнольдса Re превосходит критическую величину $Re_{кр}$. Число Рейнольдса

$$Re = \rho_0 v h / \mu, \quad (4.116)$$

где ρ_0 — плотность воздуха; v — линейная скорость потока; μ — вязкость воздуха; h — характерный геометрический размер сечения трубы [2.4]: $h \approx 4S/L$, где S — площадь поперечного сечения; L — его периметр. Для круглого сечения h равно его диаметру. Для труб с постоянным сечением минимальное значение критического числа Рейнольдса $Re_{кр} \approx 2000$. Если труба имеет переменное сечение, то критическое число зависит от перепада давления в соседних сечениях: повышается для сужающихся труб и понижается для расширяющихся. В экспериментах на физических моделях речевого тракта было найдено, что критическое число $Re_{кр} \approx 1800$ [4.47]. На величину $Re_{кр}$ влияет также степень шеро-

ховатости стенок трубы и возмущения потока на входе трубы. Колебания голосовых складок, сопровождающие артикуляцию звонких фрикативных *З, Ж, В*, создают интенсивные возмущения воздушного потока, поэтому для этих звуков можно ожидать понижения критического числа по сравнению с их глухими аналогами — звуками *С, Ш, Ф, Х*.

Плотность воздуха и вязкость μ в речевом тракте иные, чем в свободной среде, из-за наличия влаги на стенках, а сами стенки не являются абсолютно жесткими. Влияние этих факторов на условия шумообразования пока никем не исследовалось.

При резком расширении трубы скорость потока на некотором расстоянии от этого расширения еще не успевает упасть, поэтому в речевом тракте турбулизация потока, сопровождающаяся возникновением шумов, почти всегда появляется на выходе из голосовой щели. Небольшие шумы в звонких звуках, в том числе и гласных, можно измерить, а добавление шума в синтетическую речь улучшает натуральность звучания [4.37].

Сопротивление ламинарному потоку в некотором сечении речевого тракта складывается из сопротивления вязкого трения R_v и динамического сопротивления R_d . Сопротивление трения может быть рассчитано по (4.74) для средних и малых (порядка $0,2 \text{ см}^2$) площадей сечения речевого тракта. Вклад вязкого трения ламинарной компоненты в потери давления невелик, за исключением случая, когда площадь сужения меньше $0,05 \text{ см}^2$ или перепад давления на сужении меньше 10 мм. вод. ст. Динамическое сопротивление обычно значительно превышает сопротивление вязкого трения. В выражении

$$R_d = c_x \rho_0 v / 2S \quad (4.117)$$

имеется коэффициент c_x , определяемый полуэмпирическим способом. В [4.54] принимается, что для плавного сужения $c_x = 0,5$, для плавного расширения $c_x = 0,2$, а для резкого расширения $c_x = 1,2 - 1,7$. Для расширяющегося непосредственно над голосовыми складками потока в [4.21] принимается $c_x = 0,875$, а в [4.38] исходя из экспериментов на модели гортани предлагается следующая формула для определения c_x :

$$2c_x = 1,375 - 2 \frac{S_{г.щ}}{S_0} \left(1 - \frac{S_{г.щ}}{S_0} \right),$$

где $S_{г.щ}$ — площадь голосовой щели; S_0 — площадь речевого тракта над поверхностью голосовых складок. Для расчета турбулентных процессов в речевом тракте в [4.54] используется другая формула:

$$c_x(x) = \begin{cases} c_{x1} + c_{x2} \left[\frac{S(x + \Delta x)}{S(x)} - 1 \right]^2, & S(x + \Delta x) > S(x); \\ c_{x2}, & S(x + \Delta x) \leq S(x), \end{cases} \quad (4.118)$$

т. е. принимается, что для сужающейся трубы коэффициент динамического сопротивления постоянен: $c_x = c_{x2} = 0,03$, а для расши-

ряющей трубы c_x зависит от крутизны расширения и $c_{x1}=0,05$.

В турбулентном потоке сопротивление увеличивается из-за теплового рассеяния энергии в завихрениях воздуха. В [4.4] было найдено, что турбулентное сопротивление в трубе с сужением

$$R_{тб} = \left(\frac{S_2}{S_1} - 1 \right)^2 M, \quad (4.119)$$

где S_1 — площадь поперечного сечения трубы в сужении; S_2 — площадь трубы за сужением (площадь расширяющегося участка); $M = v/c_0$ — число Маха. Как видно, уравнение (4.119) по форме весьма близко (4.118) и вносит поправку на c_x в зависимости от скорости потока. В сужениях с площадью сечения порядка $0,2 \text{ см}^2$ скорость воздушного потока находится в диапазоне 10 — 50 м/с , что соответствует числу Маха $M \approx 0,03$ — $0,14$. Таким образом, сопротивление турбулентного потока, найденное по (4.119), близко сопротивлению, вычисленному по (4.118). В [4.6] обсуждаются и более сложные зависимости коэффициента c_x от геометрических параметров и скорости потока для ламинарного и турбулентного режимов.

В трубе с переменным поперечным сечением, как это имеет место в речевом тракте, полное динамическое сопротивление

$$R_0 = \frac{\rho_0}{2} \int_0^l \frac{c_x(x) v(x)}{S(x)} dx, \quad (4.120)$$

где l — длина речевого тракта; $S(x)$ — площадь поперечного сечения; $v(x)$ — распределение линейной скорости воздушного потока вдоль речевого тракта. Поскольку число Маха для воздушных потоков в речевом тракте значительно меньше единицы, как было показано в § 2.4, справедливо условие неразрывности потока для несжимаемого газа

$$w = v(x) S(x) = \text{const.} \quad (4.121)$$

Объемная скорость w аналогична электрическому току, а давление P — напряжению. Воспользовавшись этой аналогией, представим полное падение напряжения на эквивалентной электрической схеме речевого тракта как $P_0 = w R_0$. Тогда, учитывая (4.120) и (4.121), получим неявную форму для вычисления скорости потока в речевом тракте

$$v(x) = \frac{2P_0}{\rho_0 S(x) \int_0^l \frac{c_x(x) v(x)}{S(x)} dx}, \quad (4.122)$$

где P_0 — давление в легких [4.19]. Наименьшая площадь сужения в речевом тракте для фриктивных звуков сопоставима с площадью голосовой щели, поэтому сопротивление голосовой щели также должно быть включено в (4.122). С этой целью координата $x=0$ располагается не на верхней поверхности голосовых складок, как это обычно делается, а на некотором расстоянии перед

голосовой щелью, например на входе в трахею. Распределение давления вдоль речевого тракта следующее:

$$P(x) = P_0 - \omega \int_0^x R_d(x) dx,$$

или

$$P(x) = P_0 \left(1 - \frac{\int_0^x R_d(x) dx}{\int_0^l R_d(x) dx} \right). \quad (4.123)$$

Полученное выражение позволяет определить перепад давления $\Delta P = P_0 - P(x_0)$ на голосовых складках (x_0 — расстояние от начала трахеи до верхней поверхности голосовых складок), необходимый для расчета режима автоколебаний голосовых складок для звонких фрикативных звуков (см. гл. 2). Величина ΔP отражает «обратное» влияние речевого тракта на параметры голосового источника.

Уравнение (4.122) решается итеративным способом, причем для достижения устойчивости необходимо совершить замену переменных $\bar{v} = \kappa v$, где $\kappa \approx 10^{-3}$. Рассчитанное таким образом распределение скорости потока вдоль речевого тракта, начиная от трахеи, показано на рис. 4.22. Как видно, локальные максимумы скорости и соответствующие им максимумы числа Рейнольдса приходятся на сужения с последующим расширением. В тех областях, где число Рейнольдса превышает критическое значение, возникают турбулентные шумящие источники. Временные характеристики порождаемого шума могут быть описаны функцией $\xi(t) = d\xi(t)/dt$, где $\xi(t)$ есть решение дифференциального уравнения $\xi'' + 2g\omega_0\xi' + \omega_0^2\xi = \eta(t)$; $\eta(t)$ — случайная функция, равномерно распределенная на интервале $[0,1]$; параметр $g \approx 1$, а частота, определяющая максимум спектра турбулентного шума,

$$\omega_0 \approx \frac{0,2}{\pi} \frac{|v_{max}|}{\sqrt{S_m}}.$$

Здесь v_{max} — максимальная скорость воздушного потока в шумящем сужении; S_m — площадь этого сужения [4.52, 4.54].

Максимальная скорость потока в речевом тракте зависит не только от сужения в его верхней части, но и от площади голосовой щели, поэтому характеристики шума фрикативных звуков также зависят от геометрии голосовой щели. Мощность звука, генерируемого шумящим потоком в круглой трубе, пропорциональна третьей степени от перепада давления ΔP (как для дипольного источника [4.54]): $W_P \sim \Delta P^3 d^2 / \rho^2 c^2$, где d — диаметр трубы.

Как упоминалось в § 4.1, при выводе волнового уравнения принимается, что вихревые потоки отсутствуют. Следовательно, волновое уравнение, например, в форме (4.26), (4.29) или (4.33), строго говоря, не оправдливо для расчета частотных характеристик речевого тракта при шумообразовании. В то же время

волновым уравнением можно воспользоваться как некоторым приближением, поскольку известно, что человеческое восприятие не очень критично к изменению частотных характеристик фрикативных звуков в определенном диапазоне. Следуя принятому допу-

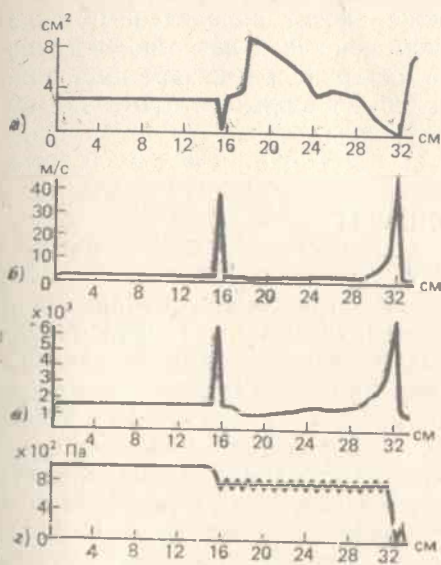


Рис. 4.22. Характеристики фрикативного С:

а — площадь поперечного сечения; б — линейная скорость воздушного потока; в — число Рейнольдса; г — давление

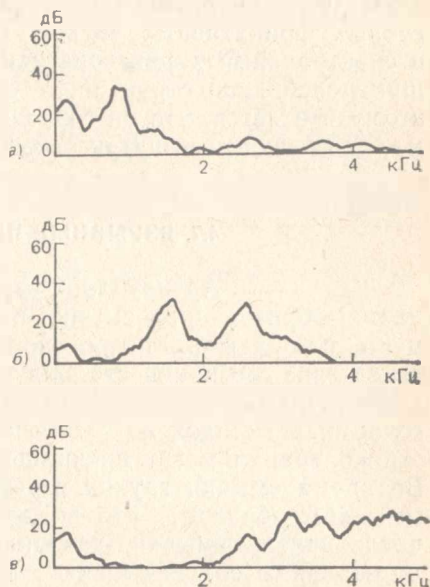


Рис. 4.23. Расчетные спектры фрикативных:

а — звук Х; б — звук Ш; в — звук С

щению, представляем шумящий поток как распределенный источник. Тогда создаваемое этим источником парциальное возбуждение временных мод согласно (4.60) определим как

$$F_k(t) = \frac{2}{\rho_0 l} \int_0^l A(x, t) \psi_k(x) dx, \quad (4.124)$$

где $\psi_k(x)$ — собственные функции речевого тракта, найденные по уравнению Вебстера. Вследствие того, что сопротивление голосовой щели для фрикативных звуков речи сопоставимо с сопротивлением речевого тракта, при решении уравнения Вебстера нужно учитывать трахею, что удлиняет речевой тракт и увеличивает число резонансов в полосе до 5 кГц. Если шум возникает в области нуля той или иной собственной функции ψ_k , то интеграл в (4.124) близок к нулю и соответствующая временная мода не возбуждается. В результате в спектре фрикативного звука образуются области с малыми энергиями, причем они могут находиться как выше по частоте от области с большой энергией, так и ниже ее.

Спектры, рассчитанные для звуков *С, Ш, Х* без учета влияния шумов голосовой щели, показаны на рис. 4.23. Субъективная оценка качества синтезированных таким образом фрикативных звуков оказывается вполне удовлетворительной.

Струя выдыхаемого воздуха, особенно для переднеязычных и губных фрикативных звуков, также вносит определенный вклад в спектральные характеристики шума, но, как известно, источник шума свободной струи носит уже квадрупольный характер, и поэтому считается, что он гораздо слабее дипольного источника шума, действующего внутри речевого тракта.

4.7. ВЗРЫВНЫЕ ПРОЦЕССЫ

Артикуляция взрывных согласных звуков сопровождается образованием смычки в речевом тракте, когда в некотором месте площадь его поперечного сечения становится равной нулю вследствие смыкания губ или подъема языка до соприкосновения с небом. Звонкие согласные обычно характеризуются колебанием голосовых складок во время смычки, хотя в некоторых случаях эти колебания могут прекращаться незадолго до взрыва смычки. Во время смычки глухих взрывных голосовые складки расходятся таким образом, что их автоколебания прекращаются вскоре после начала смычки, давление в речевом тракте увеличивается до величины подсвязочного давления, а затем резко падает после взрыва смычки, создавая так называемый импульсный источник возбуждения. В русском языке колебания голосовых складок возобновляются в среднем через 15—45 мс после взрыва смычки. Законы изменения внутриротового давления во время смычки несколько различаются для глухих и звонких взрывных: для глухих это обычно выпуклая кривая, а для звонких — вогнутая или линейная [4.48]. Место артикуляции почти не влияет ни на форму изменения, ни на максимальное значение внутриротового давления, а окружающие гласные влияют главным образом на время спада давления после взрыва смычки.

При артикуляции взрывных звуков, так же как и при артикуляции фрикативных, речевой тракт может рассматриваться как система с сосредоточенными параметрами относительно процессов изменения внутриротового давления, но, в отличие от фрикативных, здесь важную роль играют упругость воздуха в объеме легких и податливость стенок легких и речевого тракта. Механическая модель и ее электрический аналог для апериодических процессов в тракте и подсвязочной области были предложены в [2.81] (рис. 4.24) и подробно исследованы в [4.48]. В этой модели учитываются сопротивление речевого тракта R_0 , инерционность L_w , упругость C_w и потери R_w в податливых стенках, упругость воздуха в речевом тракте $C_{тр}$, сопротивление голосовой щели $R_{г.щ}$, инерционность L_w , упругость C_w и потери R_w в стенках легких, упругость воздуха C_l и потери R_l в легких, а также упругость

воздуха C_1 , C_2 и потери R_1 , R_2 в бронхах и трахее. Кроме того, учитывается поток воздуха через голосовую щель вследствие изменения объема речевого тракта.

Кинорентгенограммы речевого тракта показывают, что артикуляция звонких согласных отличается от артикуляции глухих со-

гласных расширением фарингиальной области за счет либо смещения языка вперед, либо опускания гортани. Максимальная скорость изменения линейных размеров фарингиальной области, судя по рис. 4.25, составляет примерно 5—7 см/с. Если принять, что ширина фарингиальной области равна 3—4 см, а ее длина — 6—8 см, то максимальная объемная скорость воздушного потока, дополнительно протекающего через голосовую щель из подсвязочной области в речевой тракт, находится в диапазоне 90—220 см³/с. Эта величина сравнима с объемной скоростью потока, создаваемого перепадом подсвязочного и надсвязочного давлений. По-

явление этой скорости объясняет продолжение фонации на интервале звонкой смычки, несмотря на непрерывно поступающий воздух из легких в речевой тракт. Следует отметить, однако, что для данного диктора максимальная скорость увеличения объема приходится не на середину или конец смычки, а на конечный участок переходного процесса артикуляторных органов, движущихся к смычке. В то же время нужно иметь в виду, что объем фарингиальной области может изменяться также и за счет опускания гортани [2.61], так что обычно этот объем возрастает на протяжении всей звонкой смычки [4.59].

Об изменении надсвязочного давления можно косвенно судить по изменениям амплитуды речевого сигнала и частоты основного тона на интервале звонкой смычки. При этом наблюдаются различные виды изменений: амплитуда речевого сигнала может сохраняться постоянной, нарастать с последующим падением либо падать до нуля в конце смычки. Очевидно, эти явления отражают индивидуальное разнообразие в управлении подсвязочным давлением и скоростью изменения объема фарингиальной области. В частности, для диктора, артикуляционные движения которого

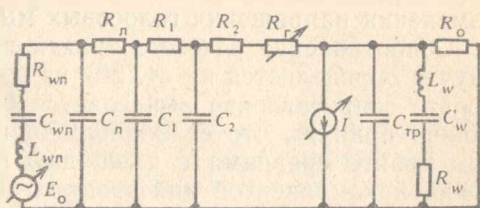


Рис. 4.24. Электроакустическая модель речевого тракта

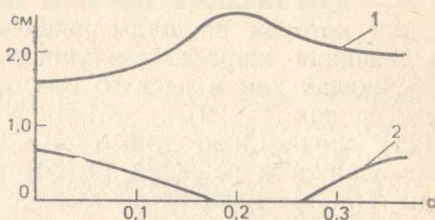


Рис. 4.25. Ширина фарингиальной области (1) и расстояние от кончика языка до твердого неба (2) в слове АДА

показаны на рис. 4.25, характерно падение амплитуды речевого сигнала к концу звонкой смычки. Частота основного тона во время смычки обычно остается постоянной или немного падает, но встречаются и случаи ее возрастания, причем даже на фоне уменьшающейся амплитуды колебаний, что свидетельствует также и об изменении напряжения голосовых мышц.

Сопротивление речевого тракта постоянному току R_0 в общем случае определяется по (4.120) из предыдущего параграфа. Обычно это сопротивление вычисляется более простым способом. Так, можно принять, что если площадь наибольшего сужения в речевом тракте сравнима с площадью голосовой щели (или ее постоянной компонентой для звонких звуков), то остальные участки речевого тракта вносят незначительный вклад в общую величину сопротивления

$$R_0 = \rho_0 c_x(x_0) v(x_0) / 2S(x_0), \quad (4.125)$$

где x_0 — максимальное значение координаты вдоль речевого тракта, для которой площадь поперечного сечения S минимальная; v — линейная скорость воздушного потока в этом сечении; c_x — коэффициент динамического сопротивления. Количественное сравнение оценок (4.120) и (4.125) показывает, что, хотя в статике (4.125) справедливо для $S(x_0) \leq 0,2 \text{ см}^2$, условие динамической согласованности решений (4.120) и (4.125) при переменной площади $S(x, t)$ значительно сужает область применения (4.125).

Если площадь наибольшего сужения в речевом тракте значительно больше площади голосовой щели, то сопротивление речевого тракта слабо зависит от его формы и может быть принято за постоянную величину, причем $R_0 \ll R_{\Gamma}$.

Упругость воздуха внутри речевого тракта определяется как

$$C_{\text{тр}} = \int_0^{x_m} S(x) dx / \rho_0 c_0^2,$$

где в случае смычки x_m равно координате вдоль речевого тракта, в которой площадь поперечного сечения S впервые обращается в нуль, а в остальных случаях x_m равно длине речевого тракта; c_0 — скорость звука. Моделирование аperiодических процессов в речевом тракте, проведенное в [4.48], показало, что величиной упругости воздуха $C_{\text{тр}}$ можно пренебречь по сравнению с упругостью стенок C_w (например, при объеме речевого тракта в 100 см^3 $C_{\text{тр}} \approx 6 \cdot 10^{-5} \text{ см}^4 \cdot \text{с}^2/\text{г}$, тогда как в зависимости от степени напряжения стенок тракта при площади их поверхности 100 см^2 $C_w = 4,5 \cdot 10^{-4} \text{—} 1,2 \cdot 10^{-3} \text{ см}^4 \cdot \text{с}^2/\text{г}$ [4.39]). По данным той же работы [4.39] удельное активное сопротивление стенок тракта находится в диапазоне $R_w = 10,6 \text{—} 8 \text{ г/см}^4 \cdot \text{с}$, а индуктивное сопротивление $L_w = 0,015 \text{—} 0,021 \text{ г/см}^4$. Изменение напряжения стенок тракта приводит к изменению максимального значения внутриротового давления на $5 \text{—} 10 \text{ мм вод. ст.}$ при давлении в легких 80 мм вод. ст. [4.48] и в то же время делает переходные процессы более затянутыми, увеличивая время достижения критического значения над-

связочного давления, при котором происходит срыв автоколебаний голосовых складок. Пожертвовав некоторыми подробностями описания этих процессов, можно упростить схему влияния податливости стенок (отбросить их активные потери и индуктивность и оставить лишь упругую компоненту).

Сопротивление голосовой щели

$$R_{г.щ} = \frac{\rho_0 c_{х.г.щ} v_{г.щ}}{2 S_{г.щ}} + R_{\tau}, \quad (4.126)$$

где $v_{г.щ}$ — линейная скорость воздушного потока через голосовую щель; $S_{г.щ}$ — площадь голосовой щели; $c_{х.г.щ}$ — коэффициент динамического сопротивления; R_{τ} — сопротивление вязкого трения. На интервале звонкой смычки при продолжающейся фонации скорость $v_{г.щ}$ и площадь $S_{г.щ}$ меняются быстро по сравнению с изменением давления в речевом тракте. Поэтому сопротивление голосовой щели можно представить его средним значением:

$$\bar{R}_{г.щ} = \frac{1}{T_0} \int_0^{T_1} R_{г.щ}(t) dt,$$

где T_0 — период основного тона; T_1 — длительность открытой голосовой щели. Величину $\bar{R}_{г.щ}$ можно приближенно оценить, проаппроксимировав форму $S_{г.щ}(t)$ в виде равнобедренного треугольника и приняв $V_{г.щ} = V_{max} = \text{const}$. Средняя площадь голосовой щели на периоде основного тона $S_{г.щ} \approx (0,15—0,25) S_{г.щ max}$, где $S_{г.щ max}$ — максимальное значение площади голосовой щели. Подставив $\bar{S}_{г.щ}$ в (4.126), получим оценку среднего сопротивления $\bar{R}_{г.щ}$. Для частоты основного тона F_0 вблизи 100 Гц и максимальной площади голосовой щели $S_{г.щ max} = 0,2 \text{ см}^2$ средняя площадь $\bar{S}_{г.щ} = 0,04 \text{ см}^2$, а среднее сопротивление голосовой щели $\bar{R}_{г.щ} = 70 \text{ г/см}^4 \cdot \text{с}$.

Упругость воздуха в легких $C_{л} = 0,003$ при объеме $V_{л} = 4500 \text{ см}^3$, а податливость стенок легких по оценке [2.81] $C_{шл} \approx 0,1$. Пренебрегая потерями в трахее и представляя воздушную полость под голосовыми складками в виде одного объема, получаем упрощенную схему для аperiодических процессов, показанную на рис. 4.26. Эта схема описывается следующей системой уравнений:

$$\left. \begin{aligned} I_1 R_{шл} + I_1' L_{шл} + \frac{1}{C_{шл}} \int I_1 dt &= P_{л} - P_0; \\ I_2 &= C_{л} P_0'; \\ P_{\tau} &= P_0 - (I_4 + I_5 + I_6) R_{г.щ}; \\ I_1 &= I_2 + I_4 + I_5 + I_6; \\ I_4 &= V_{\tau}'; \\ I_5 &= C_{ст} P_{\tau}'; \\ I_6 R_{\tau} &= P_{\tau}, \end{aligned} \right\} \quad (4.127)$$

где P_l — давление, создаваемое мышцами диафрагмы; P_0 — под-
связочное давление; $P_{тр}$ — давление в речевом тракте; $V_{тр}$ — объем
речевого тракта; I_4 — ток, возникающий вследствие изменения
объема $V_{тр}$. По оценке [2.81] инерционность стенок легких $L_{wl} \approx$
 $\approx 0,037$, а сопротивление $R_{wl} \approx 3$. Принимая во внимание наблю-

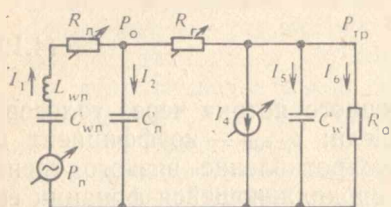


Рис. 4.26. Упрощенная электроакустическая модель речевого тракта

дающееся существенное падение
подсвязочного давления синхронно с колебаниями площади голо-
совой щели во время фонации, необходимо сделать вывод о за-
висимости легочного сопротивле-
ния от потока через голосовую
щель. Действительно, по мере вет-
вления бронхиол в легких пло-
щадь каждого канала уменьшает-
ся, хотя общая площадь попереч-
ного сечения растет, начиная с

глубины 20 см ниже голосовых складок. Для оценки легочного со-
противления в i -м канале перепишем (4.126), используя вместо ли-
нейной скорости v объемную скорость $w = vS$:

$$R_{li} = \frac{\rho_0 c_{xli} w_i}{2 S_i^2} + R_{ti} \quad (4.128)$$

Объемная скорость в каждом сечении легких остается постоян-
ной, тогда как площадь поперечного сечения S_i с ростом числа
ветвлений уменьшается. В итоге сопротивление легких можно
представить в виде двух компонент: сопротивления трения $R_{л.т}$,
не зависящего от объемной скорости, и динамического сопротивле-
ния $R_{л.д}$, зависящего от объемной скорости: $R_l = R_{л.т} + R_{л.д}$, где
 $R_{л.т} \approx 3$, а максимальное значение $R_{л.д}$, судя по результатам мо-
делирования, достигает 10—15.

Необходимо отметить, что система (4.127) является нелиней-
ной вследствие зависимости сопротивлений R_l , $R_{г.щ}$ и R_0 от про-
текающего через них тока. На рис. 4.27 и 4.28 показано измене-
ние давления и скорости потока в речевом тракте при моделиро-
вании слогов АТА и АДА, выполненном при условии, что ток $I_4 = 0$
и индуктивность легких $L_{wl} = 0$. Из этих рисунков видно, что им-
пульсное возбуждение возникает как для взрывных глухих звуков,
так и для звонких, но амплитуда импульса для звонких суще-
ственно меньше амплитуды импульса для глухих, так что обычно
в реальной речи импульсное возбуждение для звонких незаметно
на фоне колебаний. Однако у некоторых дикторов импульсное
возбуждение для звонких звуков отчетливо выражено, и это свя-
зано, конечно, с индивидуальными особенностями управления пло-
щадью голосовой щели и легочным давлением.

В русском языке взрыв и последующий шумовой интервал сла-
бее всего выражены для глухого согласного П. Это объясняется
тем, что условия излучения на губах соответствуют квадруполь-

ному источнику, интенсивность которого значительно меньше интенсивности дипольных источников, создаваемых при взрыве согласных *T* и *K*. Взрыв на *K* обычно имеет большую амплитуду, чем взрыв на *T*, причем чаще наблюдается двойное, реже — многократное возбуждение. Причина этого явления состоит в том,

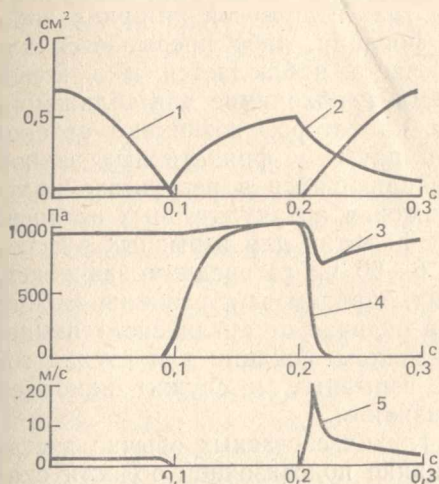


Рис. 4.27. Аперiodические процессы при моделировании слога *ATA*:

1 — площадь сужения; 2 — площадь голо-
совой щели; 3 — подсвязочное давление;
4 — надсвязочное давление; 5 — поток

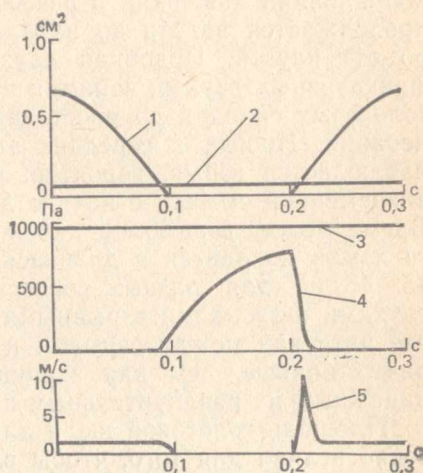


Рис. 4.28. Аперiodические процессы при моделировании слога *ADA*.

Обозначения те же, что и на рис. 4.27

что, как и в случае голосового источника, сигнал возбуждения при взрыве создается производной от объемной скорости, которая при взрыве имеет положительный и отрицательный выбросы, вызывающие двукратное возбуждение речевого тракта. Дополнительное возбуждение может быть создано и в результате неравномерности движения артикуляторных органов, в том числе и при повторных кратких смычках. Между первым и вторым импульсами для согласного *K* проходит в среднем 5—12 мс, что близко к данным для английского языка [4.26].

Глухая смычка в действительности не является акустически однородным явлением, а содержит последовательность сегментов, характеристики которых определяются координацией артикуляторных движений и движений голосовых складок. Обычно фонация прекращается через 20—30 мс после начала смычки. За это время голосовые складки расходятся с постепенным уменьшением амплитуды колебаний на расстояние, при котором автоколебания уже не могут поддерживаться. Такой же участок затухающих колебаний наблюдается и на глухих фриктивных звуках, причем в этом случае к голосовому возбуждению примешивается и шум турбулентности. Затем следует пауза, в течение которой в рече-

вом тракте накапливается давление, быстро падающее после взрыва смычки. В русском языке вслед за импульсным возбуждением взрыва (и на его фоне) обычно возникает шумовое возбуждение, продолжающееся до тех пор, пока артикуляторные органы и голосовые складки не достигнут расстояния, при котором число Рейнольдса падает ниже порога турбулизации. В зависимости от координации движений в речевом тракте шумовой интервал либо продолжается вплоть до начала фонации, либо прерывается короткой паузой. Подобная пауза часто наблюдается и в конце фрикативных звуков, занимая время, необходимое для сближения голосовых складок до расстояния, на котором возникают автоколебания. Иногда в середине этой паузы у фрикативных звуков наблюдается взрыв, вероятно, появляющийся в результате кратковременной смычки в начале движения артикуляторных органов. Время между взрывом и началом фонации для взрывных русского языка находится в диапазоне 6—60 мс со средним значением 15—43 мс для разных согласных, причем нет разницы между глухими и звонкими взрывными, в отличие от английского языка, где интервал между взрывом и началом фонации для глухих гораздо больше, чем для звонких взрывных, и служит наиболее надежным их различительным признаком.

Площадь голосовой щели для глухих согласных обычно достаточно велика для того, чтобы влияние подсвязочной области сказалось на акустических характеристиках речевого тракта. Экспериментальные и теоретические исследования показывают, что в этот момент число резонансов в заданном частотном диапазоне увеличивается и наиболее заметный из них находится в области 1700 Гц ([4.18, 4.27], см. также § 4.3). Импульсный источник взрыва и последующий за ним шумовой источник могут быть расположены в различных местах речевого тракта в зависимости от места артикуляции, поэтому акустические характеристики возбуждаемого сигнала подчиняются тем же закономерностям, что и характеристики фрикативных звуков. В частности, при совпадении места артикуляции с нулем какой-либо собственной функции подавляются соответствующие частотные области в спектре речевого сигнала. Это явление иллюстрируется экспериментами по восприятию слогов типа «фрикативный — гласный» при укорочении длительности фрикативного: сначала вместо фрикативного воспринимается аффрикат (Ц или Ч), затем — глухой взрывной с тем же местом артикуляции, что и фрикативный.

Амплитудно-частотные характеристики взрыва для английского языка исследовались в [4.22—4.24, 4.41, 4.50, 4.56, 4.58]. Результаты этих исследований несколько противоречивы, однако по ним можно заключить, что место артикуляции звонких взрывных с достаточной степенью надежности опознается главным образом по амплитудно-частотному спектру речевого сигнала, вычисленному в интервале 10—20 мс сразу после начала формантных переходов, а также по направлению этих переходов. Для глухих взрывных направление формантных переходов играет второсте-

пенную роль, а восприятие места артикуляции опирается на АЧХ сигнала и их динамику [4.50] на интервале 10—20 мс сразу после взрыва. Имеется существенная разница в спектральных характеристиках между звонкими и глухими согласными, а также между формантными переходами от гласного к смычке и от смычки к гласному. Для звонких взрывных во время смычки колебания голосовых складок обычно продолжают вплоть до ее взрыва, поэтому максимальная задержка появления импульса голосового возбуждения после начала переходного процесса равна периоду основного тона, т. е. меньше 10—12 мс. И хотя, как показано в § 4.2, скорость переходных процессов для формантных частот наибольшей именно на начальном участке, переходы формант все же достаточно отчетливо проявляются при взрыве звонких согласных, доставляя информацию о месте артикуляции.

Совершенно иная ситуация имеет место при взрыве глухих согласных: здесь время между началом движения артикуляторных органов и появлением первого импульса голосового возбуждения обычно больше 10 мс, а для заднеязычного *К* для русского языка в среднем равно 28 мс перед безударными гласными и 43 мс перед ударными. За это время переходные процессы формантных частот для многих гласных звуков успевают почти закончиться, и по оставшимся участкам обычно трудно судить о месте артикуляции взрывного согласного. Этот эффект сильнее проявляется в английском языке, чем в русском, поскольку там длительность интервала между взрывом глухой смычки и началом фонации больше. Именно этим явлением объясняется малая информативность как спектра в начале формантных переходов, так и направления этих переходов для глухих взрывных. Имеются сведения, что этот эффект существует и в русском языке. Информация о месте артикуляции сохраняется в формантных переходах от гласного к смычке независимо от того, является ли она звонкой или глухой, поскольку на этом участке фонация продолжается вплоть до начала смычки. В зависимости от последующего гласного скорость формантных переходов после взрыва глухой смычки различна, поэтому первый импульс голосового источника возбудит акустические колебания на различных частотах. Это явление действительно наблюдается в речевых сигналах [4.55].

Как будет показано в гл. 5, место участка в речевом тракте с наибольшим сужением оказывает заметное влияние на АЧХ речевого сигнала, хотя конфигурация остальных участков также влияет на них. Поэтому спектр речевого сигнала и его изменения непосредственно после взрыва в наибольшей степени характеризуют место артикуляции взрывных согласных, как это было установлено в [4.22—4.24, 4.50], хотя и не являются инвариантами места артикуляции. В то же время, различия в акустических характеристиках между глухими и звонкими взрывными, рассмотренные выше, разрешают некоторые противоречия во мнениях, касающихся роли тех или иных сегментов речевого сигнала при восприятии взрывных согласных.

СИСТЕМА УПРАВЛЕНИЯ АРТИКУЛЯЦИЕЙ

5.1. ЭЛЕМЕНТЫ СТРУКТУРЫ СИСТЕМЫ УПРАВЛЕНИЯ

В процессе управления речеобразованием участвуют области коры обоих полушарий мозга, верхние сегменты спинного мозга и мозжечок, болезни которого вызывают нарушение временных характеристик речи (скорости, ритма), влияют на заключительные стадии речевых движений и могут вызывать заикания [5.23, 5.94, 5.95]. В системе управления мышечными сокращениями большую роль играют так называемые двигательные единицы, каждая из которых содержит α -мотонейрон и иннервируемые им мышечные волокна [5.11, 5.25]. Мотонейроны черепных мышц находятся в верхних сегментах спинного мозга, и каждый из них иннервирует некоторое количество мышечных волокон, распределенных по поперечному сечению мышцы. Различают α -мотонейроны и γ -мотонейроны, управляющие волокнами мышечных рецепторов — веретен. α -мотонейрон охвачен обратной связью через клетку Реншоу, которая также получает сигналы от других α -мотонейронов. α -мотонейроны генерируют последовательности импульсов с различными частотами.

Большинство мышечных волокон мышц рта обладает способностью быстро сокращаться [5.138, 5.156]. В подбородочно-язычной мышце, а также других мышцах речевого аппарата в изометрическом режиме (т. е. при создании напряжения с сохранением постоянной длины мышцы) наблюдаются частоты следования импульсов в диапазоне 13—30 имп./с [5.115], тогда как измерения в процессе артикуляции показали существование частот до 200 имп./с. Эти частоты соответствуют так называемому изотоническому режиму (т. е. сокращению с постоянной нагрузкой), характерному для процессов артикуляции.

Моторные единицы различаются не только по временным характеристикам, но и по величине развиваемого ими усилия. Например, моторные единицы передней головки двубрюшной мышцы, опускающей нижнюю челюсть, развивают усилия 0,01—4,98 г со средним значением 0,26 г, а длительность их переходных процессов колеблется между 16 и 56 мс со средним значением 29 мс [5.116]. Двигательные единицы жевательной мышцы, поднимающей нижнюю челюсть, гораздо сильнее. Они создают усилия 0,1—205 г со средним значением 83 г, а их переходные процессы длятся 30—50 мс. В процессах артикуляции соблюдается так называемый

мый «принцип размера», согласно которому для медленных движений с малыми амплитудами активизируются мелкие двигательные единицы, создающие малые усилия, а для быстрых движений с большими амплитудами активизируются крупные двигательные единицы, создающие большие усилия.

Амплитуда импульсов, возбуждающих крупные двигательные единицы, примерно в 3—4 раза больше амплитуды импульсов мелких двигательных единиц [5.134], но уровень измеряемой ЭМГ (электромиограммы) при больших мышечных усилиях увеличивается в 100—1000 раз, поэтому предполагается, что при больших усилиях происходит синхронизация разрядов двигательных единиц [5.25]. В пользу этого предположения свидетельствует отличие между поведением модели, основанной на принципе асинхронного разряда мотонейронов со случайным распределением межимпульсных интервалов [5.20, 5.21] (см. также § 5.5), и измеренной ЭМГ. Вопрос о синхронизации окончательно еще не решен, и рассматриваются различные механизмы, лежащие в основе увеличения уровня измеряемых ЭМГ при больших мышечных усилиях. К проблеме синхронизации может иметь отношение так называемое явление «затягивания», состоящее в синхронизации частот множества независимых генераторов, работающих на общую нагрузку. Это явление обнаруживается, в частности, и в электрической активности мозга [5.9].

5.2. РЕЦЕПТОРЫ

Обратная связь — это передача в центральную нервную систему информации о движениях артикуляторного аппарата. Информация передается через акустический канал и слуховой анализатор и через рецепторы мышц. Рецепторы, находящиеся в мышцах, играют роль датчиков и передают сведения о состоянии мышечной системы в соответствующие центры управления. Часть этой информации необходима для обеспечения надлежащей точности сокращения мышц, другая часть используется для координации мышечных движений для достижения заданных целей артикуляции. В настоящее время специфика рецепторов, принимающих участие в управлении артикуляцией, мало изучена. Частично этот недостаток может быть восполнен за счет использования описаний рецепторов, содержащихся в других мышцах человека и животных, но следует учитывать ограниченные возможности такого переноса свойств сравнительно хорошо изученных рецепторов на рецепторы речевых мышц. Фактически пока можно говорить лишь о том, какую роль рецепторы могли бы играть в управлении артикуляцией, а не о конкретных схемах управления. Обратные связи могут быть локальными и могут использоваться для коррекции характеристик отдельных участков нейромышечной системы и глобальными, охватывающими всю систему. Чтобы получить хотя бы приблизительное представление о свойствах обратной связи, нуж-

но оценить динамические и частотные характеристики включенных в нее рецепторов.

В мышцах артикуляторных органов и тканях полости рта находятся рецепторы с быстрым откликом, реагирующие на скорость и ускорение (мейснеровы частицы и тельца Паччини), и рецепторы с медленным откликом, реагирующие на смещение и скорость (тельца Руффини, органы Гольджи и веретена). Одни рецепторы являются линейными элементами, а другие — обладают нелинейными характеристиками. Количественное описание характеристик рецепторов дает ясное представление о возможных свойствах обратной связи в системе управления артикуляцией и позволяет расчитать некоторые процессы управления.

Рецепторы сухожилий (органы Гольджи). Эти рецепторы находятся в сухожилиях в области перехода мышечных волокон в сухожилия [5.11]. В системе речевых мышц органы Гольджи содержатся только в мышцах, управляющих движениями нижней челюсти. Долгое время считалось, что эти рецепторы играют роль датчика «аварийного» натяжения, подавая сигнал о приближении к пределу механической прочности мышц. Имеются сведения о том, что эти рецепторы способны также непрерывно реагировать на степень сокращения мышцы, причем пороги их возбуждения сильно различаются для активного и пассивного движений [5.122]. Более того, обнаружено, что рецепторы сухожилий могут быть возбуждены от сокращения лишь одной двигательной единицы, т. е. при очень слабом сокращении мышцы. Такая чувствительность позволяет органам Гольджи принимать участие в управлении артикуляционными движениями, но, насколько известно, прямых исследований на этот счет не проводилось. Согласно [5.93] рецепторы сухожилий дают отклик, пропорциональный как положению, так и скорости сокращения мышцы. В широком диапазоне напряжений (до 2500 г) зависимость стационарной частоты их разряда от напряжения мышцы в изометрическом режиме существенно нелинейна и может быть аппроксимирована логарифмической функцией.

Вибротактильные рецепторы (тельца Паччини). Эти рецепторы передают информацию о легком давлении, очень малых амплитудах движений и ускорениях [5.93]. Их чувствительность к амплитуде вибраций очень велика. Например, в области частот 200—400 Гц она порядка 10^{-5} мм [5.32]. Тельца Паччини представляют собой многослойные эллипсоидальные капсулы длиной 1—4 мм и обычно их находят поблизости от сухожильных рецепторов, основных нервов и кровеносных сосудов [5.43]. Идентификация этих рецепторов основывается на характерной зависимости их отклика от частоты: при раздражении поверхности кожи вибрирующими контактами большого диаметра (около 3 см²) частотная характеристика отклика телец Паччини имеет вид U-образной кривой с минимумом в области 250—300 Гц, а весь диапазон частот отклика занимает область от 40 до 800 Гц [5.158]. На частотах ниже 40 Гц в отклике преобладает реакция другого вида рецепторов, не зависящая от частоты. Этот вид рецепторов может быть возбужден при

синусоидальной стимуляции контактами малой площади (около 0,2 см²). В [5.157] было обнаружено, что тельца Паччини имеются в передней области нижней поверхности языка, тогда как в задней поверхности они не найдены. Тот факт, что тельца Паччини находятся на нижней поверхности языка, вызывает некоторое удивление и требует надлежащего физиологического объяснения. Чувствительность телец Паччини к вибрациям в широком диапазоне частот могла бы быть источником информации по крайней мере о частоте первой форманты и форме соответствующей пространственной моды (распределение давления вдоль оси речевого тракта).

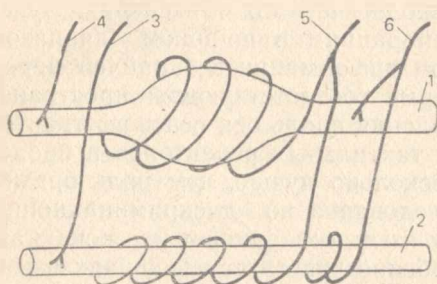
Тактильные рецепторы. Роль тактильных рецепторов в процессах речеобразования изучена несколько лучше, чем роль органов Гольджи и телец Паччини. Исследования по дискриминационной чувствительности показали, что язык соперничает с кончиками пальцев в чувствительности. Особенно чувствителен кончик языка, за ним следуют губы и нёбо [5.138]. Минимально различимое расстояние между двумя точками касания на кончике языка равно 1,1 мм [5.99], а нижней части языка — около 10 мм [5.48]. Механорецепторы твердого нёба имеют отклик в полосе частот 10—400 Гц и реагируют на смещение в 10—16 мкм [5.99]. Проводящие волокна, идущие от тактильных рецепторов, имеют большой диаметр (попадают в одну группу с волокнами α -мотонейронов) и, следовательно, обладают большой скоростью проведения сигналов возбуждения [5.128]. На vyšших обзьянах исследовались проводящие пути от рецепторов языка к коре мозга и мозжечку [5.49]. Стимуляция тактильных механорецепторов вызывает потенциалы вдоль всего пути средней лемнискально-таламокортикальной проекции язычного нерва. Эти рецепторы представлены значительным объемом ткани мозжечка. В [5.138] указывается, что тактильные рецепторы относятся к числу наиболее чувствительных. Электрическая стимуляция тактильных рецепторов достигает мозжечка через 15—20 мс и вызывает реакцию в коре больших полушарий мозга лишь ненамного позже, чем стимуляция веретен.

Отклик тактильных рецепторов пропорционален степени деформации языка и содержит, таким образом, информацию либо о касании языком другой поверхности речевого тракта, либо о его локальных деформациях. На основании этого в [5.49] делается вывод о том, что тактильные рецепторы играют решающую роль на завершающих этапах артикуляционных движений, когда касание языком нёба позволяет точнее определить его положение и форму. В связи с этим в [5.49] выдвигается гипотеза о существовании некой управляемой единицы артикуляции, которая не обязательно совпадает с какой-либо речевой единицей. Ближе всего к этой единице находится слог, оканчивающийся согласным звуком.

Моделирование системы управления артикуляцией, проведенное в [5.145], показало, что учет тактильных рецепторов существенно улучшает характеристики фрикативных звуков речи, для которых особенно высоки требования по точности артикуляции.

Веретена. Веретена являются рецепторами, включенными в цепь

управления сокращением мышц. Считается, что эти рецепторы играют основную роль в управлении мышцами. Веретена могут содержать два вида волокон — ядерно-сумчатые и ядерно-цепочечные (рис. 5.1), причем в веретенах языка преобладают ядерно-цепочечные волокна [5.99]. Веретена расположены параллельно



мышечным волокнам, и их реакция зависит от разности длин внешних (мышечных) и внутренних (принадлежащих самому веретену)

Рис. 5.1. Мышечные веретена:

1 — ядерно-сумчатое волокно; 2 — ядерно-цепочечное волокно; 3 — статическое γ -управление; 4 — динамическое γ -управление; 5 — первичные окончания; 6 — вторичные окончания

волокон. Веретена языка имеют длину несколько миллиметров и содержат от четырех до восьми внутренних волокон диаметром 6—28 мкм, заключенных в капсулу, наполненную жидкостью [5.99, 5.54]. К веретенам подходят управляющие пути (так называемое γ -управление), а от веретен отходят окончания — первичные I и вторичные II. Вторичные окончания обнаружены только у немногих из самых крупных веретен языка [5.99]. Диаметр подводящих γ -путей 2—8 мкм, и соответственно скорость проведения близка к 30 м/с [5.146]. Каждое γ -волокно иннервирует несколько веретен [5.11]. Кроме γ -волокон к веретенам могут подходить ответвления от волокон, иннервирующих α -мотонейроны. Задержка реакции сокращения веретена на управляющий импульс — примерно 6 мс, задержка отклика веретена языка в коре мозга 5—12 мс, время от сигнала веретена до реакции мышц языка — не больше 20 мс, а для мышц нижней челюсти — около 15 мс [5.37]. Общая задержка отклика мышцы на γ -возбуждение по оценке [5.146] 30—50 мс. Это время складывается из времени передачи по нервному волокну и синапсам (около 10 мс), реакции веретена и мышечного волокна. Центральная задержка для γ -мотонейронов равна 3 мс [5.27].

Отклик ядерно-сумчатых веретен пропорционален длине мышцы и скорости ее сокращения, эти веретена управляются так называемыми динамическими γ -мотонейронами. Отклик ядерно-цепочечных веретен пропорционален только длине мышцы, и они управляются статическими γ -мотонейронами. Порог чувствительности ядерно-цепочечных волокон значительно ниже, чем ядерно-сумчатых, следовательно, их волокна сокращаются быстрее. Вторичные окончания находятся главным образом на ядерно-цепочечных волокнах, а первичные окончания примерно поровну распределены на обоих видах волокон [5.122], так что возбуждение ядерно-цепочечных волокон порождает отклик и в первичных, и во вторичных окончаниях, тогда как возбуждение динамического веретена дает отклик преимущественно у первичных окончаний. Диаметр первич-

ных окончаний 12—20 мкм, а вторичных — 5—12 мкм [5.27, 5.146]. В [5.136] было показано, что различия в характеристиках отклика первичных и вторичных окончаний объясняется различием вязко-упругих свойств ядерно-сумчатых и ядерно-цепочечных волокон.

Характеристика отклика вторичных окончаний — линейная (коэффициент усиления 0,1), а у первичных окончаний — существенно нелинейная с большой крутизной в зоне малых амплитуд (коэффициент усиления порядка 100), которая становится пологой при укорочении, превышающем 0,5 мм (коэффициент усиления — порядка 0,1) [5.11]. Амплитудная характеристика первичных окончаний приближенно может быть представлена в виде логарифмической зависимости. Максимальная динамическая чувствительность веретен — около 10^{-6} мм [5.32].

Гистологические исследования установили наличие веретен в мышцах голосовых складок [5.73], мышцах гортани [5.88] и лицевых мышцах [5.87]. Веретена обнаружены в языке — немного в его кончике, больше в задней части кончика и очень много в верхней продольной мышце и поперечной мышце [5.54]. Мышцы, поднимающие нижнюю челюсть, содержат веретена, тогда как в опускающих ее мышцах веретен почти нет [5.99]. Ранее считалось, что мышцы, иннервируемые черепно-мозговыми нервами, имеют значительно меньше веретен, чем мышцы, управляемые спинномозговыми нервами. Так, по оценке [5.146] в черепных мышцах находится примерно пять веретен на 1 г мышечной ткани, тогда как в мышцах, управляемых спинным мозгом, — до 20 веретен на 1 г мышечной ткани, а гистологические исследования челюстно-подъязычной мышцы вообще не обнаружили в ней присутствия веретен. Мнение о бедности речевых мышц веретенами было опровергнуто [5.49], где на основе гистологических и электрофизиологических исследований языка высших обезьян была доказана его высокая насыщенность веретенами. Оказалось, что веретена ориентированы по всем пространственным осям, хотя и не ортогонально, а под некоторыми углами, к тому же изменяющимися при следовании вдоль языка. Кроме того, сами веретена в языке обезьян имеют менее сложную структуру, чем, например, веретена мышц руки человека — у них меньше внутренних волокон, и они не отличаются заметно по диаметру, т. е. не разделяются на два класса.

Пространственная ориентация веретен дает возможность оценки положения и деформации языка по всем направлениям (см. § 5.7). Было обнаружено, что некоторые веретена откликаются только на одно направление движений языка — вправо, влево, вверх или вниз, тогда как в большинстве случаев веретена реагируют менее специфическим образом, например уменьшая скорость импульсации при повороте налево и увеличивая — при повороте направо. Электрическая стимуляция веретен позволила определить время задержки отклика в коре полушарий мозга, оно оказалось около 4—5 мс.

Частотные характеристики первичных и вторичных окончаний показаны на рис. 5.2. Частота f_0 , на которой происходит излом характеристики, одинакова для обоих видов окончаний и равна при-

мерно 0,5—1 Гц [5.22, 5.150]. Частота и наклон частотных характеристик позволяют оценить параметры передаточных функций веретен с обоими видами окончаний. Прежде всего видно, что форма характеристик одинакова, а сами характеристики лишь смеще-

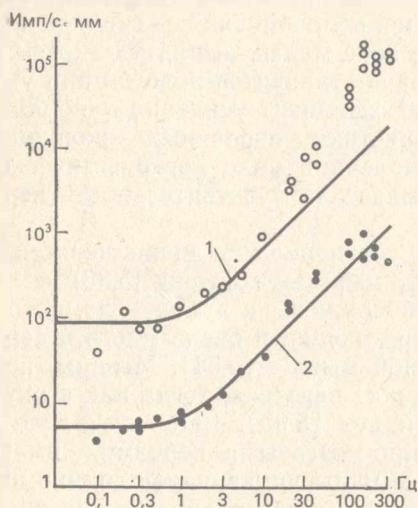


Рис. 5.2. Частотные характеристики веретен:

1 — первичные окончания; 2 — вторичные окончания

ную относительно друг друга, следовательно, и вид передаточных функций также должен быть одинаков. Различие касается лишь параметров функций, что подтверждается одинаковой частотой излома. Однако численные соотношения параметров передаточных функций в обоих случаях таковы, что чувствительность вторичных окончаний на два порядка ниже чувствительности первичных. Поэтому вторичные окончания подразделяют на «статические» и «динамические». В приведённом примере разница в коэффициентах усиления, которые определяются по отклику на нулевой частоте, составляют два, а не три порядка, которые указываются в [5.11]. Чувствительность веретен зависит от их расположения в мышце — веретена, находящиеся ближе к поверхности мышцы, имеют меньшую чувствительность, чем веретена, расположенные глубоко в мышце.

По первичным и вторичным окончаниям идут сигналы, содержащие компоненту, пропорциональную скорости сокращения мышечного волокна, поэтому с учетом инерционности сокращения внутренних волокон веретен их передаточную функцию в общем виде можно представить как реальное дифференцирующее звено (подробнее о свойствах передаточных функций см. в § 5.7):

$$W(s) = \kappa (T_1 s + 1) / (T_2 s + 1),$$

где $T_2 = \kappa T_1$. Электрическая схема этого звена показана на рис. 5.3. Постоянная времени T_1 может быть найдена по частоте f_0 , на которой амплитудно-частотная характеристика этого звена испытывает излом. Для случая, показанного на рис. 5.2, $f_0 = 0,5$ Гц и $T_1 = 1/2\pi f_0 \approx 0,3$ с. Принимая $\kappa = 0,1$, получаем $T_2 = 0,03$ с. Тогда для первичных окончаний передаточная функция

$$W_I(s) = 0,1 (0,3 s + 1) / (0,03 s + 1),$$

а для вторичных окончаний

$$W_{II}(s) = 10 (0,3 s + 1) / (3 s + 1).$$

По данным [5.27] переходные процессы ядерно-цепочечного волокна при его сокращении длятся примерно 35—100 мс, а при расслаблении — 30—450 мс (имеется в виду отклик на единичное возбуждение статического γ -мотонейрона). Указанные данные соответствуют полученным выше оценкам передаточных функций.

В откликах веретен на начало и конец мышечного сокращения иногда присутствует всплеск активности, достигающий до 0,5 амплитуды максимального отклика. Появление всплеска объясняют эффектом сухого трения, поскольку его величина зависит от начального положения мышцы и всплеск отсутствует при ритмическом раздражении или параболическом входном сигнале, имеющем постоянное ускорение [5.11, 5.27]. Однако гипотеза сухого трения весьма уязвима, так как пока не найдены механизмы, которые порождали бы этот эффект во внутренних волокнах веретен. Более правдоподобным представляется объяснение начальных и конечных всплесков активности свойствами систем с распределенными параметрами. Как было показано в гл. 1, при переменной граничных условиях в отклике растягиваемой или сжимаемой упругой системы появляется компонента, пропорциональная второй производной от смещения. Динамическая компонента возникает также и при изломе временной характеристики внешнего усилия, который часто присутствует в экспериментах с динамической нагрузкой мышц (см., например, [5.121]).

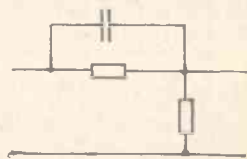


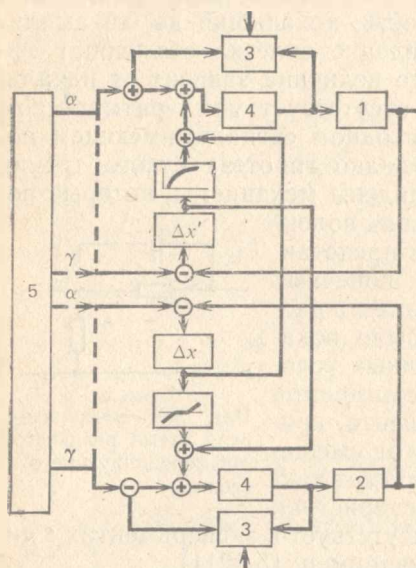
Рис. 5.3. Электрическая схема реального дифференцирующего звена

Статические окончания дают отклик как на растяжение, так и на укорочение мышечных волокон, тогда как динамические окончания здесь проявляют нелинейные свойства, реагируя только на сокращение. Следовательно, динамические веретена представляют собой существенно нелинейные системы с логарифмическим амплитудным откликом, существующим лишь в положительной области значений рассогласования между длиной внутренних волокон веретена и мышечных волокон. Как статические, так и динамические веретена являются системами с переменными параметрами, поскольку под влиянием сигналов по γ -путям изменяются их коэффициент усиления и порог срабатывания. В [1.13] показано, что γ -управление может смещать рабочую точку на характеристике мышцы «напряжение — смещение» и, таким образом, адаптироваться к различным условиям работы, удовлетворяя различным критериям оценки качества системы управления.

Рассмотрим теперь несколько подробнее схему включения веретен в систему управления мышцами (рис. 5.4). Из центральной нервной системы к мышце идут три управляющих пути: от α -мотонейронов, от статических γ -мотонейронов и от динамических γ -мотонейронов, причем примерно в 10% случаев α -пути ответвляются на управляющие входы веретен (так называемое β -управление). Сигналы обратной связи от веретен суммируются с α -управлением на α -мотонейронах мышцы и (через вставные нейроны) ее антаго-

нистами [5.150]. Разновидности γ -управления подразумевают возможность раздельного управления положением и скоростью сокращения мышц, реализуя различные критерии управления в зависимости от решаемой задачи.

Во временной области отклик динамического веретена



$$x_d(t) = \begin{cases} \kappa_d \lg \frac{d \Delta x}{d'}, & \Delta x > a_d; \\ 0, & \Delta x \leq a_d, \end{cases}$$

где Δx — сигнал рассогласования между заданной длиной мышцы $x_{вх}$ и фактической ее длиной $x_{вых}$; k_d — коэффициент усиления; a_d — динамический порог срабатывания. Для статического отклика веретена

$$x_{\text{CT}}(t) = \begin{cases} \kappa_{\text{CT}} \Delta x, & \Delta x > a_{\text{CT}}; \\ 0, & \Delta x \leq a_{\text{CT}}, \end{cases}$$

где $a_{\text{ст}}$ — статический порог срабатывания.

При превышении порогов срабатывания обоих веретен в первичном окончании сигнал имеет компоненты

Рис. 5.4. Схема управления двигательной единицей:

1 — физические мышечные волокна; 2 — тонические мышечные волокна; 3 — клетки Реншоу; 4 — α -мотонейрон; 5 — центральная нервная система

$$x_I(t) = \kappa_0 \lg \frac{d\Delta x}{dt} + \kappa_2 \Delta x, \quad (5.1)$$

а во вторичном окончании

$$x_{II}(t) = \kappa_3 \lg \frac{d\Delta x}{dt} + \kappa_4 \Delta x. \quad (5.2)$$

В этих формулах $\kappa_1 \gg \kappa_3$ и $\kappa_4 \gg \kappa_2$. Учитывая данные [5.149] о том, что до частот 6 Гц α -мотонейрон не искажает динамики входного сигнала, допустим, что α -мотонейрон играет роль чисто пропорционального звена с коэффициентом усиления κ_α . Тогда при $\Delta x > a_{\text{ст}}$, $d\Delta x/dt > a_{\text{д}}$ на мышечное волокно поступает сигнал

$$x(t) = \kappa_\alpha \left(x_\alpha + \kappa_1 \lg \frac{d\Delta x}{dt} + \kappa_2 \Delta x \right). \quad (5.3)$$

Если взять линейный участок характеристики мышцы, то ее отклик на воздействие сигналов вида (5.3) может быть представлен как суперпозиция откликов двух систем: системы с разомкнутым управлением $x_1(t) = F_1(x_\alpha, \kappa_\alpha)$ и замкнутой системы с отрицательной обратной связью $x_2(t) = F_2(\Delta x, k_1, k_2, k_3, k_4, \kappa_\alpha)$:

$$x_{\text{BBLX}}(t) = x_1(t) + x_2(t).$$

Такое комбинированное — замкнутое и разомкнутое — управление мышечным сокращением обеспечивает гибкость при управлении,

позволяя для каждой конкретной задачи выбрать оптимальное соотношение вкладов каждого вида систем. Это свойство, в частности, объясняет большую устойчивость системы управления артикуляцией при искусственном нарушении обратной связи (см. § 5.4).

В действительности мотонейрон обладает порогом срабатывания, т. е. является нелинейным звеном. Его линеаризация понадобилась нам для того, чтобы в наиболее отчетливой форме выявить наличие комбинированной структуры системы управления сокращением мышечных волокон. Нелинейность мотонейрона не изменяет этой структуры, а лишь влияет на динамику процессов управления, существенная нелинейность которых является принципиально важным свойством системы управления движениями вообще и артикуляции в частности. Игнорирование нелинейности системы управления движениями может привести к физиологически нереальным моделям, тогда как учет нелинейности дает возможность описать целый ряд явлений (см. § 5.6) [5.81].

Довольно редко встречающееся β -управление представляет собой чистый вид автоматического регулирования по рассогласованию. В этом случае отсутствует возможность коррекции характеристик рецепторов со стороны центральной нервной системы и, следовательно, диапазон использования β -управления ограничен. γ -управление обеспечивает приспособление характеристик цепи обратной связи к особенностям задачи, решаемой в данный момент системой управления движениями. Обычно сигналы α - и γ -управления возникают одновременно. В особых случаях при совершении ритмических движений активность γ -путей создается с опережением α -активности, как это наблюдается в управлении дыхательными мышцами [5.61], а также при очень медленных движениях, в случае анестезии, у децеребрированных кошек [5.74].

Частотные и динамические характеристики веретен распределены в довольно широком диапазоне. Вероятно, это дает возможность улучшить точность работы каждого веретена за счет сокращения диапазона его входных и выходных сигналов. Но эти различия в характеристиках требуют отдельного управления различными группами веретен, т. е. формирования различных γ -сигналов, так что γ -управление предлагает гораздо лучшие характеристики системы управления, чем β -управление [5.89].

Система управления артикуляцией обладает свойствами, отличными от свойств системы управления другими движениями [5.128]. Например, установлено, что, в отличие от локомоции, рефлексорные пути в речеобразовании не работают: стретч-рефлекс не обнаружен ни в движениях языка, ни в движениях нижней челюсти [5.37]. Иными словами, обратная связь от веретен (и от других механорецепторов) замыкается не на α -мотонейронах, а на более высоких уровнях системы управления артикуляцией, что позволяет ввести между выходом рецептора и входом α -мотонейрона дополнительный блок обработки сигналов обратной связи с целью улучшения характеристик системы управления.

5.3. СКОРОСТЬ РАСПРОСТРАНЕНИЯ ВОЗБУЖДЕНИЯ

Скорость распространения нервных импульсов (возбуждения) зависит от диаметра проводящих путей. Для двигательных нервов она равна 50—70 м/с, а для γ -волокон — 15—50 м/с [5.11]. Скорость распространения сигналов в окончаниях двигательных нервов, расположенных внутри мышц, на порядок меньше: 2—5 м/с [5.148], что объясняется ветвлением нервных волокон (после каждого ветвления диаметр волокон уменьшается), а также отсутствием у нервных окончаний миелиновой оболочки. Каждое γ -волокно иннервирует несколько мышечных волокон более или менее синхронно. Используя свойства проведения сигналов в отдельной двигательной единице, удалось построить методику измерения скорости распространения сигналов в мышце с помощью

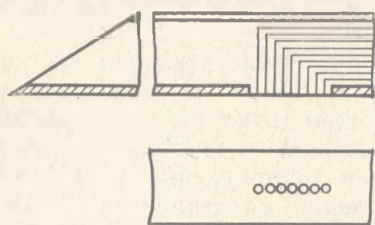


Рис. 5.5. Многоканальный электрод для измерения скорости распространения возбуждения

многоканальных электродов [5.148]. Многоканальный электрод (рис. 5.5) позволяет отыскать импульсы, принадлежащие одной и той же двигательной единице, и, так как расстояние между электродами известно, а разница во времени появления этих импульсов может быть измерена, вычислить скорость распространения не представляется трудной задачей. Амплитуда, форма и длительность импульсов, принадлежащих одной двигательной

единице, очень близки, поэтому распознавание импульсов осуществляется путем их визуального сравнения.

На скорость распространения сигналов в мышечном волокне влияют такие факторы, как утомление мышцы вследствие длительного напряжения, кровоснабжение, температура, возраст человека. Так, в [5.148] установлено, что в результате длительного напряжения скорость распространения 3,09—3,37 м/с через 30 мин упала примерно вдвое. Прекращение кровоснабжения мышцы путем пережатия кровеносных сосудов через 20—40 мин также приводит к падению скорости распространения примерно вдвое. Исследование физиологических процессов, лежащих в основе таких изменений, показало, что причиной служит нарушение процесса обмена веществ [5.126]. В следующих параграфах будет показано, что для правильной интерпретации электромиограмм и анализа качества системы управления артикуляцией необходимо знать скорость распространения сигналов и общую задержку сигналов управления в нервных цепях. Были проведены специальные эксперименты по измерению скорости распространения возбуждения в подбородочной мышце, управляющей подъемом нижней губы. Для каждой обнаруженной двигательной единицы производилось усреднение интервалов между импульсами каждые 2 с, при этом поддерживалось минимальное напряжение этой мышцы с тем, чтобы зафиксировать

разряды только от одной двигательной единицы. Всего было исследовано 13 двигательных единиц, распределение скоростей распространения в которых находилось в диапазоне 1,85—5,66 м/с, а средняя скорость по всем двигательным единицам была равной 4,1 м/с.

Несмотря на то, что длина нервных волокон, расположенных внутри мышц, значительно меньше длины путей, связывающих мышцу с центрами управления, малая скорость распространения по внутримышечным нервным волокнам может привести к задержкам, сравнимым с общей задержкой сигналов управления. Средняя задержка на синапсах в мышце равна 5 мс, т. е. сравнима со временем пробега сигнала управления от α -мотонейрона до мышцы. Скорость распространения сигналов по внешним нервным волокнам обычно остается постоянной, тогда как скорость распространения для внутримышечных волокон, как мы видели, может вдвое уменьшиться. Следовательно, соответствующее увеличение задержки может поставить в некоторых случаях систему управления на грань устойчивости и даже привести к автоколебаниям.

5.4. ОБРАТНАЯ СВЯЗЬ

Из теории автоматического регулирования известно, что обратная связь улучшает динамические характеристики системы управления, повышает точность и помогает компенсировать возмущения, вносимые внешней средой и изменением параметров системы. Существуют глобальные и локальные обратные связи. Глобальная обратная связь заключается в подаче сигнала с выхода системы на ее вход, а локальная обратная связь охватывает отдельные звенья внутри системы управления. Глобальная обратная связь оценивает окончательный результат действий системы и сравнивает его с заданной программой. Если обнаруживаются различия, то формируется сигнал рассогласования, который используется для коррекции движений. Глобальная обратная связь должна быть отрицательной, поскольку положительная обратная связь увеличивает отклонение от заданного закона поведения объекта управления и создает неустойчивость. В зависимости от коэффициента усиления в цепи обратной связи она может быть более или менее глубокой, и при коэффициенте усиления, равном нулю, обратная связь прерывается и система управления превращается в разомкнутую. Преимуществом разомкнутой системы управления является ее устойчивость, тогда как замкнутые системы склонны к неустойчивости. Это обстоятельство особенно важно для анализа системы управления движениями вообще и речи в частности, так как задержки в нейронных цепях уменьшают запас устойчивости.

Впервые представление о процессах речеобразования как результате действий системы управления с обратной связью было развито в [5.63]. Недоступность высших уровней речевой системы

управления для исследований привела к созданию множества абстрактных лингвистических и физиологических моделей, лишь косвенно отражающих истинные характеристики этой системы. Ни одна из этих моделей не сумела объяснить все особенности процессов речеобразования, что вполне естественно, поскольку трудно только из наблюдений за выходом системы управления — артикуляцией и акустическим сигналом — вывести обоснованные заключения об относительной роли центральной нервной системы и периферической обратной связи, причем имеется опасность отделить функционирование периферической системы от процессов речеобразования. Поэтому особенную важность приобретает математическое моделирование, позволяющее использовать знания из других областей, в частности теории автоматического управления.

Разнообразные виды обратной связи в речеобразовании можно разделить на две группы: акустическую обратную связь, оценивающую акустические характеристики речевого сигнала, и механическую обратную связь, контролирующую движения артикуляторных органов и форму речевого тракта. Датчиками механической (проприоцептивной) обратной связи служат всевозможные рецепторы, часть которых была описана в § 5.2. Основными источниками сведений о возможной роли обратной связи в процессах речеобразования служат клинические исследования заболеваний мозга, острые опыты на животных по исключению обратной связи путем перерезки нервных каналов (деафферентации), искажения и маскировка акустического сигнала шумом, анестезия ротовой полости и создание всевозможных помех процессу речеобразования.

Метод анестезии основан на том, что нервные волокна меньшего диаметра, передающие, например, сигналы γ -управления к вставкам, поддаются анестезии быстрее и угнетаются дольше, чем волокна большого диаметра. Анестезия поверхности языка или нервов, управляющих речевым аппаратом, в течение некоторого времени широко применялась для выяснения роли механической обратной связи. Анализ результатов этих исследований выявил некоторые методологические недостатки, свойственные этому методу. Во-первых, при анестезии неизвестно, какие виды обратной связи и в какой степени угнетены. Во-вторых, остается неизвестной и степень подавления сигналов из центральной нервной системы. Так, имеются свидетельства ослабления челюстно-подъязычной мышцы при анестезии подбородочного нерва [5.36] и ее влияния на центральные механизмы при наркозах в клинических условиях. Кроме того, влияние анестезии на речь разных людей неодинаково, вплоть до полной невосприимчивости, что может быть следствием разной чувствительности к используемому препарату, различий в анатомии и разброса в глубине и месте анестезии. Нередко использовавшаяся в этих экспериментах оценка влияния анестезии на слух недостаточна, поскольку имеются доказательства существования артикуляционной компен-

сации. Например, в [5.35] было найдено, что при анестезии подбородочной ветви тройничного нерва движения нижней челюсти были нарушены, а речь воспринималась как нормальная. При этом подъем нижней челюсти уменьшился в среднем на 3 мм, но это было компенсировано увеличенной амплитудой движения губ.

Анестезия влияет на пассивные и активные движения по-разному. Установлено, что при анестезии γ -путей при пассивном растяжении веретена ведут себя так же, как и до анестезии, но при произвольных сокращениях α - и γ -отклики уменьшаются и изометрическое сокращение ослабляется [5.74].

Результаты экспериментов по анестезии дали основание утверждать что, хотя достигаемые эффекты выражены довольно слабо, а иногда они противоречивы, анестезия уменьшает скорость и увеличивает амплитуду движений. Похоже, что периферическая обратная связь важна для управления не столько временными характеристиками, сколько положением артикуляторных органов [5.138]. В [5.85] отмечается, что тактильно-кинестетические помехи с помощью анестезии мало влияют на разборчивость отдельных слов и связной речи. Однако люди с врожденными сенсорными дефектами иногда обладают очень неразборчивой речью, так что не исключено, что при анестезии играет роль прошлый опыт сенсорных связей. При общей малой повреждаемости речи согласные все же повреждаются сильнее, чем гласные, хотя между исследователями нет единодушия о преобладающем типе повреждений — указываются ошибки типа замен, изменения места артикуляции или искажений внутри одного и того же класса фонем. При анестезии подбородочного нерва длительность согласных звуков увеличивается, хотя и незначительно — от 1 до 20 мс [5.82]. Изменения внутриротового давления также невелики и не обнаруживают какой-либо четкой тенденции, за исключением назальных звуков, для которых давление заметно повышается.

Картина нарушений при анестезии детей оказывается в общем такой же, как и для взрослых, т. е. удлиняются звуки, главным образом согласные, тогда как гласные даже укорачиваются. Особенно удлиняются фрикативные и прежде всего *s*. Единственное отличие состоит в большей повреждаемости гласных, артикуляция которых приближается к нейтральной позиции.

Наибольшая повреждаемость *tʃ*, *dʒ*, *z*, *ʃ*, *t*, *l* и особенно *s* отмечается при двусторонней анестезии подбородочного [5.47] и тройничного [5.48] нервов. При этом в [5.47] все ошибки являются ошибками места артикуляции с наиболее частым переходом *s* — *ʃ*, хотя пересечение границ незначительно. Интересно отметить, что в этих экспериментах не наблюдались ошибки ни для *d*, ни для *n*. В [5.48] параллельно с анестезией проводились измерения ЭМГ, в которых было обнаружено, что активность че-

люстно-подъязычной мышцы и передней головки двубрюшной мышцы резко падала, а активность некоторых других мышц, например подбородочно-язычной, повышалась. Тот факт, что в этих экспериментах больше всего повреждаются фрикативные звуки, свидетельствует о наиболее высоких требованиях к точности системы управления артикуляцией этих звуков.

Обстоятельные исследования влияния анестезии на процессы артикуляции были проведены в [5.76], где в различных комбинациях применялись поверхностная анестезия языка и блокировка язычного нерва наряду с бинауральной маскировкой белым шумом высокого уровня (100 дБ). Одновременно регистрировались спектрограммы и палатограммы. Наибольшее влияние проявилось при совместном воздействии всех факторов. Оно состояло в увеличении длительности звуков, большей площади контакта языка с нёбом, сужении центрального желобка языка и сдвиге места артикуляции вперед. На основании результатов этих экспериментов был сделан вывод о том, что для временной регуляции сегментов речи и площади контакта языка с нёбом скорее важна не тактильная рецепция, а акустическая обратная связь. Однако на наш взгляд малое влияние нарушений отдельных видов обратной связи указывает на высокую надежность и гибкость комплекса различных видов обратных связей, обеспечиваемых разными типами рецепторов. Наблюдения за клиническими случаями нарушений речевой деятельности, в которых происходит обширное нарушение механизмов обратной связи на высших уровнях, демонстрируют невозможность нормальной реализации процесса речеобразования без обратных связей. Кроме того, как отмечается в [5.37], тот факт, что механорецепторы языка не включены в рефлекторную цепь в процессах речеобразования, объясняет малую эффективность анестезии язычного и тройничного нервов.

Создание механических помех движениям артикуляторных органов позволяет оценить такую характеристику обратной связи, как ее задержка. Увеличение сопротивления воздушному потоку во время фонации через 30—80 мс приводит к сокращению внутренних межреберных мышц, управляющих увеличением давления в легких [5.142]. В других экспериментах время задержки обратной связи оказывается гораздо меньше. Например, при создании сопротивления закрытию губ при артикуляции губных взрывных увеличение мышечного напряжения задерживается всего лишь на несколько миллисекунд [5.147]. В аналогичных экспериментах получены задержки порядка 15—20 мс [5.64]. В экспериментах по внезапному растяжению углов [5.113] губ было установлено, что рефлекторная активность губных мышц появляется через 14—17 мс, а амплитуда рефлекса зависит от логарифма скорости растяжения (или отпускания), что, кстати, отличается от рефлекса на растяжение, наблюдаемого у животных, где амплитуда отклика зависит от амплитуды смещения.

Акустическая обратная связь, в отличие от проприоцептивной,

является глобальной, поскольку она оценивает конечный результат процесса речеобразования — акустические характеристики речевого сигнала. Звуки, произносимые человеком, достигают слухового анализатора двумя путями: по воздуху и за счет костной проводимости. Поскольку скорость распространения звука по костям больше скорости звука в воздухе, в слуховом анализаторе при прослушивании собственной речи должно возникать наложение акустических сигналов со сдвигом около 0,2 мс.

Роль акустической обратной связи обнаруживается при анализе речи глухих, у которых создается своя, специфическая артикуляторная система, отличная от артикуляции людей с нормальным слухом. Некоторые признаки звуков для глухих остаются неразличимыми, и тогда эти звуки появляются в их речи, случайно чередуясь [5.129]. Попытки подавления акустической обратной связи с помощью шума или ускорения темпа речи обычно не оказывают заметного влияния [5.135].

Взаимосвязь процессов речеобразования и восприятия сказывается во взаимном влиянии длительного проговаривания каких-либо звуко сочетаний на их восприятие и длительного прослушивания звуко сочетаний на их характеристики при речеобразовании [5.55, 5.56]. Обнаружено, например, что после прослушивания серии слогов *ph*i и *bi* (английский язык) диктор воспроизводит слоги *pi* и *ti* с меньшим временем между взрывом смычки и появлением основного тона, т. е. они приближаются по характеристикам к звонким взрывным, тогда как в обратной ситуации характеристики *bi* и *di* не меняются. Границы восприятия естественных и синтетических звуков также сдвигаются в сторону сужения после их длительного прослушивания. Это объясняется в [5.55, 5.56] утомлением тех отделов системы управления речью, которые заведуют артикуляцией и проприоцепцией.

Некоторые количественные оценки характеристик акустической обратной связи получены в экспериментах по прослушиванию собственной речи с задержкой во времени. Впервые такие эксперименты были описаны в [5.100]. Было обнаружено, что при увеличении времени задержки до 80 мс человек начинает «заикаться», повторяя отдельные фонемы, а при задержке до 1 с повторяет целые слова. Максимальный эффект возникает при задержке 200 мс, причем этот интервал не зависит от длины слова, т. е. контекста [5.84]. Напротив, речь заикающихся людей становится более плавной при задержке акустического сигнала или маскировке звука белым шумом [5.66]. Это явление хорошо известно в теории автоматического регулирования — увеличение задержки в системе обратной связи сначала вызывает неустойчивость, а при дальнейшем увеличении задержки система вновь возвращается в область устойчивости. Искусственно вызванное заикание сопровождается явлениями, отличающимися от явлений, наблюдающихся во время естественного заикания. В частности, регистрация электрической активности мышц гортани в экспериментах с задержкой восприятия показывает возникновение коле-

баний мышечной активности, не похожих на колебания во время естественного заикания [5.60, 5.67]. Вместе с тем схожесть внешнего эффекта позволяет проводить далеко идущие аналогии и относительно механизмов возникновения естественного заикания.

Как было показано в гл. 3, ширина полосы частот АЧХ передаточных функций артикуляторных органов равна 5—10 Гц. По оценке [5.37] максимально допустимое время задержки, при превышении которой в системе возникают автоколебания, составляет 58 мс для полосы шириной 8 Гц и 41 мс для полосы шириной 10 Гц. К сожалению, не указывается, для какого коэффициента усиления получены эти оценки, но принимая их как ориентировочные, нетрудно заметить, что запас устойчивости в системе артикуляции относительно времени задержки не очень велик — порядка 10—30 мс. Этот запас устойчивости может быть исчерпан при уменьшении скорости распространения потенциалов в мышечных волокнах (см. § 5.3), увеличении центральной задержки или изменении коэффициента усиления в системе управления. Физиологический тремор с частотами осцилляции 8—19 Гц в [5.149] связывается с задержками 10—50 мс в цепи обратной связи, включающей спинной мозг, тогда как патологический тремор (например, при паркинсонизме) с частотами 4—6 Гц связывают с задержками 50—120 мс в центральной нервной системе. Задержка возбуждения внутренних мышц может быть на 30 мс больше, чем лицевых мышц [5.91].

В многоконтурной системе, какой является система управления артикуляцией, где требуемая конфигурация речевого тракта может быть создана активизацией мышц в разной степени, обратная связь может замыкаться на другом контуре вместо уничтожения сигнала рассогласования в каком-либо данном контуре. В этом случае глобальная обратная связь по-прежнему остается отрицательной, но локально сигнал рассогласования может приводить к увеличению отклонения в другом контуре [5.62, 5.81]. Например, при ограничении движения нижней челюсти или внесении в это движение внезапных помех не увеличивается активность мышц, которая компенсировала бы эти помехи, а повышается активность мышц, управляющих движениями губ [5.64]. Этот тип обратной связи, по-видимому, доминирует в процессах управления артикуляцией, создавая значительные трудности для аналитической оценки области устойчивости и качества процессов управления.

5.5. ЭЛЕКТРИЧЕСКОЕ ПОЛЕ ПОТЕНЦИАЛОВ МЫШЕЧНОГО ВОЛОКНА

Измерения электрической активности мышц (так называемой электромиографии или, сокращенно, ЭМГ) дают важную информацию о процессах управления сокращением мышц.

Эту информацию можно использовать для оценки механизмов управления артикуляцией только в том случае, если соблюдена высокая точность измерения ЭМГ, исключающая неправильную интерпретацию формы управляющих сигналов. В противном случае возможны ошибочные заключения не только об отдельных характеристиках системы управления, но и о принципиально важных ее свойствах — таких, как критерии качества регулирования [5.81]. В измерениях обычно используют пару электродов (биполярное отведение), либо расположенных на коже (поверхностное отведение), либо внедренных в мышцу в виде тонких (диаметром около 50 мкм) проволочек. Поверхностные электроды не создают болезненных ощущений, но ими нельзя пользоваться при исследовании глубоко расположенных мышц вследствие помех, создаваемых другими, находящимися ближе к поверхности мышцами.

Для правильной интерпретации ЭМГ сигналов нужно иметь представление об электрических процессах в мышце и о влиянии того или иного метода измерения на свойства получаемых сигналов. Такой анализ был проведен в [5.50, 5.107—5.109] в предположении, что мышечная ткань однородна и изотропна по своим электрическим свойствам, диаметр мышечного волокна не изменяется по его длине, электроды расположены вдоль волокна, вдали от его окончания и точки иннервации. Скорость распространения потенциалов по мышечному волокну очень мала (см. § 5.3), поэтому электрическое поле можно считать стационарным. Тогда для одиночного волокна с потенциалом $\Phi(t)$ частотные характеристики электрического поля $\Psi(j\omega)$ описываются уравнением Лапласа $\nabla^2 \Psi(j\omega) = 0$, где ∇^2 — оператор Лапласа. Принимая сферическую симметрию поля с граничными условиями на расстоянии $r \rightarrow \infty$ как $\lim_{r \rightarrow \infty} \Psi(j\omega) = 0$, получаем решение

$$\Psi(j\omega) = S_0(j\omega) \left\{ \int_{-\infty}^0 [(x-x_0)^2 + h^2]^{-\frac{1}{2}} e^{-j\frac{\omega x}{v}} dx + \right. \\ \left. + \int_0^{\infty} [(x-x_0)^2 + h^2]^{-\frac{1}{2}} e^{-j\frac{\omega x}{v}} dx \right\}, \quad (5.4)$$

где x — координата вдоль волокна; x_0 — координата точки отведения потенциала; h — кратчайшее расстояние от электрода до волокна; v — скорость распространения потенциала по волокну; ω — круговая частота; S_0 — амплитудно-частотная характеристика электрического источника. Две компоненты в (5.4) соответствуют волнам, распространяющимся по волокну в положительном и отрицательном направлениях от точки иннервации. Если электрод расположен достаточно далеко от точки иннервации, т. е. $x_0 \rightarrow \infty$, то (5.4) принимает вид

$$\Psi(j\omega) = \frac{K_0 \left(\frac{\omega h}{v} \right)}{K_0 \left(\frac{\omega a}{v} \right)} e^{-j\frac{\omega x_0}{v}} \Psi_a(j\omega), \quad (5.5)$$

где K_0 — модифицированная функция Бесселя 2-го рода нулевого порядка; a — радиус волокна; $\Psi_a(j\omega)$ — спектр сигнала на поверхности волокна. Отражение на границе кожа — воздух может быть учтено наложением симметричного поля, но оно оказывает малое влияние на вид решения.

Если отведение потенциалов производится с помощью пары электродов, разнесенных на расстояние $2d$, то разностное поле

$$\Delta \Psi(j\omega) = \Psi(j\omega)|_{x_0-d} - \Psi(j\omega)|_{x_0+d}.$$

Для поверхностных электродов, находящихся вдали от точки иннервации,

$$\Delta \Psi(j\omega) = 2j \sin \frac{\omega d}{v} \frac{K_0\left(\frac{\omega h}{v}\right)}{K_0\left(\frac{\omega a}{v}\right)} e^{-j \frac{\omega x_0}{v}} \Psi_a(j\omega). \quad (5.6)$$

Характеристики поля, измеряемого проволочными электродами, аналогичны (5.6). Как видно из сравнения (5.5) и (5.6), биполярное отведение вносит искажение в измеряемый сигнал в виде фильтровой функции $\sin \omega d/v$, которая создает нули (или минимумы) в спектральной огибающей на частотах $\omega d/v = = k\pi$, $k=0, 1, 2, \dots$ Этим свойством можно воспользоваться для измерения скорости распространения потенциала по частоте f_1 первого минимума в измеренном спектре: $v=2df_1$.

Вследствие конечной длины волокна вдоль него устанавливаются стоячие волны (рис. 5.6). Ошибки в оценке электрического поля, возникающие по этой причине, определяются как

$$|\varepsilon| < \left[\frac{\omega l}{v} K_0\left(\frac{\omega h}{v}\right) \right]^{-1},$$

где l — расстояние между электродом и концом волокна.

Рассматривая поле, создаваемое множеством волокон одной и той же двигательной единицы, записываем

$$\Phi_{DE}(t) = \sum_{n=1}^N \Phi(t - \tau_n, h_n, v_n, a_n),$$

где τ_n — время прибытия потенциала в n -м волокне; h_n — расстояние от электрода до n -го волокна, v_n и a_n — скорость распространения импульса и радиус n -го волокна. Распределение интервалов между импульсами, регистрируемыми от одной и той же двигательной единицы, при малых частотах следования импульсов обычно несимметрично [5.24] и лучше всего аппроксимируется распределением Вейбеля [5.111]

$$p(\tau) = \frac{\kappa}{\beta} \left(\frac{\tau - \tau_0}{\beta} \right)^{\kappa-1} e^{-\left(\frac{\tau - \tau_0}{\beta} \right)^{\kappa}},$$

где $p(\tau)$ — плотность вероятности; κ, β, τ — параметры распределения. При уменьшении периода следования импульсов распределение становится более симметричным и может быть аппроксимировано гауссовским распределением [5.24]. Если же принять во внимание, что обычно одна двигательная единица содер-

жит несколько десятков волокон, то согласно закону больших чисел часто считают, что распределение времени появления импульсов есть гауссово. Тогда измеренный биполярными электродами спектр электрического поля двигательной единицы, содержащей N волокон, по [5.50]

$$\Psi_{DE}(j\omega) = 2NM \left[\sin \frac{\omega d}{v} \right] \frac{K_0 \left(\frac{\omega \bar{h}}{v} \right)}{K_0 \left(\frac{\omega \bar{a}}{v} \right)} e^{-\frac{\sigma_\tau^2 \omega^2}{2}} \Psi_a(j\omega),$$

где $M[\cdot]$ — математическое ожидание; $\bar{v}$, $\bar{h}$, $\bar{a}$ — средние величины скорости распространения, расстояния до электродов и радиуса волокон; σ_τ — дисперсия времени появления импульсов, которая определяется из выражения

$$\sigma_\tau^2(j\omega) = 4NM \left[\sin^2 \frac{\omega d}{v} \right] \frac{K_0 \left(\frac{\omega \bar{h}}{v} \right)}{K_0 \left(\frac{\omega \bar{a}}{v} \right)} e^{-2(\sigma_\tau \omega)^2} \Psi_a(j\omega) - M^2 \left[\sin \frac{\omega d}{v} \right] e^{-(\sigma_\tau \omega)^2}. \quad (5.7)$$

Обычно принимается, что вследствие дисперсии длительность сигнала от двигательной единицы больше, чем от единичного волокна, и находится в диапазоне 1—13 мс [5.112], причем эта длительность укладывается примерно в шесть стандартных отклонений σ_τ , где $0,5 \leq \sigma_\tau \leq 2$ мс. Эта дисперсия приводит к относительному уменьшению огибающей спектра $\Psi_{DE}(j\omega)$ на высоких частотах — область влияния статистики межимпульсных интервалов находится в полосе 0—40 Гц [5.112], за пределами которой вид спектра определяется характеристиками самих импульсов. Увеличение числа волокон в двигательной единице повышает уровень низкочастотных компонент [5.108].

Допущение гауссова распределения сильно упрощает аналитические построения

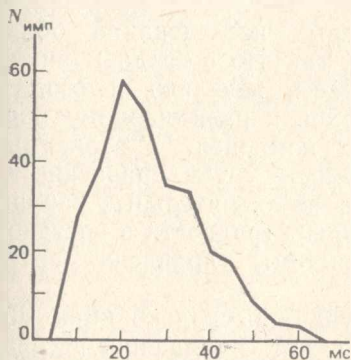


Рис. 5.7. Гистограмма межимпульсных расстояний для нижней кольцевой мышцы губы (нормальный темп артикуляции)

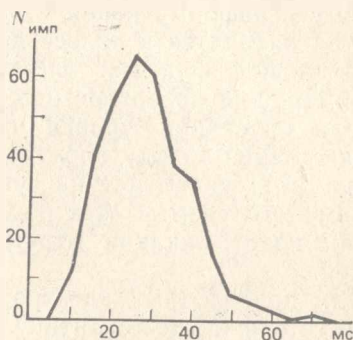


Рис. 5.8. Гистограмма межимпульсных расстояний для нижней кольцевой мышцы губы (ускоренный темп артикуляции)

ния, но нужно помнить о приближенном характере получаемых решений. В наших экспериментах по отведению потенциалов от нижней кольцевой губной мышцы с помощью поверхностных электродов производилось построение гистограмм межимпульсных расстояний для нормального и ускоренного произнесения бессмысленных звуко сочетаний типа БАБОБА и т. д. Как видно из рис. 5.7 и 5.8, для нормального темпа артикуляции гистограмма весьма несимметрична, тогда как для ускоренного темпа максимум сдвигается в сторону больших расстояний, вследствие чего гауссово распределение для этого случая будет более приемлемым, чем для предыдущего.

При анализе поля множества двигательных единиц в [5.108] принимается, что их активность некоррелирована, и применим квадратичный закон суммирования. В результате суммарный спектр мощности сигналов, измеряемых поверхностными биполярными электродами, представляется как

$$W(\omega) = \frac{C(\omega)}{2} [H(\omega, h_{min}) - H(\omega, h_{max})],$$

$$\text{где } C(\omega) = 4 W_a(\omega) \sin^2 \frac{\omega d}{v} Q(\omega) \left(\frac{v}{\omega a} \right)^2 \frac{1}{K_0 \left(\frac{\omega a}{v} \right)};$$

$$H(\omega, h) = \left(\frac{\omega h}{v} \right)^2 \left[K_1^2 \left(\frac{\omega h}{v} \right) - K_0^2 \left(\frac{\omega h}{v} \right) \right].$$

Здесь $W_a(\omega)$ — спектр мощности электрического сигнала; h_{min} — кратчайшее расстояние от электрода до мышцы; h_{max} — расстояние от электрода до дальней границы мышцы; K_1 — модифицированная функция Бесселя 2-го рода первого порядка; $Q(\omega)$ — суммирующая функция; $Q(\omega) = 1 + (N-1)e^{-(\sigma\tau\omega)^2}$, где N — среднее число волокон в двигательной единице; σ_τ — среднее стандартное отклонение времени прибытия импульсов.

Проведенный анализ характеристик измеряемого электрического поля потенциалов мышечного волокна показывает, что влияние вида и расположения электродов начинает сказываться главным образом на частотах выше 10 Гц. В гл. 3 мы видели, что наиболее информативная для речевых исследований область спектра находится в полосе до 10 Гц, так что в среднем влияние фильтрующей функции электродов здесь невелико и сводится, например, для поверхностных электродов к равномерному подъему или опусканию уровня частотных компонент в зависимости от расстояния между электродами. Вместе с тем уравнение дисперсии (5.7) указывает на существование значительных вариаций в форме измеряемых спектров, что создает трудности в постановке задачи идентификации параметров системы управления артикуляцией.

Для проволочных электродов ситуация несколько иная. Дело в том, что в процессе артикуляции за счет сокращения исследуемой мышцы расстояние между электродами может меняться синхронно с движениями, изменяя тем самым соотношения между частотными компонентами спектра измеряемых сигналов, например увеличение расстояния приводит на малых частотах к пропорциональному изменению наклона спектра. Устранить это явле-

ние можно за счет предлагаемого Хики склеивания проволочных электродов таким образом, чтобы расстояние между ними всегда оставалось неизменным. В этом случае, так же как и для поверхностных электродов, можно произвести коррекцию наклона спектра измеренного сигнала по известной величине межэлектродного расстояния.

Способы обработки измеренных электромиограмм для получения наиболее достоверной информации о процессах управления, в том числе артикуляцией, давно служат предметом исследований. К настоящему времени почти повсеместно применяют так называемый метод интегрирования (или сглаживания), при котором отрицательные значения измеренного сигнала выпрямляют, а затем подвергают интегрированию с постоянной времени 20—40 мс. Учитывая нестабильность измерений, связанную со статистической природой распределения возбуждающих импульсов в пространстве и времени, следует усреднять результаты измерений не менее чем по 20 реализациям одного и того же звуко сочетания. Измеренный таким образом сигнал учитывает как изменение частоты импульсации, так и вовлечение в действие более крупных моторных единиц, порождающих импульсы с большей амплитудой. Однако здесь не происходит точной демодуляции ни частотно-импульсной, ни амплитудно-импульсной модуляции, с помощью которой кодируется сигнал, управляющий сокращением мышц.

Для более строгого подхода к измерению формы нейромоторных команд неоднократно пытались использовать разнообразные счетчики импульсов. В [5.86] предлагается определять число импульсов в скользящем окне длиной 170 мс, причем для повышения помехоустойчивости вводится зона нечувствительности к импульсам с малой амплитудой. Утверждается, что средний коэффициент корреляции сигнала при таком способе измерений почти в 3 раза выше, чем при интегрировании, и форма сигналов становится более устойчивой от одной реализации звуко сочетания к другой. Вместе с тем способ подсчета числа импульсов подвергается критике за игнорирование информации, содержащейся в амплитуде интерференционных сигналов, поэтому он не находит широкого применения в практике речевых исследований.

Рассмотрим несколько подробнее характер искажений временных параметров измеряемых сигналов, создаваемых при частотно-импульсной модуляции нейромоторных сигналов и обработке ЭМГ. Существуют два вида модуляций, посредством которых кодируются нейромоторные команды. Первый вид модуляции — частотно-импульсная, отображающая амплитуду управляющей команды в виде частоты следования импульсов. Этот вид модуляции наблюдается при измерении активности отдельной двигательной единицы. При увеличении амплитуды управляющих сигналов возрастает число активных моторных единиц. По мере возрастания амплитуды входного сигнала активизируются более крупные двига-

тельные единицы, генерирующие импульсы большей амплитуды и создающие большие удельные усилия. Это второй вид модуляции — амплитудно-импульсная.

Если обозначить длительность одного импульса, генерируемого моторной единицей, τ , а период следования импульсов — T , то очевидно, что область пропорциональности между амплитудой управляющего сигнала и частотой следования импульсов $F=1/T$ ограничена величиной $F_{max}=1/\tau$, за которой наступает насыщение, так что коэффициент передачи частотно-импульсной модуляции представляет собой нелинейную характеристику (рис. 5.9). Частота следования импульсов F не только отображает амплитуду управляющего сигнала, она еще ограничивает спектр передаваемого сигнала частотой $F_c \leq F/2$. Это свойство частотно-импульсной

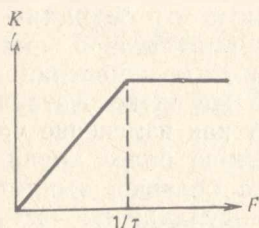


Рис. 5.9. Коэффициент передачи частотно-импульсной модуляции

модуляции следует из теоремы Котельникова об отсчетах. В результате полоса передаваемого сигнала оказывается наибольшей при больших значениях амплитуды управляющего сигнала, которым соответствуют большие частоты следования импульсов, и наименьшей — при малых значениях, т. е. одновременно с амплитудой передаваемого сигнала изменяется и полоса передаваемых частот. Таким образом, для медленных двигательных единиц с частотами 1—10 имп./с максимальная полоса передаваемого сигнала не превышает 0,5—5 Гц, тогда как для

быстрых двигательных единиц, существующих в речевых мышцах, максимальная полоса пропускания теоретически может достигать 50—100 Гц.

Реальная полоса пропускания определяется не только пределом Котельникова, но и способом восстановления (демодуляции) сигнала. Если, например, при амплитудно-импульсной модуляции происходит ступенчатая аппроксимация, т. е. от отсчета к отсчету сохраняется постоянное значение восстанавливаемой функции, равное предыдущему отсчету, то при $T=5$ мс ошибка аппроксимации на частоте $F=100$ Гц достигает 80—90%, а на частоте 10 Гц — 10%. Если используется линейная интерполяция (эквивалентная интегрированию поступающих импульсов), то ошибка на частоте $F=10$ Гц составляет уже только 5% [4.12]. Такое фильтрующее свойство частотно-импульсной модуляции, существующей в нервной системе, следует принимать во внимание при построении детальных моделей системы управления мышечными сокращениями. Когда требуется максимальная скорость изменения напряжения в мышце, для сохранения спектра передаваемого сигнала следует регулировать амплитуду воздействия путем изменения числа активных двигательных единиц, каждая из которых генерирует импульсы с максимально доступной частотой, а не с помощью частотно-импульсной модуляции. Заметим, что зависимость полосы частот передаваемого сигнала от его амплитуды

органически присуща частотно-импульсной модуляции. Она проявляется при любом методе измерения и обработки ЭМГ.

Входной сигнал при частотно-импульсной модуляции восстанавливается путем измерения межимпульсных интервалов — это наиболее точный способ. Метод, при котором подсчитывается число импульсов n на некотором скользящем интервале времени T_n и вычисляется средняя частота $\bar{F} = n/T_n$, весьма помехоустойчив, но при этом дополнительно сужается полоса пропускания до величины, обратно пропорциональной интервалу T_n . Широко применяемый метод интегрирования ЭМГ на интервале T_n при измерении активности отдельной двигательной единицы дает функцию, близкую к оценке, получаемой методом средней частоты:

$$\varphi_i(t) = \frac{1}{T_n} \int_{t-T_n}^t |\psi_i(\theta)| d\theta. \quad (5.8)$$

Здесь $\psi_i(\theta)$ — последовательность импульсов i -й двигательной единицы. Как видно, (5.8) зависит не только от частоты следования импульсов, но и от их формы и амплитуды. При исследовании отдельной двигательной единицы это свойство вносит дополнительные (помимо ограничения полосы) погрешности в оценку входного сигнала. Но это же свойство становится достоинством в случае активизации многих двигательных единиц. Когда активизируются другие двигательные единицы, волокна которых находятся в сфере измерения электродов, то пропорциональность между входным сигналом и измеренной частотой следования импульсов сохраняется в среднем до частоты F_{max} . Термин «в среднем» имеет статистический смысл, причем предполагается, что импульсы различных двигательных единиц не синхронизированы и сдвиг между ними есть случайная величина, распределенная по нормальному закону. В том случае, когда импульсы от различных двигательных единиц достигают электродов в одно и то же время, измерение частоты теряет смысл, а их амплитуда складывается с коэффициентами, зависящими от внутреннего сопротивления электрических генераторов импульсов различных двигательных единиц и сопротивления тканей, находящихся между электродами и волокнами двигательных единиц. В [5.20] было показано, что зависимость интегрированной ЭМГ $\varphi(t)$ при больших интервалах интегрирования

$$\varphi(t) = \frac{1}{T_n} \int_{t-T_n}^t \sum_{i=1}^N |\psi_i(\theta)| d\theta$$

от числа активных двигательных единиц N носит нелинейный характер — амплитуда $\varphi(t)$ пропорциональна $\kappa \sqrt{N}$, где $\psi_i(\theta)$ — временная последовательность импульсов i -й двигательной единицы, а коэффициент пропорциональности $\kappa = \sqrt{2/3\pi} a \sqrt{\tau F}$, где a — амплитуда импульсов; τ — длительность одиночного импульса; F — средняя частота их следования. При этом предполагает-

ся, что амплитуда и форма импульсов каждой двигательной единицы постоянны, время прихода импульсов от различных двигательных единиц несинхронно и вследствие большого числа действующих двигательных единиц справедлива центральная предельная теорема, согласно которой случайная величина $\sum_{i=1}^N |\psi_i(\theta)|$ имеет гауссовское распределение. При тех же предположениях в [5.21] оценивается ожидаемое число пересечений нулевого уровня в интерференционной ЭМГ на некотором интервале. Оказалось, что среднее число пересечений за единицу времени есть

$$F_{\infty} = \frac{\sqrt{6}}{\pi} \frac{2}{\tau}, \quad (5.9)$$

т. е. постоянная величина, не зависящая от формы входного сигнала, а определяющаяся только длительностью импульсов. Полученная оценка несколько больше $F_{max} = 1/\tau$, что связано с двухфазной формой импульсов, аппроксимированных в [5.21] в виде треугольников, симметричных относительно нулевой линии. Оценка частоты импульсов по (5.9) отражает эффект насыщения частотно-импульсной модуляции при большом числе активных двигательных единиц. Если бы удалось автоматически распознать формы импульсов, принадлежащих разным двигательным единицам, и измерить межимпульсные интервалы в последовательностях импульсов, принадлежащих одной и той же двигательной единице, то порог насыщения можно было бы повысить.

При развитии больших усилий активизируются более крупные двигательные единицы, и в этом случае имеют место как частотно-импульсная, так и амплитудно-импульсная модуляции. Простое измерение межимпульсных интервалов приведет к погрешности в оценке формы управляющего сигнала, поэтому интегрирование ЭМГ выглядит более предпочтительным, хотя и не свободно от некоторых погрешностей. Моделирование эффекта активизации более крупных двигательных единиц в [5.125] показывает, что в противоположность частотно-импульсной модуляции между развиваемой силой и частотой следования импульсов в данном случае имеется пропорциональная зависимость.

5.6. АВТОНОМНЫЙ КОНТУР УПРАВЛЕНИЯ

Физиологические исследования позволяют описать схему управления отдельной двигательной единицей (рис. 5.10), несмотря на то, что известны не все характеристики передаточных функций блоков этой системы. Возникающее намерение совершить какое-либо движение приводит к формированию управляющего сигнала, который в высших структурах центральной нервной системы (блок 1) преобразуется в сигналы, управляющие α -мотонейронами (блок 2). Сигналы с выхода α -мотонейрона возбуж-

дают мышечные волокна некоторой двигательной единицы (блок 3), которые, сокращаясь, изменяют напряжение или длину мышцы (блок 4), что приводит к смещению объекта управления (блок 5). В контуре управления существует ряд обратных связей. Так, α -мотонейрон охвачен локальной отрицательной обратной связью через клетки Реншоу (блок 6), которые также получают

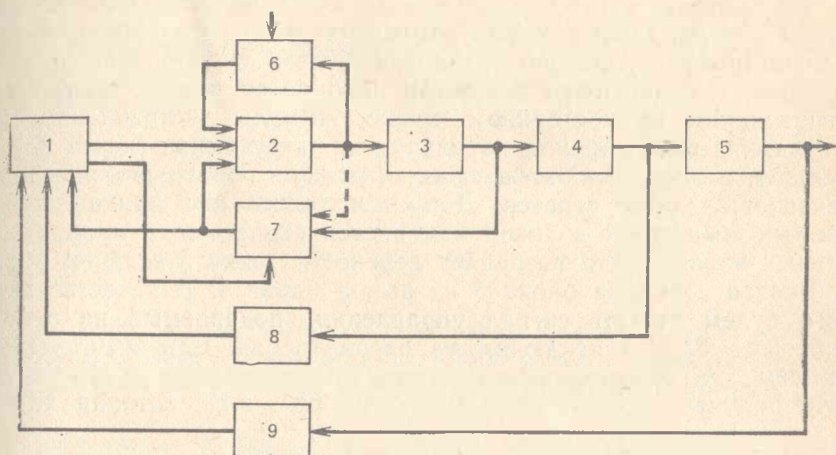


Рис. 5.10. Контур управления двигательной единицей:

1 — центральная нервная система; 2 — α -мотонейрон; 3 — мышечные волокна; 4 — мышечная ткань; 5 — объект управления; 6 — клетки Реншоу; 7 — веретена; 8 — рецепторы; 9 — акустический анализатор

сигналы от других α -мотонейронов и в некоторых случаях могут быть заторможены вплоть до полного прерывания этой обратной связи. Локальная обратная связь через клетки Реншоу улучшает динамические характеристики мотонейрона, делая его амплитудно-частотную характеристику более равномерной [5.53].

Сокращение мышечных волокон возбуждает веретена (блок 7), которые посылают сигналы обратной связи на α -мотонейроны (не обязательно только своей мышцы). Характер отклика веретена на сокращение мышечных волокон зависит от его настройки, задаваемой центральной нервной системой через γ -управление. Изменение напряжения и положения мышцы активизирует остальные рецепторы (блок 8), сигналы обратной связи от которых вместе с сигналами от веретен корректируют управляющие сигналы, что условно показано на рис. 5.10 в виде обратных связей на блок 1. Наконец, смещения объекта управления (например, нижней челюсти) приводят к изменению акустических характеристик речевого сигнала и служат основой для акустической обратной связи (блок 9), замыкающейся где-то в высших структурах центральной нервной системы.

Математические модели системы управления мышечной активностью уже давно служат предметом исследований [5.13, 5.16,

5.33, 5.150]. Специфическим моделям системы управления артикуляцией также посвящен ряд работ [5.28, 5.42, 5.63, 5.143—5.145]. Очевидная сложность системы управления артикуляцией и уровень наших знаний пока не позволяют создать полное количественное описание этой системы, однако исследование ее отдельных элементов вполне возможно.

Рассмотрим контур управления отдельной двигательной единицей с точки зрения теории автоматического регулирования. Для максимального упрощения анализа будем считать контур автономным, т. е. разорвем все связи, идущие от других контуров, и зафиксируем на постоянном уровне сигналы γ -управления. Исключим из рассмотрения локальную обратную связь через клетки Реншоу, а также глобальную акустическую обратную связь и все рецепторы, кроме веретен. Допустим также, что длина возбуждаемых мышечных волокон изменяется синхронно с длиной пассивных волокон, что позволяет перенести точку отведения обратной связи с выхода блока 3 на выход блока 4. В качестве входного будем считать сигнал управления, подаваемый на α -мотонейрон, а блок 1 исключим из рассмотрения. Сам α -мотонейрон опишем как безынерционное пропорциональное звено. В результате получаем схему, показанную на рис. 5.11, которая хотя и

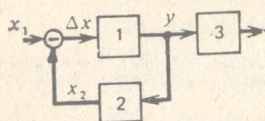


Рис. 5.11. Схема контура управления:

1 — мышца; 2 — объект управления; 3 — рецептор

довольно упрощенно представляет реальные физиологические структуры и связи, но обладает тем достоинством, что для нее можно до конца довести процедуру анализа, используемую в теории автоматического регулирования. Такой контур эквивалентен схеме β -управления или контуру, охваченному обратной связью в схеме с фиксированным γ -управлением. В полученной схеме характеристика рецептора есть некоторая обобщенная характеристика множества веретен. Несмотря на столь простой вид этой системы, ее анализ позволяет оценить точность, устойчивость и качество переходных процессов.

В этом параграфе мы будем пользоваться понятием передаточной функции, основанной на преобразовании Лапласа $X(p) = \int_0^{\infty} x(t)e^{-pt}dt$, где $x(t)$ — некоторая функция времени; $X(p)$ — ее изображение по Лапласу; p — комплексная переменная. Обозначим оператор преобразования по Лапласу как $L[x(t)]$ и перечислим некоторые его свойства, которые понадобятся в дальнейшем.

1. Теорема сложения: преобразование Лапласа суммы функций $x_1(t)$ и $x_2(t)$ есть сумма преобразований Лапласа:

$$L[x_1(t) + x_2(t)] = X_1(p) + X_2(p).$$

2. Теорема умножения: постоянный множитель может быть вынесен за оператор Лапласа:

$$L[\kappa x(t)] = \kappa X(p).$$

3. Теорема запаздывания: задержка сигнала на τ эквивалентна умножению на $e^{-p\tau}$ в лапласовском изображении:

$$L[x(t-\tau)] = e^{-p\tau} X(p).$$

4. Теорема о конечном значении:

$$x(t) = \lim_{t \rightarrow \infty} pX(p).$$

5. Дифференцирование сигнала n раз во временной области при нулевых начальных условиях соответствует в изображении Лапласа умножению на p^n , а при ненулевых начальных условиях

$$L\left[\frac{d^n x}{dt^n}\right] = p^n X(p) - \sum_{k=1}^n x^{(k-1)}(0) p^{n-k}.$$

6. Интегрирование сигнала во временной области эквивалентно делению на p в преобразовании Лапласа. Более подробные сведения о свойствах преобразования Лапласа можно найти в многочисленных руководствах, например в [5.14, 3.7].

Передаточная функция $W(p)$ какой-либо системы управления определяется как отношение преобразования Лапласа выходного сигнала $X_{\text{вых}}(p)$ к преобразованию Лапласа входного сигнала $X_{\text{вх}}(p)$ при нулевых начальных условиях: $W(p) = X_{\text{вых}}(p)/X_{\text{вх}}(p)$. При этих условиях преобразование Лапласа по форме идентично операторному методу, так что, например, для системы второго порядка с постоянными коэффициентами, описываемой обыкновенным дифференциальным уравнением $a_0 \frac{d^2 x_{\text{вых}}}{dt^2} + a_1 \frac{dx_{\text{вых}}}{dt} + a_2 x_{\text{вых}} = \kappa x_{\text{вх}}$, преобразование

Лапласа есть

$$(a_0 p^2 + a_1 p + a_2) X_{\text{вых}}(p) = \kappa X_{\text{вх}}(p),$$

откуда передаточная функция $W(p) = \kappa/(a_0 p^2 + a_1 p + a_2)$. Вообще передаточная функция линейной системы с постоянными коэффициентами без запаздывания может быть представлена в дробно-рациональном виде

$$W(p) = \frac{M(p)}{N(p)} = \frac{a_0 p^n + a_1 p^{n-1} + \dots + a_n}{b_0 p^m + b_1 p^{m-1} + \dots + b_m}.$$

Пользуясь свойствами преобразования Лапласа, можно получить содержательные оценки даже для такой весьма упрощенной системы управления, которая показана на рис. 5.11. Пусть передаточная функция блока 1 — $W_1(p)$, а блока 2 — $W_2(p)$. Запишем соотношения между сигналами в каждой точке такой системы:

$$\Delta X(p) = X_1(p) - X_2(p); \quad Y(p) = \Delta X(p) W_1(p); \quad X_2(p) = Y(p) W_2(p).$$

Исключая промежуточные величины ΔX и X_2 , находим передаточную функцию замкнутой системы

$$W(p) = Y(p)/X_1(p) = W_1(p)/(1 + W_1(p) W_2(p)).$$

Для оценки точности найдем передаточную функцию относительно сигнала рассогласования ΔX

$$V(p) = \frac{\Delta X(p)}{X_1(p)} = \frac{1}{1 + W_1(p) W_2(p)}. \quad (5.10)$$

Представим передаточные функции $W_1(p)$ и $W_2(p)$ в виде дробей:

$$W_1(p) = \kappa_1 \frac{N_1(p)}{M_1(p)}; \quad W_2(p) = \kappa_2 \frac{N_2(p)}{M_2(p)}, \quad (5.11)$$

где κ_1 и κ_2 — коэффициенты усиления в прямой и обратной цепях. Подадим на вход этой системы ступенчатый сигнал с амплитудой A . Его преобразование Лапласа есть $X(p) = A/p$. Теперь можно оценить статическую ошибку замкнутой системы, т. е. разность между значениями входного и выходного сигналов в установившемся режиме. Для этого воспользуемся теоремой о конечном значении, которая с учетом (5.10) принимает вид

$$\Delta x(t) = \lim_{t \rightarrow \infty} p \Delta X(p) = \lim_{p \rightarrow 0} p \frac{A}{p} \frac{M_1(p) M_2(p)}{N_1(p) N_2(p) + \kappa_1 \kappa_2 M_1(p) M_2(p)}.$$

Если $N_1(p)$, $N_2(p)$, $M_1(p)$ и $M_2(p)$ — рациональные или экспоненциальные функции от p , то при $p \rightarrow 0$ каждый из этих членов стремится к 1, поскольку согласно (5.11) свободные члены вынесены в виде множителей κ_1 и κ_2 . Тогда $\lim_{p \rightarrow 0} p \Delta X(p) = A/(1 + \kappa_1 \kappa_2)$, и от-

носительная статическая ошибка

$$\varepsilon = 1/(1 + \kappa), \quad (5.12)$$

где $\kappa = \kappa_1 \kappa_2$ — общий коэффициент усиления прямой и обратной цепей.

Таким образом, при данных условиях замкнутая система имеет ненулевую ошибку установившегося уровня, причем она обратно пропорциональна общему коэффициенту усиления. Такие системы называются *статическими*. Если в $M_1(p)$ или $M_2(p)$ имеется свободный множитель p , что соответствует включению в передаточную функцию $W_1(p)$ или $W_2(p)$ интегрирующего звена, то статическая ошибка

$$\lim_{p \rightarrow 0} p \Delta X(p) = p \frac{A}{p} \frac{p}{p + W_1(p) W_2(p)} = 0.$$

Системы, обеспечивающие нулевую ошибку установившегося уровня, называются *астатическими*. Вводя p^2 в знаменатель $W_1(p)$ или $W_2(p)$, можно свести к нулю и динамическую ошибку, т. е. погрешность отслеживания заданной скорости изменения входного сигнала. Такие системы называются *дважды астатическими*. Уничтожение динамической ошибки, может привести к опасности возникновения неустойчивости, для преодоления которой приходится вводить специальную коррекцию, усложняющую систему управления.

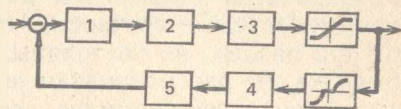
Обращаясь к системе управления артикуляцией, напомним, что, как показано в гл. 3, относительная неустойчивость устано-

вившегося положения различных артикуляторных органов составляет всего 3—5%. Следует учитывать, что в эти оценки входит и погрешность измерений. Кроме того, неизвестный управляющий сигнал также подвержен определенным колебаниям от одной реализации заданного движения к другой. Измерения вариаций положения артикуляторных органов сами по себе не определяют статической ошибки, поскольку неизвестен установившийся уровень входного сигнала, но имеются косвенные доказательства справедливости этой оценки. Известно, что положение кончика языка при артикуляции фрикативных звуков не произвольно, а определяется условиями возникновения турбулентных шумов, причем диапазон возможных положений очень узок. Несомненно, система управления артикуляцией осведомлена об узости этого диапазона и задает управляющие сигналы с надлежащей точностью. Тот факт, что отклонения положения кончика языка для фрикативных звуков от реализации к реализации не превышают 3—4%, свидетельствует о применимости критерия относительного разброса положений артикуляторных органов к оценке статической ошибки. С учетом вышесказанного статическая ошибка системы управления не превышает 3—5%, и даже может быть равна нулю. Если все же допустить существование статической ошибки, то по (5.12) можно найти требуемый коэффициент усиления — для приведенных ошибок он должен быть не менее 30. Допустимость такого коэффициента усиления в системе управления определяется запасом ее устойчивости. Разброс в поддержании скорости артикуляторных органов значительно выше — 20—60%, так что система управления артикуляцией не может обладать астатизмом выше первого порядка.

Вопрос об астатизме системы управления артикуляцией может быть решен путем анализа области ее устойчивости в зависимости от значения коэффициента усиления и запаздывания в цепях управления. Для проведения такого анализа конкретизируем схему контура управления, приводя ее к новому виду (рис. 5.12). В этой

Рис. 5.12. Контур управления:

1 — задержка; 2 — сокращающийся элемент мышцы; 3 — объект управления; 4 — рецептор; 5 — задержка



схеме учтено запаздывание как в прямой, так и в обратной цепях управления, а также три вида нелинейностей: ограничение движению артикуляторного органа, представленное как кусочно-линейная характеристика на выходе объекта управления; порог срабатывания, или зона нечувствительности рецептора, и логарифмический характер отклика рецептора. В соответствии с данными [5.99, 5.13] характеристика рецептора принимается как характеристика множества первичных окончаний веретен, откликающихся на скорость сокращения мышцы или положения артикуляторного органа.

Объект управления вместе с вязко-инерционными мышечными структурами представлен блоком 3 с передаточной функцией $W_2(p) = \kappa_2 / (Tp^2 + 2\xi p + 1)$. Передаточная функция рецептора (блок 4) представлена реальным дифференцирующим звеном $W_3(p) = -[\kappa_3(T_4 p + 1)] / (T_5 p + 1)$. Механизм развития мышечного сокращения описывается блоком 2 с передаточной функцией $W_1(p) = \kappa_1 / [p(T_1 p + 1)]$, которая выбрана соответственно известному свойству суммации сокращений отдельных мышечных волокон. При достаточно большой частоте сокращения и продолжительности активного состояния, меньшей критического интервала утомления мышцы, свойства суммации могут быть приближенно описаны характеристикой либо реального интегратора (т. е. $W_1(p)$), либо апериодического звена: $W_{1a}(p) = \kappa_1 / (T_1 p + 1)$. Выбор той или иной схемы весьма важен для определения астатизма системы управления, поэтому в дальнейшем будут исследоваться оба способа представления этого звена.

Блоки 1 и 5 содержат звенья сосредоточенной задержки, их передаточные функции представлены соответственно как $e^{-p\tau_1}$ и $e^{-p\tau_2}$. Нелинейную статическую характеристику рецептора запишем как

$$x_{\text{вых}} = \begin{cases} \ln \kappa_p x_{\text{вх}}, & x_{\text{вх}} > a_0; \\ 0, & x_{\text{вх}} \leq a_0, \end{cases}$$

где κ_p — эквивалентный коэффициент внутреннего усиления рецептора. При этом необходимо допустить существование специальных средств, регулирующих этот коэффициент усиления при изменении порога a_0 исходя из условия $\ln \kappa_p a_0 = 0$ или $\kappa_p = 1/a_0$. Если в этой области значений сигнала рецептора логарифмическое представление имеет физический смысл, то уменьшение порога возбуждения a_0 должно сопровождаться соответствующим увеличением коэффициента усиления. Эта связь между порогом возбуждения и коэффициентом усиления может быть реализована механизмом γ -управления. В дальнейшем динамика перестройки рецептора не рассматривается — считается, что процессы регулирования протекают при разных, но постоянных во времени значениях порога a_0 . К тому же мы рассматриваем не конкретный отдельный рецептор, а некоторую эквивалентную схему, учитывающую свойства множества рецепторов, среди которых имеются рецепторы с различными значениями порогов и коэффициентов усиления.

Эквивалентную передаточную функцию нелинейности контура учитывающую все виды нелинейности, найдем методом гармонического баланса, основанным на аппроксимации нелинейности первой гармоники разложения ее характеристики в ряд Фурье [5.30]. Входное воздействие принимается синусоидальным с амплитудой A . Тогда выходной сигнал $x_{\text{вых}}(t) = F(x_{\text{вх}}) \approx \approx A[a(A) \sin \omega t + b(A) \cos \omega t]$, где $F(x_{\text{вх}})$ — статическая характеристика нелинейного элемента. Величины $a(A)$ и $b(A)$ определяются как коэффициенты при первой гармонике:

$$a(A) = \frac{\omega}{\pi A} \int_0^{2\pi/\omega} F(A \sin \omega t) \sin \omega t dt;$$

$$b(A) = \frac{\omega}{\pi A} \int_0^{2\pi/\omega} F(A \sin \omega t) \cos \omega t dt.$$

Таким образом, эквивалентная передаточная функция нелинейности $J(A) = a(A) + jb(A)$. Если нелинейный элемент обладает гистерезисом, то коэффициент $b(A)$ зависит от площади S петли гистерезиса $b(A) = S/\pi A^2$, а для однозначных характеристик $b(A) = 0$ и передаточная функция $J(A) = a(A)$. Поскольку в данном случае нелинейность ограничения и логарифматор рецептора не разделены другими звеньями, их можно рассматривать как одну нелинейность и, вводя новую переменную $\psi = \omega t$, получаем:

$$J(A) = \frac{4}{\pi A} \left[\int_{\alpha_0}^{\alpha_1} (\ln \kappa_p A \sin \psi) \sin \psi d\psi + \int_{\alpha_1}^{\pi/4} (\ln \kappa_p a_1) \sin \psi d\psi \right] =$$

$$= \frac{4}{\pi A} \left[\ln \frac{a_1}{a_0} \frac{1 + \sqrt{1 - \left(\frac{a_0}{A}\right)^2}}{1 + \sqrt{1 - \left(\frac{a_1}{A}\right)^2}} + \sqrt{1 - \left(\frac{a_1}{A}\right)^2} - \sqrt{1 - \left(\frac{a_0}{A}\right)^2} \right],$$

где a_1 — диапазон свободного движения артикуляторного органа до ограничения, и $\alpha_0 = \arcsin a_0/A$; $\alpha_1 = \arcsin a_1/A$. На рис. 5.13 показаны статические характеристики эквивалентной передаточной функции рассматриваемой нелинейности $J(A)$ при пороге ограничения $a_1 = 1,5$ см и различных порогах срабатывания a_0 . Как видно, изменение порога a_0 в сравнительно небольшом диапазоне приводит к существенному изменению характеристики отклика нелинейного элемента, что может сыграть важную роль в процессах управления артикуляцией.

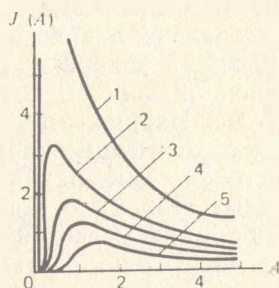


Рис. 5.13. Статические характеристики эквивалентной передаточной функции нелинейности ($a_1 = 1,5$ см):
1 — $a_0 = 0,01$ см; 2 — $a_0 = 0,1$ см;
3 — $a_0 = 0,2$ см; 4 — $a_0 = 0,3$ см; 5 —
 $a_0 = 0,5$ см

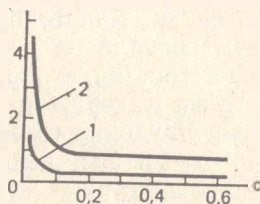


Рис. 5.14. Области устойчивости статической системы ($a_1 = 0,6$ см):
1 — $a_0 = 0,05$ см; 2 — $a_0 = 0,25$ см

Теперь передаточную функцию рассматриваемой системы можно записать как

$$W(p) = \frac{\kappa_1 \kappa_2 e^{-p\tau_1} (T_5 p + 1)}{K e^{-p\tau} J(A) (T_4 p + 1) + p (T_1 p + 1) (T_5 p + 1) (T^2 p + 2\xi p + 1)}, \quad (5.13)$$

где $K = \kappa_1 \kappa_2 \kappa_3$; $\tau = \tau_1 + \tau_2$. Нас интересует влияние общего коэффициента усиления K и общего запаздывания τ на устойчивость системы, поэтому заменим аргумент p на $j\omega$ и решим (5.13) относительно $K(\omega, A)$ и $\tau(\omega, A)$:

$$K = \frac{\sqrt{\Phi_1^2(\omega) + \Phi_2^2(\omega)}}{J(A) [1 + (\omega T_4)^2]}; \quad \tau = \frac{1}{\omega} \operatorname{arctg} \frac{\Phi_1(\omega)}{\Phi_2(\omega)},$$

где $\Phi_1(\omega) = \omega T_4 \{-\omega^4 T T_1 T_5 + \omega^2 [T + T_1 T_5 + 2\xi(T_1 + T_5)] - 1\} - \omega^3 [2\xi T_1 T_5 + T(T_1 + T_5)] + \omega(2\xi + T_1 + T_5)$;

$\Phi_2(\omega) = -\omega^4 T T_1 T_5 + \omega^2 [T + T_1 T_5 + 2\xi(T + T_5)] - 1 + \omega T_4 \{\omega^3 [2\xi T_1 T_5 + T(T_1 + T_5)] - \omega(2\xi + T_1 + T_5)\}$.

Воспользовавшись так называемым методом D -разбиения на плоскости (K, τ) , можно определить область устойчивости системы и предельные значения коэффициента K , при которых система находится на грани устойчивости. Для этого представим знаменатель (5.13) в виде комплексной функции $U(\omega) + jV(\omega)$ и составим якобиан

$$\begin{vmatrix} \frac{\partial U}{\partial K} & \frac{\partial V}{\partial K} \\ \frac{\partial U}{\partial \tau} & \frac{\partial V}{\partial \tau} \end{vmatrix}.$$

Знак этого якобиана указывает, по какую сторону от границы D -области система устойчива. На рис. 5.14 показаны две области устойчивости для статической системы, т. е. для системы, в которой передаточная функция $W_1(p)$ представлена в виде апериодического звена. Как говорилось в § 5.2, время задержки между появлением сигнала на входе α -мотонейрона и началом сокращения мышцы составляет примерно 20—50 мс. Из рис. 5.14 видно, что при таких задержках величины общего коэффициента усиления, соответствующие границе устойчивости системы, находятся в диапазоне 1—10. Эти величины K согласно теореме о конечном значении соответствуют статическим ошибкам не менее 10%, что противоречит наблюдаемым ошибкам. Следовательно, система управления артикуляцией скорее всего обладает астатизмом первого порядка.

Известно, что наличие нелинейности оказывает большое влияние на устойчивость системы управления и в ряде случаев может привести к автоколебаниям. Это явление не всегда вредно — например, звук P в русском языке явно образуется путем автоко-

лебаний кончика языка, тогда как другие переднеязычные звуки артикулируются без автоколебаний. Следовательно, автоколебания в звуке P являются управляемым процессом и возникают за счет надлежащего выбора параметров системы управления.

Амплитуда и частота автоколебаний в рассматриваемом контуре определяются решением уравнения $Y(j\omega)J(A) + 1 = 0$, где $Y(j\omega)$ — фазочастотная характеристика линейной части системы. Графическое решение этого уравнения весьма удобно, оно сводится к поиску пересечения годографом функции $Z(j\omega) = -1/Y(j\omega)$ вещественной оси на интервале, охватываемом значениями эквивалентной передаточной функции $J(A)$ (рис. 5.15). В данном случае $J(A)$ является чисто вещественной функцией (хотя могут появиться и мнимые ее значения, если учесть гистерезис мышцы). На рис. 5.15 показаны годографы $Z(j\omega)$ при различных значениях общего коэффициента усиления K и запаздывания τ , а интервал значений $J(A)$ на вещественной оси соответствует порогам $a_0 = 0,1$ см и $a_1 = 1,5$ см.

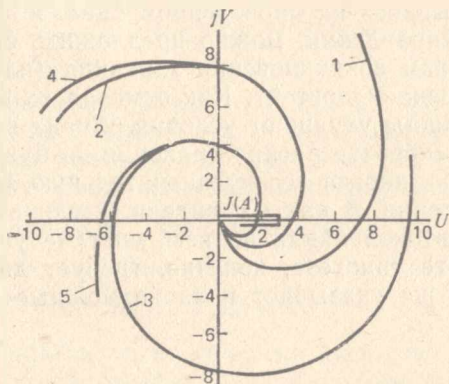


Рис. 5.15. Годографы $Z(j\omega)$ и эквивалентная передаточная функция нелинейности $J(a)$:

1 — $h=2$; $\tau=0,1$; 2 — $h=2$; $\tau=0,2$; 3 — $h=2$; $\tau=0,4$; 4 — $h=4$; $\tau=0,2$; 5 — $h=8$; $\tau=0,1$

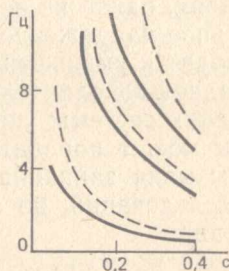


Рис. 5.16. Частота автоколебаний как функция запаздывания:

— — — $T=10^{-4}$, $\xi=-0,004$;
— — — $T=10^{-3}$, $\xi=-0,04$

Анализ устойчивости показывает, что в зависимости от времени запаздывания в системе могут возникнуть автоколебания с частотами порядка 1, 5 и 10 Гц (рис. 5.16). Какие из этих частот будут реально присутствовать в автоколебаниях, зависит в первую очередь от порога рецептора a_0 , который оказывает большое влияние на динамику регулирования. В частности, имеется возможность управления качеством переходных процессов и реализации разных критериев при смене режимов работы системы регулирования.

Автор далек от того, чтобы распространять свойства рассмотренной весьма упрощенной схемы автономного контура на всю систему управления артикуляцией, однако ряд известных явлений в

процессах речеобразования дает основания для проведения весьма важных аналогий. Одно из таких явлений состоит в возникновении заикания при прослушивании собственной речи с задержкой во времени (см. § 5.4). В описанной выше модели автономного контура при увеличении задержки автоколебания также возникают на частотах следования фонем, слогов и слов. Очевидно, что как в реальных условиях, так и в модели действует один и тот же механизм возникновения автоколебаний, в основе которого лежит увеличение задержки в системе управления, хотя, несомненно, в реальных условиях процессы, вызывающие искусственное заикание, значительно сложнее. При изучении естественных заиканий выяснилось, что они сопровождаются возникновением так называемого «блока гортани», в результате чего нарушается координация между сокращениями мышц, сводящих и разводящих голосовые складки, в результате чего в этих мышцах развиваются большие напряжения [5.66, 5.67]. По-видимому, это является следствием нарушения временного контроля сокращений этих мышц, однако непосредственная причина возникновения заиканий при этом все же оставалась неясной. Основываясь на проведенном анализе поведения автономного контура управления, можно предложить следующий гипотетический механизм возникновения заикания (или, точнее говоря, одной из его разновидностей). Как отмечалось в § 5.3, длительное напряжение мышцы ухудшает условия обмена веществ и приводит к уменьшению скорости распространения возбуждения, что создает дополнительную, иногда весьма значительную задержку в цепях системы управления. А как мы видели, увеличение задержки может породить автоколебания во всей системе управления, вызывая заикания. Эта гипотеза, конечно, требует дальнейших исследований, но она же указывает и на возможные их направления.

5.7. ПРОСТРАНСТВЕННО-ВРЕМЕННАЯ ОРГАНИЗАЦИЯ СИСТЕМЫ УПРАВЛЕНИЯ

УПРАВЛЯЕМОСТЬ АКУСТИЧЕСКИХ ПРОЦЕССОВ

Свойства рассмотренного в предыдущем разделе автономного контура определяют лишь локальные, или «дифференциальные», характеристики системы управления артикуляцией, которая в отличие от большинства систем управления движениями человеческого тела, является системой с распределенными параметрами, причем распределенность параметров наблюдается по крайней мере на трех уровнях. Прежде всего это акустический уровень, описываемый волновыми колебательными процессами в речевом тракте. Акустические характеристики речевого сигнала определяются длиной и конфигурацией речевого тракта. Изменения конфигурации в значительной мере зависят от непрерывных

упругих деформаций языка, составляющих второй уровень распределенности. Деформации языка в свою очередь создаются в результате непрерывного или кусочно-непрерывного распределения усилий, создаваемого внутренними и внешними мышцами (третий уровень распределенности). Теория управления системами с распределенными параметрами сложнее теории управления системами с сосредоточенными параметрами и не так подробно разработана, но она доказала свою эффективность во многих технических приложениях. Она основывается на мощном математическом аппарате дифференциальных уравнений в частных производных и интегральных уравнений, обеспечивающем содержательные подходы к решению широкого круга задач. В частности, имеются примеры построения систем управления объектами, описываемыми волновыми уравнениями.

Некоторые из результатов, полученных при разработке технических систем, непосредственно применимы к анализу системы управления речеобразованием. Например, в [5.5] показано, что для системы, описываемой волновым уравнением,

$$\frac{\partial^2 u}{\partial x^2} = \frac{1}{c_0^2} \frac{\partial^2 u}{\partial t^2}, \quad 0 \leq x \leq l, \quad t > 0, \quad (5.14)$$

с граничными условиями $u(0, t) = f(t)$; $u(l, t) = 0$, где $f(t)$ — управляющее воздействие, за исключением специального вида начальных условий, не существует такого $f(t)$, которое перевело бы систему (5.14) из начального неравновесного состояния

$$u(x, 0) = u_0(x); \quad \partial u(x, 0)/\partial t = 0$$

в состояние равновесия в момент $t = T$

$$u(x, T) = 0; \quad \partial u(x, T)/\partial t = 0,$$

где $T \leq 2l/c_0$, т. е. за время, меньшее двукратного пробега звуковой волной отрезка длиной l . Применительно к речевому тракту длиной $l = 17,5$ см это означает, что переходные процессы акустических колебаний не могут быть короче 1 мс при любых характеристиках источника голосового возбуждения. При $T > 2l/c_0$ задача успокоения колебаний имеет не единственное решение.

Для системы управления речеобразованием наиболее важным является вопрос об управляемости резонансных колебаний. Он формулируется следующим образом: каково число управляемых резонансов и какие значения могут принимать их параметры — частоты и амплитуды? Из практики речевых исследований известно, что качество речевого сигнала субъективно не улучшается при расширении полосы частот выше 5 кГц и что достаточная разборчивость сохраняется при ограничении полосы до 3,5 кГц. Это дает основания утверждать, что смысловая информация в речевом сигнале передается (для вокализованных звуков) параметрами первых трех-четырех резонансов. Полученные данные, однако, не дают ответа на вопрос о принципиальной возможности управления параметрами и более высоких резонансов. Кроме того,

известно, что между параметрами первых резонансов в процессах речеобразования существует определенная корреляция, но опять-таки неизвестна ее природа — то ли она является результатом целенаправленных действий системы управления, то ли следствием ограничений, которые система управления не в состоянии преодолеть. Рассмотрим этот вопрос более подробно.

Пусть для функции площади поперечного сечения речевого тракта $S(x)$ и соответствующего уравнения Вебстера

$$\frac{1}{S} \frac{\partial}{\partial x} \left(S \frac{\partial P}{\partial x} \right) - \frac{1}{c_0^2} \frac{\partial^2 P}{\partial t^2} = 0 \quad (5.15)$$

имеется множество собственных чисел $\{\lambda^{(0)}_i\}$ и собственных функций $\{\psi^{(0)}_i(\lambda^{(0)}_i, x)\}$. Обозначим через $S_0(x)$ исходную функцию площади поперечного сечения, подвергнем $S_0(x)$ небольшому изменению и оценим его влияние на собственные числа $\lambda^{(0)}_i$, которые, как известно, определяют частоту резонансных колебаний. Пусть новая площадь поперечного сечения

$$S(x) = S_0(x) + S_1(x), \quad (5.16)$$

где $S_1(x)$ — возмущение, причем $S_1(x) \ll S_0(x)$ при любом x . Тогда можно применить метод Шредингера [4.10] и определить новые собственные числа $\lambda^{(m)}_i$ и собственные функции $\psi^{(m)}_i$ для m -го приближения, представив неизвестные собственные функции возмущенной системы в виде разложения по собственным функциям невозмущенной системы

$$\psi^{(m)}_i(x) = \psi^{(0)}_i(x) + \sum_{m \neq i} \alpha_{mi} \psi^{(0)}_m(x), \quad (5.17)$$

где α_{mi} — малая величина. Возмущенные значения собственных чисел представим как

$$\lambda^2_i = \lambda^{(0)2}_i (1 + \eta_i), \quad (5.18)$$

где η_i — также малая величина.

Разделим переменные в уравнении (5.15) и получим систему $(SP')' + \lambda^2 SP = 0$; $T'' + \lambda^2 c_0^2 T = 0$.

$$(5.19)$$

Подставим (5.16) — (5.18) в (5.19), пренебрежем членами второго порядка малости и примем во внимание, что по определению $(S_0 P')' + \lambda^{(0)2} S_0 P = 0$ и, как следствие, $(S_0 \psi^{(0)}_n)' + \lambda^{(0)2}_n S_0 \psi^{(0)}_n = 0$. Умножим каждый член получившегося уравнения на $\psi^{(0)}_i$ и возьмем определенный интеграл на интервале от 0 до l . В силу ортогональности собственных функций $\psi^{(0)}_h$ с весом S_0 справедливо соотношение

$$\int_0^l S_0 \psi^{(0)}_n \psi^{(0)}_i dx = 0 \quad \text{при } i \neq n,$$

откуда легко можно получить

$$\int_0^l (S_0 \psi^{(0)}_m)' \psi^{(0)}_i dx = 0 \quad \text{при } i \neq m.$$

Тогда для поправки первого порядка имеем

$$\eta_i = -\frac{1}{\lambda_i^{(0)^2}} \int_0^l [S_1(x) \psi_i^{(0)'}(x)]' \psi_i^{(0)}(x) dx + \int_0^l S_1(x) \psi_i^{(0)^2}(x) dx.$$

Возьмем первый интеграл по частям и получим

$$\eta_i = \frac{1}{\lambda_i^{(0)^2}} \int_0^l S_1(x) [\psi_i^{(0)'}(x)]^2 dx - \int_0^l S_1(x) \psi_i^{(0)^2}(x) dx - \\ - S_1(x) \psi_i^{(0)}(x) \psi_i^{(0)'}(x) \Big|_0^l. \quad (5.20)$$

Полагая, что при $x=l$ сигнал свободно излучается (т. е. $\psi_i(l)=0$), а при $x=0$ речевой тракт заканчивается абсолютно жесткой стенкой (т. е. $\psi_i'(0)=0$), избавимся от последнего члена в (5.20). Если принять, что возмущение сконцентрировано в одной точке x_0 , т. е. описывается δ -функцией: $S_1(x) = \epsilon \delta(x-x_0)$, — то поправки на собственные числа возмущений системы найдем как

$$\eta_i = \frac{\epsilon \{ [\psi_i^{(0)'}(x_0)]^2 - [\lambda_i^{(0)} \psi_i^{(0)}(x_0)]^2 \}}{\lambda_i^{(0)^2}}.$$

Отсюда видно, что наибольшее влияние на частоту i -го резонанса оказывает возмущение, приложенное либо в узле i -й собственной функции ($\psi_i^{(0)}(x_0)=0$), либо в ее пучности ($\psi_i^{(0)'}(x_0)=0$). Увеличение или уменьшение частоты i -го резонанса возмущенной системы зависит от знака ϵ . Например, при сужении ($\epsilon < 0$) в первом случае частота резонанса понижается, а во втором — повышается. Частота i -го резонанса останется неизменной, когда

$$[\psi_i^{(0)'}(x_0)]^2 = [\lambda_i^{(0)} \psi_i^{(0)}(x_0)]^2.$$

Тот факт, что в речевом тракте должны существовать области, изменение площади поперечного сечения в которых приводит либо к сильному, либо к слабому изменению акустических характеристик речевого сигнала, был использован в так называемой «квантовой гипотезе» Стивенса [5.152], согласно которой каждый класс звуков любого языка порождается множеством конфигураций речевого тракта, относительно которых акустические характеристики устойчивы, т. е. мало изменяются при изменении конфигурации тракта в пределах заданного множества форм.

Различная чувствительность акустических характеристик речевого сигнала к изменению формы отдельных участков при данной конфигурации речевого тракта накладывает более жесткие требования к точности управления формой одних участков и допускает отклонения формы других участков без заметного влияния на резонансные частоты. Этот фактор необходимо учитывать при анализе формы речевого тракта в различных экспериментах.

Допустим, что мы хотим максимально увеличить частоту i -го

Если координаты x_i точек приложения сосредоточенных возмущений зафиксированы, то система (5.21) линейна относительно амплитуд возмущений и должна удовлетворять всем требованиям, предъявляемым к разрешимости линейных систем: при $N > M$ условие существования решения состоит в равенстве рангов матриц $\|A\|$ и $\|B\|$, где

$$\|A\| = \begin{vmatrix} a_1(x_1) & a_1(x_2) & \dots & a_1(x_M) \\ a_2(x_1) & a_2(x_2) & \dots & a_2(x_M) \\ \dots & \dots & \dots & \dots \\ a_N(x_1) & a_N(x_2) & \dots & a_N(x_M) \end{vmatrix},$$

$$\|B\| = \begin{vmatrix} a_1(x_1) & a_1(x_2) & \dots & a_1(x_M) & \lambda_1^{(0)2} \eta_1 \\ a_2(x_1) & a_2(x_2) & \dots & a_2(x_M) & \lambda_2^{(0)2} \eta_2 \\ \dots & \dots & \dots & \dots & \dots \\ a_N(x_1) & a_N(x_2) & \dots & a_N(x_N) & \lambda_N^{(0)2} \eta_N \end{vmatrix},$$

а при $N=M$ условие разрешимости есть неравенство нулю определителя системы:

$$D = \begin{vmatrix} a_1(x_1) & a_1(x_2) & \dots & a_1(x_M) \\ a_2(x_1) & a_2(x_2) & \dots & a_2(x_M) \\ \dots & \dots & \dots & \dots \\ a_N(x_1) & a_N(x_2) & \dots & a_N(x_M) \end{vmatrix}.$$

На возможные решения наложены дополнительные ограничения, вытекающие из физической реализуемости, — величина ε_m не может превышать некоторых значений как в положительной, так и в отрицательной областях: $-\varepsilon_{m1} \leq \varepsilon_m \leq \varepsilon_{m2}$, где ε_{m2} определяется максимально возможным раскрытием речевого тракта, а ε_{m1} — условием $S_0(x_m) - \varepsilon_{m1} \geq 0$. Тем самым усиливаются ограничения и на способность речевого тракта создавать заданное сочетание значений резонансных частот. Вместе с тем при $M > N$ появляется возможность достичь того же влияния на акустические характеристики с помощью различных комбинаций управляющих воздействий.

Реально локальное сужение в речевом тракте не может быть сконцентрировано в одной точке, оно распределено на некотором интервале, поэтому рассмотрим возможности управления характеристиками речевого тракта в общем виде, не накладывая специальных ограничений на вид возмущения $S_1(x)$. Попробуем определить вид этого возмущения в случае, когда необходимо изменить на заданную величину η_i каждое собственное число $\lambda_i^{(0)}$ в пределах $i \leq N$. Тогда из (5.20) имеем систему интегральных уравнений, где искомая функция $S_1(x)$ находится под знаком интегралов:

$$\left\| \frac{a_i}{\lambda_i^{(0)2}} - b_i - c_i = \eta_i \right\|. \quad (5.22)$$

$$\text{где} \\ a_i = \int_0^l S_1(x) [\psi_i^{(0)'}(x)]^2 dx; \quad (5.23)$$

$$b_i = \int_0^l S_1(x) [\psi_i^{(0)}(x)]^2 dx; \quad (5.24)$$

$$c_i = S_1(l) \psi_i^{(0)}(l) \psi_i^{(0)'}(l) - S_1(0) \psi_i^{(0)'}(0) \psi_i^{(0)}(0). \quad (5.25)$$

Нетрудно видеть, что, когда система (5.22) разрешима, искомая функция $S_1(x)$ может оказаться не единственной, поскольку условия разрешимости касаются лишь коэффициента a_i , b_i и c_i , остающихся неизменными при любых преобразованиях функции $S_1(x)$, сохраняющих неизменными интегралы (5.23), (5.24) и концевые значения (5.25). Таким образом, заданное изменение частотных характеристик речевого тракта может быть достигнуто путем использования его различных конфигураций. Это свойство хорошо известно в акустической теории речеобразования.

Очевидно, что для $N=1$, т. е. при необходимости управлять единственным резонансом, сравнительно легко найти возмущающую функцию $S_1(x)$, удовлетворяющую вычисленным из (5.22) значениям a_i , b_i и c_i . При увеличении N это становится труднее, но даже определив требуемое возмущение $S_1(x)$, необходимо сопоставить его с условиями физической реализуемости. Кинематика речевого тракта в большинстве случаев позволяет произвести не более трех локальных сужений одновременно — на губах, кончике языка и в районе нёбной занавески. Следовательно, наибольшего эффекта в управлении можно ожидать лишь для трех первых резонансов, поскольку для высших резонансов эффективность локальной деформации падает из-за того, что на область деформации приходится и узел, и пучность соответствующей собственной функции.

Однако даже на управляемость первых трех резонансов наложены ограничения, как это видно из анализа условий разрешимости системы (5.22), и дополнительные ограничения в виде условий физической реализуемости, так что могут быть реализованы не все возможные возмущения формы речевого тракта и в результате изменения резонансных частот могут оказаться зависимыми. Еще большая степень зависимости между частотами резонансов должна быть свойственна гласным звукам, поскольку их артикуляция характеризуется не локальной, а распределенной деформацией речевого тракта.

УПРАВЛЯЕМОСТЬ АРТИКУЛЯЦИИ

В гл. 3 показано, что конфигурация губ и языка описывается бесконечным рядом $u(\varphi, t) = \sum_{k=1}^{\infty} \psi_k(\varphi) T_k(t)$, где u — отклонение от равновесного положения; $T_k(t)$ — k -я временная

мода; $\psi_k(\varphi)$ — k -я пространственная мода (собственная функция). Для языка, например, эти моды являются решениями системы:

$$(E J_z \psi_k'')'' + \psi_k'' - (\lambda_k^2 - c) \psi_k = 0,$$

$$T_k'' + 2g T_k' + \omega_k^2 T_k = - \frac{1}{\|\psi_k\|^2} \int_0^{\varphi_{\max}} \frac{\partial^2 M(\varphi, t)}{\partial \varphi^2} \psi_k(\varphi) d\varphi,$$

где E — модуль упругости; J_z — момент инерции; λ_k — собственные числа; $2g = r/\rho$; r — погонная вязкость; ρ — погонная плотность; R_0 — средний радиус языка в нейтральном положении; $M(\varphi, t)$ — распределенный изгибающий момент, создаваемый внешними и внутренними мышцами; $\|\psi_k\|^2$ — норма собственных функций; $\varphi_{\max}$ — максимальное значение угла в полярной системе координат (соответствующее кончику языка); c — погонная упругость подстилающих тканей и $\omega_k^2 = (cR_0^2 + \lambda_k^2)/\rho R_0^2$.

В § 3.4 было показано, что при определенных предположениях о распределении усилий внутренних мышц по поперечному сечению языка произвести замену $-\partial^2 M/\partial \varphi^2 = A(\varphi)F(t)$, где $A(\varphi)$ — функция, учитывающая момент инерции сечения, положение и протяженность сокращающихся мышечных волокон; $F(t)$ — закон изменения во времени усилия, создаваемого внутренними мышцами.

Для упрощения анализа представим усилия, создаваемые внешними мышцами в виде суперпозиции δ -функций. Тогда уравнение вынужденных движений k -й временной моды имеет вид

$$T_k'' + 2g T_k' + \omega_k^2 T_k = \int_0^{\varphi_{\max}} \left[A(\varphi) F(t) + \sum_{j=1}^J a_j \delta(\varphi - \varphi_j) \right] \psi_k(\varphi) d\varphi,$$

где коэффициент a_j учитывает амплитуду и угол между нормалью к поверхности и направлением усилия, создаваемого внешней мышцей. Собственные функции ψ_k будем считать нормированными к единице.

Длина волокон внутренних мышц меньше общей длины языка, поэтому они могут сокращаться на разных его участках. Примером этого служит сокращение волокон внутренних мышц на передней части языка, которое приводит к подъему или опусканию его кончика. Однако создаваемые внутренними мышцами усилия всегда распределены на более или менее протяженном участке и поэтому возбуждают одновременно почти все временные моды, что соответствует изменению формы языка на всем его протяжении. Усилия внешних мышц, напротив, могут быть сосредоточены на более или менее узких участках, поэтому анализ возможностей управления формой языка будет более показательным, если принять во внимание лишь внешние мышцы, т. е. положить $F(t) = 0$. Тогда имеем очевидное преобразование $T_k'' + 2g T_k' + \omega_k^2 T_k = \sum_{j=1}^J a_j \psi_k(\varphi_j)$, откуда следует, что амплитуда k -й временной моды, возбуждаемой сосредоточенными условиями,

пропорциональна значению k -й собственной функции в точке приложения этого усилия. В частности, если φ_j совпадает с узлом (нулем) k -й собственной функции, то эта временная мода будет отсутствовать в движении языка, а k -я пространственная мода не изменит своего вклада в его форму. Такие точки называются точками неуправляемости [5.5], причем даже расширение воздействия на окрестность этой точки не делает ее точкой управляемости, за исключением особых случаев распределения усилий на окрестности. Если же φ_j совпадает с пучностью k -й собственной функции, то вклад этой моды в форму языка и характер его движений будет максимальным. Перечисленные свойства дают основание предположить в отдельных случаях, что внешние мышцы играют большую роль в управлении формой языка, чем внутренние.

$$\left. \begin{aligned} A_1 &= a_1 \psi_1(\varphi_1) + a_2 \psi_1(\varphi_2) + \dots + a_J \psi_1(\varphi_J); \\ A_2 &= a_1 \psi_2(\varphi_1) + a_2 \psi_2(\varphi_2) + \dots + a_J \psi_2(\varphi_J); \\ &\vdots \\ A_N &= a_1 \psi_N(\varphi_1) + a_2 \psi_N(\varphi_2) + \dots + a_J \psi_N(\varphi_J), \end{aligned} \right\} \quad (5.26)$$

Если точки приложения усилий φ_j фиксированы, то система (5.26) линейна относительно неизвестных величин усилий a_j и разрешима при хорошо известных условиях (см. предыдущий параграф). Это означает, что не все требуемые положения вектора $(A_1, A_2, \dots, A_N)$ могут быть реализованы с помощью управления величинами усилий a_j . Ситуация принципиально не изменится и в случае непрерывного распределения усилий вдоль языка, которое создает, например, челюстно-язычная мышца. Изменение граничных условий на корне языка и изменение жесткости его тканей расширяют возможности управления за счет смещения узлов и пучностей пространственных мод, однако основное свойство такой распределенной системы, как язык, остается неизмен-

ным — реализуемы не все желаемые формы языка. Это ограничение усиливается условиями физической реализуемости управляющих воздействий a_j . То же самое относится и к реализуемости заданного закона изменения формы языка во времени — система управления не в состоянии создать произвольные траектории в пространстве параметров $(A_1, A_2, \dots, A_N)$ от одного состояния к другому.

Задача контроля истинного положения управляемого объекта с целью создания обратной связи в системах с распределенными параметрами решается сложнее, чем в сосредоточенных системах. Допустим, что мы в состоянии измерить отклонения языка от нейтрального положения в любой точке f_j . Тогда, измеряя положения точек в каждый момент t имеем линейную систему:

$$\left. \begin{aligned} u(\varphi_1, t) &= T_1(t)\psi_1(\varphi_1) + T_2(t)\psi_2(\varphi_1) + \dots + T_N(t)\psi_N(\varphi_1); \\ u(\varphi_2, t) &= T_1(t)\psi_1(\varphi_2) + T_2(t)\psi_2(\varphi_2) + \dots + T_N(t)\psi_N(\varphi_2); \\ &\vdots \\ u(\varphi_N, t) &= T_1(t)\psi_1(\varphi_N) + T_2(t)\psi_2(\varphi_N) + \dots + T_N(t)\psi_N(\varphi_N), \end{aligned} \right\} \quad (5.27)$$

Отметим, что для решения системы (5.27) и сокращения числа вычислительных операций путем специального размещения точек измерения нужно знать вид собственных функций ψ_k , распределение их узлов и собственных чисел. Иными словами, при модальном управлении формой языка система управления артикуляцией должна быть каким-либо образом осведомлена о текущих пространственно-временных характеристиках языка как упругого тела, т. е. располагать математической моделью языка.

шения между модами языка по данным измерения его положения в отдельных точках. Источником такой информации служат механорецепторы и веретена, расположенные в тканях языка. Оценим способность веретен измерять локальные деформации языка. Рассмотрим прямоугольный элемент языка длиной l и толщиной d в системе координат xoy , где ось x направлена вдоль языка, а y — перпендикулярно к его поверхности (рис. 5.17, где штрихпунктирной линией показана средняя линия элемента). Совместим левые грани исходного и деформированного элементов, тогда правая грань деформированного элемента повернется относительно своей левой грани на угол α . Аппроксимируем форму деформированного элемента трапецией с параллельными основаниями, причем допустим, что после деформации длина граней осталась прежней, нижнее основание удлинилось, а верхнее — укоротилось на Δl .

Пусть изменение длины основания Δl регистрируется веретеном, расположенным на нижнем основании. Тогда из геометрических построений найдем, что

$$\sin \frac{\alpha}{2} \approx \frac{\Delta l}{d},$$

и изменение координаты x_A точки A , переместившейся в точку B , есть $\Delta x \approx (l - \Delta l) \cos \alpha/2 - l \approx -\Delta l$, так как для малого угла $\alpha \cos \alpha/2 \approx 1$, и $\Delta y \approx (l - \Delta l) \sin \alpha/2$, или, пренебрегая величиной второго порядка малости, имеем $\Delta y \approx \Delta l(l/d)$. Как нетрудно ви-

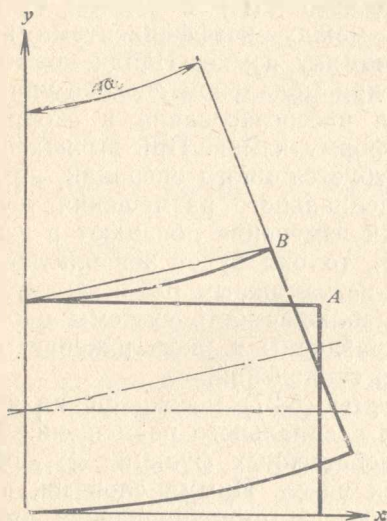


Рис. 5.17. Деформация элемента языка

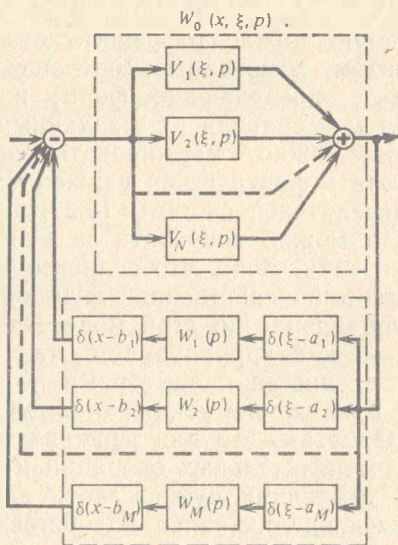


Рис. 5.18. Схема модального управления языком:

$V_j(\xi, p)$ — моды языка; $W_j(p)$ — передаточные функции рецепторов

десь, отношение l/d характеризует чувствительность измерения приращения вертикальной координаты точки Δy с помощью оценки изменения длины веретена Δl , расположенного горизонтально. Эта чувствительность довольно мала, поскольку l — это исходная длина внутренних волокон веретена и $l \ll d$. Здесь источником ошибок является оценка меняющейся в процессе артикуляции толщины языка d , но это не единственный вид ошибок. Измерение приращений координат Δx и Δy верхней точки B правой грани деформированного элемента производится относительно ее старого положения A , т. е. относительно положений левой грани этого элемента, которое считается известным, но которое также смещается в процессах деформации языка. Следовательно, оценка новых положений координат верхней поверхности языка должна производиться путем интегрирования дифференциальных приращений Δx и Δy вдоль по длине языка, что приводит к накоплению ошибок. Точную оценку погрешности измерения координат деформированной поверхности дать затруднительно, однако ясно, что величина этой ошибки не пренебрежимо мала.

Накапливающиеся ошибки измерения формы языка с помощью веретен могут быть скорректированы тактильными рецепторами, например, в момент касания верхними и боковыми частями поверхности языка нёба и верхних зубов. Поэтому можно согласиться с мнением, высказанным в [5.49], о важной роли тактильных рецепторов на завершающих стадиях артикуляционных движений не только для взрывных согласных, обладающих стадией смычки, но и для фрикативных. Об этом свидетельствуют и уже упоминавшиеся эксперименты с моделью управления артикуляцией фрикативных звуков, в которую были включены и тактильные рецепторы.

К анализу систем с распределенными параметрами, так же как и к анализу систем с сосредоточенными параметрами, применим аппарат передаточных функций. Но передаточные функции этих систем обладают несколько иными свойствами, чем передаточные функции сосредоточенных систем. Здесь первоначальным является понятие импульсной переходной функции — отклика системы на единственный короткий импульс в момент t_0 , приложенный в некоторой точке x_0 пространственной области D , т. е. $\delta(x-x_0)\delta(t-t_0)$. Импульсная переходная функция есть не что иное, как функция Грина, характеризующая решение соответствующего уравнения в частных производных, и обозначается она как $G(x, \xi, t_1, t_2)$, где x, t_1 — аргументы входного воздействия; ξ, t_2 — аргументы выходного отклика. Если система стационарна, т. е. ее параметры не изменяются во времени, то импульсная переходная функция зависит лишь от разности аргументов $\tau = t_1 - t_2$, и тогда передаточная функция такой системы $W(x, \xi, p)$ есть преобразование Лапласа от $G(x, \xi, \tau)$ по аргументу τ :

$$W(x, \xi, p) = \int_0^{\infty} G(x, \xi, \tau) e^{-p\tau} d\tau.$$

Если преобразование Лапласа от входного сигнала есть $w(\xi, p)$, то преобразование Лапласа от выходного сигнала

$$U(x, p) = \int_D W(x, \xi, p) w(\xi, p) d\xi,$$

а во временной области $u(x, t)$ вычисляется как обратное преобразование Лапласа:

$$u(x, t) = \frac{1}{2\pi} \int_{c-j\infty}^{c+j\infty} U(x, p) e^{pt} dp,$$

где c меньше всех действительных частей полюсов функции $U(x, p)$.

С помощью аппарата передаточных функций можно анализировать системы, состоящие из различных соединений блоков: последовательного, параллельного, замыкания обратной связью и т. д., однако во многих случаях решения оказываются более сложными, чем для систем с сосредоточенными параметрами. Если рассмотреть общий случай обратной связи в системе управления движениями языка, то множество рецепторов лучше всего представить в виде распределенного измерителя с передаточной функцией $W_p(x, \xi, p)$. Тогда передаточная функция замкнутой системы $W(x, \xi, p)$ по [5.5] записывается как интегральное линейное уравнений Фредгольма второго рода:

$$W(x, \xi, p) = \int_D W_1(x, \eta, p) W(\eta, \xi, p) d\eta + W_0(x, \xi, p),$$

где $W_1(x, \eta, p)$ — ядро этого интегрального уравнения, определяемое как передаточная функция последовательного соединения блоков обратной и прямой связи; D — область пространственного определения выходного сигнала; $W_0(x, \xi, p)$ — передаточная функция объекта управления. Теория интегральных уравнений Фредгольма достаточно хорошо разработана, и, более того, можно найти схемы, в которых это уравнение вырождается и решение может быть найдено в явном и даже алгебраическом виде. Во всяком случае здесь применим метод последовательных приближений.

Если принять, что форма языка регулируется с помощью модального управления и как измерения, так и управляющие воздействия производятся в отдельных точках с координатами соответственно a_i и b_i , то схема управления упрощается (рис. 5.18). Передаточная функция такой схемы [5.5]

$$W(x, \xi, p) = \sum_{i=1}^N W_i(p) W_0(x_i, b_i, p) X_i(\xi, p) + W_0(x, \xi, p),$$

где $W_0(x, \xi, p)$ — передаточная функция языка, которая легко находится как сумма передаточных функций первых N мод; W_i — передаточные функции цепей обратной связи, а $X_i(\xi, p)$ находятся из системы линейных уравнений:

$$\|\delta_{ij} - W_0(a_i, b_j, p)\| \cdot \|X_i(\xi, p)\| = \|W_0(a_i, \xi, p)\|;$$

$$\delta_{ij} = \delta(x - a_i) \delta(x - b_j).$$

В терминах обратной связи можно также сравнительно легко решить вопрос о влиянии ограничения движения языка при соприкосновении его с нёбом или верхними зубами, тогда как решение этой задачи в терминах граничных условий для упругой пластины приводит к возникновению нелинейных уравнений.

Располагая передаточной функцией распределенной системы управления, можно заранее оценить запас ее устойчивости, достижимость заданного состояния, качество переходных процессов и другие характеристики еще до фактической реализации намечаемых движений. В частности, к системам с распределенными параметрами применим критерий устойчивости Найквиста, состоящий в требовании, чтобы число оборотов вектора $W(j\omega)$, $-\infty \leq \omega \leq \infty$, относительно точки $(-1, 0)$ на комплексной плоскости было равно нулю. Здесь $W(j\omega)$ — функция, получаемая из передаточной функции замкнутой системы $W(p)$ путем замены p на $j\omega$ [5.30]. В [5.5] показано, что если известны собственные функции и собственные числа объекта управления, то можно определить степень управляемости системы с заданными параметрами, и если она управляема, синтезировать управляющее воздействие, оптимальное по некоторому критерию. В отношении системы управления артикуляцией это означает, что в случае существования в центральной нервной системе внутренней модели, получающей информацию о сравнительно медленно меняющихся параметрах объекта управления с помощью механорецепторов и веретен, появляется возможность формирования управляющих сигналов, не требующих быстросредействующей обратной связи; при этом изменяются в случае необходимости параметры самой системы управления для достижения желаемого результата.

КРИТЕРИИ УПРАВЛЕНИЯ АРТИКУЛЯЦИЕЙ

В последние годы основное внимание уделяется поиску критериев управления процессами речеобразования. Ясно, что успех этого поиска невозможен без учета информационных характеристик речевого сигнала и свойств восприятия, которые изучены далеко не достаточно. В настоящее время экспериментальные данные о физиологии движений и специфике процессов речеобразования, а также современные методы анализа и синтеза систем автоматического управления позволяют установить некоторые из критериев управления, действующих на различных уровнях и этапах речеобразования. Прежде всего остановимся на ограничениях, накладываемых на управляющие команды безотносительно к содержанию речевого сигнала.

Из анализа ошибок речеобразования следует, что существует такой уровень, на котором элементы речевого потока представлены во вневременной форме и лишь на следующем уровне происходят их отбор и упорядочение во времени [5.1, 5.68, 5.118]. В качестве таких элементов могут выступать слова, слоги, отдель-

ные звуки и артикуляционные признаки (место и способ артикуляции, звонкость — глухость и т. д.). Очевидным свойством этого уровня является ограничение на число элементов разного рода, помещающихся в оперативной памяти системы управления артикуляцией. Различные психофизические эксперименты позволяют определить примерное число запоминаемых элементов — 5—9 [5.22, 3.9, 5.68, 5.120]. Ограничения кратковременной памяти распространяются и на уровень нейромоторных команд, как это следует из систематического изменения длительности элементов речевого потока [5.96, 5.97, 5.101, 5.104, 5.105]. Наиболее подробно это явление исследовано для английского и шведского языков, но имеются и работы, посвященные русскому языку [5.3]. Оно заключается в том, что длительность элементов речевого потока зависит от их количества. Общая закономерность такова: при увеличении числа элементов их длительность сокращается. Степень сокращения зависит от окружения и положения данного элемента в слове и фразе [5.129, 5.130]. Установлено, что соблюдение этой закономерности существенно улучшает натуральность звучания и разборчивость синтетической речи [5.51, 5.98].

Существуют другие закономерности: например, длительности согласного и гласного звуков в слове находятся в обратной зависимости — увеличение одной вызывает уменьшение другой. В частности, глухие согласные звуки (взрывные и фрикативные) длиннее, чем соответствующие им звонкие, и гласные перед глухими согласными короче, чем перед звонкими. Конечные сегменты в конце слов и фраз длиннее начальных. Эти явления уже не связаны с емкостью кратковременной памяти, и здесь действуют другие критерии, возможно, связанные с энергетикой источников нейромоторных импульсов и мышечных сокращений.

Фактор ограниченности энергетических ресурсов биологических структур иногда выступает в литературе как «принцип экономии усилий» и, несомненно, играет большую роль в процессах речеобразования. Расход энергии при работе мышцы считается в [5.81] одним из наиболее важных критериев качества управления, причем указывается на необходимость учета в балансе энергий и отрицательной работы, т. е. работы, совершаемой над мышцей внешними источниками. В теории автоматического управления ограничения на энергетику источника управления обычно записывают в виде нормы управляющих сигналов

$$\|u(t)\| = \left(\int_0^T |u(t)|^p dt \right)^{1/p}, \quad (5.28)$$

где $u(t)$ — управляющий сигнал; T — время управления. Различные значения параметра p соответствуют различным ограничениям, наложенным на источник управления: при $p=1$ ограничивается «расход» (например, количество жидкости газа или электричества), при $p=2$ — энергия сигнала, при $p=\infty$ амплитуда управляющего сигнала $\max|u| \leq a$. Из теории функций действи-

тельного переменного известно, что различным показателям p в (5.28) соответствуют различные метрики в пространстве сигналов управления, причем чем больше p , тем «сильнее» эта метрика, т. е. расстояние ρ_i , измеренное в метрике с показателем p_i , всегда больше расстояния ρ_j , измеренного в метрике $p_j < p_i$. Метрики, задаваемые функционалом (5.28), называются L -метриками, причем метрика L_2 при $p=2$ есть обычная евклидова метрика. Предельный случай $p=\infty$ совпадает с одной из так называемых C -метрик, а именно с метрикой C_0 , в которой расстояние измеряется как максимум модуля разности двух функций: $\rho_{C_0} = \max |u_1(t) - u_2(t)|$. Это самая слабая из C -метрик, но она мажорирует все L -метрики. Другие разновидности C -метрик накладывают ограничения на максимумы модулей производных. Так, метрика C_1 , ограничивающая первую производную, может быть интерпретирована как критерий «минимум управления», избегающий резких изменений значения управляющих сигналов.

Такая упорядоченность метрик гарантирует удовлетворение всех критериев, мажорируемых выбранной метрикой. В то же время использование более сильного критерия там, где это не вызывается необходимостью, накладывает излишние ограничения на управляющие сигналы и сужает возможности системы управления. Например, при совершении медленных движений с малой амплитудой критерий ограничения максимума модуля управляющего сигнала слишком силен. Поэтому можно ожидать, что в зависимости от характера совершаемых движений на передний план выступают различные ограничения на источник управления.

Согласование изменений длительности гласного и согласного в слове можно интерпретировать как стремление стабилизировать потребную энергию на интервале слова, а удлинение конечного сегмента — как дефицит энергии, истраченной на начальных этапах. В этом отношении существует заметная разница в образовании гласных и согласных: характеризуясь распределенной деформацией речевого тракта, артикуляция гласных требует меньше энергии, чем артикуляция согласных. Это проявляется и при изменении темпа артикуляции: например, при ускорении темпа длительность согласных обычно уменьшается незначительно по сравнению с длительностью гласных (например, в отношении 0,9 н 0,75 по [5.71]) или при существенном укорочении гласных, (примерно вдвое) скорость переходных процессов изменяется в пределах 10%. По-видимому, управление артикуляцией гласных обладает большим запасом по мощности, чем управление артикуляцией согласных. Изменение темпа артикуляции вызывает и другой эффект, связанный с ограничениями энергетики источника, — реорганизацию команд управления. Одно время считалось, что изменение темпа артикуляции приводит к пропорциональному изменению длительности команд, тогда как относительные сдвиги фаз управления различными артикуляторными органами сохраняются инвариантными. Однако вскоре было обнаружено измене-

ние фаз активизации мышц, управляющих небной занавеской, языком и губами при изменении темпа артикуляции и места ударения [5.69, 5.70, 5.79, 5.80, 3.32].

В [1.10] исследовалось влияние скорости артикуляции и места слога в звукосоочетании на сдвиг начала нейромоторных команд относительно начала и конца смычки, оцениваемой по акустическому сигналу. Электрическая активность верхней и нижней кольцевых мышц губ (ВКГ) и (НКГ), подбородочной мышцы (ПМ) и мышцы, опускающей углы губ (ОУГ), регистрировалась проволочными и поверхностными электродами.

Речевой материал состоял из отдельных звуков, 120 бессмысленных звукосоочетаний, 24 слов и 11 фраз. Звукосоочетания оканчивались слогом *БА* или *ПА*, начинаясь соответственно с *Б* или *П*, а в середине находились конструкции гласный — согласный — гласный или гласный — согласный — согласный — гласный с различными сочетаниями звуков (например, *БУБИБА*, *ПАПМАПА*, *БАБДОБА* и т. д.). Весь текст повторялся несколько раз в нормальном и примерно вдвое ускоренном темпах.

В тех случаях, когда удавалось достаточно надежно установить начало и конец моторной команды, измерялись ее длительность и время появления относительно начала и конца смычки согласного звука, определяемых по записи акустического сигнала.

Точность измерений составляла 5 мс, и результаты их усреднения округлялись также до 5 мс.

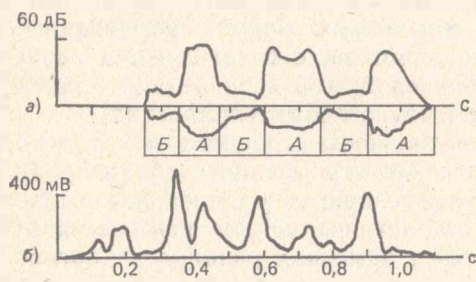


Рис. 5.19. Электрическая активность нижней кольцевой мышцы губы, регистрируемая проволочными электродами:

а — огибающая речевого сигнала;
б — потенциалы на проволочных электродах

Сигналы, регистрируемые проволочными электродами от НКГ для губных согласных, имеют три выраженные фазы, соответствующие смыканию губ, смычке и размыканию губ (рис. 5.19). При этом амплитуда сигнала на фазе размыкания в 2—4 раза больше амплитуды сигнала на фазе смыкания, а длительность сигнала размыкания примерно во столько же раз короче длительности сигнала на фазе смыкания, и это соотношение сохраняется и при ускорении темпа артикуляции, т. е. здесь явно происходит обмен длительности сигнала на его амплитуду, при котором уровень сигнала на фазе смычки для звонкого согласного *Б* заметно меньше амплитуды импульса смыкания, тогда как для глухого *П*, а также для *В*, *Ф*, *М* уровень сигнала во время смычки не понижается. При быстрой артикуляции у звука *Б* импульсы смыкания и размыкания сближаются, так что спад уровня сигнала на фазе смычки не столь велик, как при нормальном темпе. У остальных губных активность на

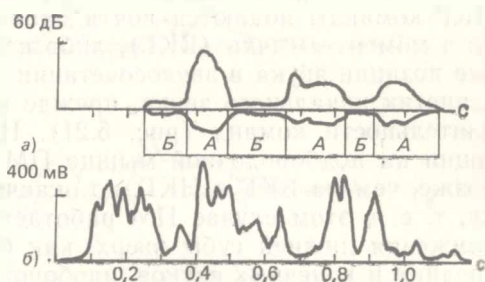
фазе смычки (или щели) при этом повышается, в результате чего разница в форме сигналов уменьшается. Интересно отметить обнаружение активности НКГ на фазе размыкания переднеязычных согласных *Д, Н, З, Л*, причем в быстром темпе артикуляции амплитуда импульса размыкания для этих звуков сравнима с амплитудой соответствующих импульсов для губных согласных.

Форма ЭМГ сигналов в звукосочетаниях зависит от места звука в последовательности: для начальных звуков она ближе к форме сигналов для отдельно произносимых звуков в том же темпе, тогда как звуки в середине и конце звукосочетания для нормального темпа имеют сигналы, сходные с сигналами быстро артикулируемых звуков. В то же время у начальных слогов *БА* наблюдается иногда обмен амплитуды импульса на фазе смыкания на его длительность, так что интеграл по времени остается примерно постоянным.

При регистрации сигналов поверхностными электродами амплитудные соотношения сигналов на разных фазах артикуляции губных согласных оказываются несколько иными, чем при регистрации проволочными электродами: амплитуда импульса размыкания обычно меньше амплитуды импульса смыкания или уровня сигнала на фазе смычки (или щели) (рис. 5.20) (за исключением звука *М*, где наблюдается весьма короткий импульс размыкания значительной амплитуды). В ряде случаев уровень сигнала на интервале смычки (для *П*) или щели (для *В, Ф*) существенно превышает амплитуду сигнала на остальных фазах. При увеличении темпа артикуляции уровень сигнала на фазе смычки (или щели) повышается. В противоположность сигналам от проволочных электродов активность НКГ для переднеязычных *Д, Н, З, Л* мало отличается от фона даже при быстром темпе артикуляции. Форма сигнала в звукосочетаниях в общем со-

Рис. 5.20. Электрическая активность нижней кольцевой мышцы губы, регистрируемая поверхностными электродами:

а — огибающая речевого сигнала;
б — потенциалы на поверхностных электродах



ответствует форме сигналов для отдельных звуков, однако для начальных звуков в звукосочетании длительность сигналов сокращается.

Характерной особенностью сигналов, регистрируемых как проволочными, так и поверхностными электродами, является широкая изменчивость их формы и временных соотношений, что побуждает исследователей сначала усреднять сигналы, а затем измерять их характеристики. Однако подобное усреднение в значительной степени маскирует структуру сигнала именно в силу его изменчивости, поэтому представляется целесообразным усреднять

Таблица 5.1

Мышца	Опережение начала и длительность моторных команд в разных темпах артикуляции, мс, для положения											
	начального				среднего				конечного			
	Норма		Быстро		Норма		Быстро		Норма		Быстро	
	опережение	длительность	опережение	длительность	опережение	длительность	опережение	длительность	опережение	длительность	опережение	длительность
Верхняя кольцевая	-210	210	-175	150	-75	95	-60	70	-85	115	-65	100
Нижняя кольцевая	-200	175	-175	140	-75	95	-55	60	-80	105	-65	90
Подбородочная	-170	125	-105	60	-50	60	-50	35	-55	60	-55	45
Опускающая угол губы	-125	90	-115	70	-60	50	-60	35	-40	50	-50	40

не сами сигналы, а их параметры. Результаты такого усреднения для поверхностных электродов примерно по 40—120 реализациям звука *Б* в бессмысленных звукосочетаниях приведены в табл. 5.1, где показаны время начала команд относительно начала звонкой смычки (опережение) и длительность этих команд для начального, среднего и конечного положений звука. Из табл. 5.1 видно, что на верхнюю и нижнюю ветви кольцевой мышцы губ ВКГ и НКГ команды подаются почти одновременно и оканчиваются либо в момент смычки (ВКГ), либо незадолго до нее (НКГ). Влияние позиции звука в звукосочетании проявляется в отличии характеристик начального звука, прежде всего во времени появления и длительности команд (рис. 5.21). Например, для начальной позиции на подбородочной мышце ПМ команда подается несколько позже, чем на ВКГ и НКГ, но оканчивается она задолго до смычки, т. е. в этом случае ПМ работает только на начальной стадии движения нижней губы вверх, как бы играя роль ускорителя¹. У средних и конечных звуков, наоборот, команды на смыкание оканчиваются несколько позже начала смычки. При этом для всех позиций звука разница между началами команд на ВКГ, НКГ и ПМ в абсолютном выражении остается почти неизменной и составляет примерно 20—40 с. Активность ОУГ также изменяется в зависимости от положения звука в звукосочетании.

Ускорение темпа артикуляции для среднего и конечного положений звука почти не влияет на начало команд у мышц ПМ и ОУГ, но зато укорачивает их длительность, хотя и не пропорци-

¹ Соотношение между началом активности ПМ и НКГ здесь получилось обратным тому, которое отмечается в [3.49].

онально темпу ускорения речи. Для начального положения звука в быстром темпе команды в 2—3 раза опережают команды для средних и конечных положений, но тем не менее все же сдвигаются по направлению к смычке по сравнению с командами в нормальном темпе. Таким образом, временные соотношения между началами команд различных мышц и их длительностями изменя-

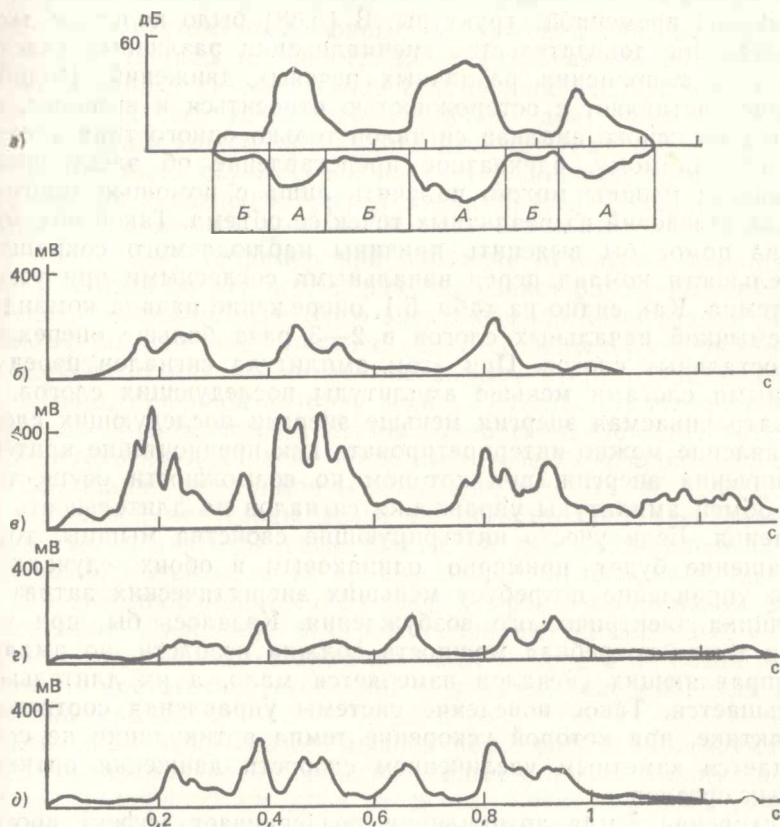


Рис. 5.21. Координация губных мышц:

а — огибающая речевого сигнала; *б* — кольцевая верхняя мышца; *в* — кольцевая нижняя мышца; *г* — мышца, опускающая углы губ; *д* — подбородочная мышца

ются в зависимости от темпа артикуляции и положения звука в звукосочетании. Анализ сигналов, регистрируемых проволочными электродами, подтверждает этот вывод. В данном случае импульсы на смыкание появляются значительно раньше у начальных звуков, но длительность их короче, чем у сигналов, регистрируемых поверхностными электродами. Время же появления и длительность импульса перед взрывом смычки почти не зависят от положения звука в слове.

Сравнивая сигналы, регистрируемые поверхностными и про-

лочными электродами, необходимо отметить устойчивое различие их формы. В зависимости от вида электродов, как было показано, изменяются и интерференционные картины (см. § 5.5), однако это явление не в состоянии изменить соотношение амплитуд импульсов на последовательных фазах артикуляции согласных. Следовательно, двигательные единицы, лежащие вблизи поверхности мышц и находящиеся в их глубине, получают сигналы неодинаковой временной структуры. В [5.38] было получено экспериментальное доказательство специализации различных участков НКГ для выполнения различных речевых движений. Подобное явление заставляет с осторожностью относиться к выводам, сделанным на основе анализа сигналов только одного типа электродов. По-видимому, адекватное представление об электрической активности мышцы можно получить лишь с помощью многочисленных отведений из различных точек ее объема. Такой объемный анализ помог бы выяснить причины наблюдаемого сокращения длительности команд перед начальными согласными при ускорении темпа. Как видно из табл. 5.1, опережение начала команд перед смычкой начальных слогов в 2—3 раза больше опережения для остальных слогов. При этом амплитуда сигналов перед начальными слогами меньше амплитуды последующих слогов, так что затрачиваемая энергия меньше энергии последующих слогов. Это явление можно интерпретировать как предпочтение критерию ограничения энергии, при котором по возможности осуществляется обмен амплитуды управления сигналами на длительность возбуждения. Если учесть интегрирующие свойства мышцы, то ее сокращение будет примерно одинаковым в обоих случаях, но такое управление потребует меньших энергетических затрат от источника электрического возбуждения. Казалось бы, при ускорении темпа потребная мощность должна возрасти, но амплитуда управляющих сигналов изменяется мало, а их длительность уменьшается. Такое поведение системы управления соответствует тактике, при которой ускорение темпа артикуляции не сопровождается заметным увеличением скорости движения артикуляторных органов.

Ускорение темпа артикуляции подчеркивает эффект противодействия раскрытию губ, демонстрируемый при регистрации проводочными электродами. Непосредственным следствием такого противодействия является повышение жесткости мышечной системы и, как результат, увеличение ее собственной частоты и улучшение динамических характеристик. Для глухих взрывных это явление объясняется необходимостью противодействия внутриротовому давлению. Очевидно, этот же эффект наблюдается при повышении уровня сигнала на фазе смычки глухих согласных при регистрации поверхностными электродами. При взрыве же губной или тем более язычной смычки уже не надо удерживать стремление губ к раскрытию, и, следовательно, увеличение их жесткости должно объясняться необходимостью быстрее развести губы на такое расстояние, где не могут возникнуть турбулентные шу-

мы, маскирующие фонетическое качество звуков. При ускорении артикуляции громкость речи повышается, что увеличивает вероятность турбулизации потока на губах даже при взрыве язычной смычки. Это обуславливает увеличенную реакцию НКГ при артикуляции звуков *Т, Д, Н, З, Л*.

Зарегистрированные изменения временных соотношений нейромоторных команд в зависимости от темпа артикуляции, положения звука в звукосочетании, осмысленности и т. д., очевидно, являются результатом изоморфного отображения пространства нейромоторных команд на пространство артикуляторных движений. При этом одна и та же степень сокращения мышцы может быть достигнута за счет возбуждения разных двигательных единиц [5.57], а один и тот же артикуляторный акт (например, движение губ) может быть осуществлен за счет различной пространственно-временной активности группы мышц [5.127]. Такая свобода в выборе управляющих сигналов создает устойчивость к повреждениям и помехам, например при параличе отдельных мышц, разговоре с пищей во рту и т. п. речь сохраняет разборчивость.

Эксперименты по отведению электрических потенциалов от жевательной мышцы при произнесении бессмысленных и осмысленных звукосочетаний также демонстрируют реорганизацию моторных команд, причиной которой может служить ограничение энергетических возможностей системы управления. Обычно считается, что жевательная мышца не принимает участия в процессах артикуляции. Наши эксперименты показали, что это не совсем так. Действительно, электроды почти не регистрируют электрической активности ни для отдельных звуков, ни для бессмысленных звукосочетаний в любом темпе произнесения. Активности не было и при произнесении осмысленных слов и фраз в нормальном темпе артикуляции. Лишь при быстром темпе на осмысленных словах и фразах возникает некоторая активность, хотя и не очень интенсивная, но достаточно устойчивая и дающая основание для весьма важных выводов.

Активность жевательной мышцы регистрировалась двумя парами проволочных электродов, введенных в левую мышцу на расстоянии примерно 1 см друг от друга. Спектрограммы речевых сигналов, в которых была обнаружена активность этой мышцы, записывались с той же скоростью, что и электромиограммы. Это позволило осуществить точную привязку во времени возникновения электрической активности к звуковому составу речевых сигналов.

Электрические сигналы жевательной мышцы представляют собой короткие группы импульсов или даже отдельные импульсы, форма и время появления которых несколько различались в отведенных потенциалах первой и второй парами электродов. Как упоминалось в § 5.1, двигательные единицы жевательной мышцы обладают относительно большой силой сокращения, поэтому активизация даже небольшого их числа может внести заметный вклад в движение нижней челюсти.

Для щелевых звуков *С, Ш, З, Ж* и аффрикат *Ц, Ч* всплески активности наблюдались только на фазе движения к щели или смычке. Активность на интервале самой щели или перед ее смыканием наблюдалась всего в несколько случаях. Для заднеязычных *К, Г* активность не была обнаружена, она появлялась только для переднеязычных и губных взрывных согласных на фазе как смыкания, так и размыкания. Чаще всего активность на фазе смыкания возникала перед мягким звуком *Т*, а на фазе размыкания — для звуков *Т, Б, В*. Отсутствие активности для заднеязычных согласных объясняется тем, что движения нижней челюсти оказывают на их артикуляционные характеристики малое влияние.

Эти эксперименты указывают на использование различных тактик активизации мышц системой управления артикуляцией. Появление активности жевательной мышцы только при артикуляции осмысленных звукосочетаний свидетельствует о стремлении системы управления сохранить надлежащую разборчивость при возмущениях в виде изменения темпа. Можно выделить три вида активности жевательной мышцы: на фазе движения к смычке, на фазе движения к щели и на фазе взрыва смычки. Активность на фазе смыкания для взрывных может объясняться либо дефицитом мощности мышц, обычно используемых в данной артикуляционной ситуации, либо глобальной оптимизацией какого-либо энергетического критерия. При движении к щели энергетические критерии, возможно, являются не единственной причиной активизации этой мышцы. Поскольку артикуляция щелевых звуков предъявляет высокие требования к точности системы управления, активность жевательной мышцы может свидетельствовать о вовлечении в процесс артикуляции депрессоров нижней челюсти (например, двубрюшной мышцы), тормозящих ее движение на заключительных этапах формирования щели, благодаря чему отсутствует перерегулирование в быстром темпе артикуляции. В описываемых экспериментах активность депрессоров не регистрировалась, но появление нейромоторных сигналов в жевательной мышце на фазе взрыва смычки взрывных звуков говорит о ее появлении. В этом случае сокращение жевательной мышцы как бы противодействует взрыву смычки.

Активность жевательной мышцы на фазе смыкания свидетельствует еще об одной важной особенности системы управления артикуляцией — нейромоторные сигналы появляются до начала или в самом начале движения, т. е. еще до того, как в результате оценки реального движения по некоторому критерию могла бы выявиться необходимость в сокращении этой мышцы. Отсюда следует, что решение о возбуждении жевательной мышцы было принято заранее, до фактического движения, только по входным условиям, задающим скорость артикуляции.

Возвращаясь к обсуждению роли энергетических критериев в формировании управляющих команд, из вышеописанных экспериментов можно заключить, что реорганизация команд при измене-

нии темпа артикуляции вполне может быть связана с оптимизацией одного из этих критериев, поскольку динамические характеристики мышц нелинейны относительно амплитуды управляющих сигналов. Характер этой нелинейности зависит от длины, скорости сокращения и прочих параметров мышц, т. е. различен для разных мышц, поэтому простое изменение темпа артикуляции приводит к смещению оптимума энергетических затрат. Здесь происходит естественное распространение энергетических критериев управления с отдельной мышцы на комплекс мышц, принимающих участие в том или ином артикуляторном движении.

Есть основания полагать, что энергетические критерии действуют и глобально, охватывая все движение артикуляторных органов, совершающихся в процессе артикуляции. Влияние энергетических критериев проявляется в частоте встречаемости звуков в любом языке. Если упорядочить звуки по частоте встречаемости таким образом, чтобы рангу $r=1$ соответствовал наиболее часто встречающийся звук, рангу $r=2$ — следующий по частоте встречаемости звук и т. д., то зависимость между частотой встречаемости p и рангом описывается законом Мандельброта $p(r) = A/(r+B)^\gamma$, где параметры B и γ зависят от языка, а параметр A отражает стоимость генерирования звука. Данный закон был получен из условия максимизации информации, передаваемой сообщением, составленным из элементов некоторого кода, при условии ограниченной стоимости сообщения [5.4, 5.119]. В составляющие стоимости звуков должны входить и энергетические затраты, в частности число одновременно активизируемых мышц. Из рис. 5.22 видно, что ранжированное распределение вероятностей фонем не коррелировано с их помехоустойчивостью, тогда как на первых местах по частоте встречаемости находятся назальные звуки M, N , для генерирования которых не требуется подъем небной занавески, осуществляемый для всех остальных звуков. Видно также, что глухие согласные встречаются чаще звонких, очевидно, потому, что они образуются без участия голосового источника и, следовательно, без активизации мышц гортани.

Известно, что в речи существуют звуки предпочтительные и неpreferchitelnye, артикуляция которых представляет некоторые трудности. Установлено, что в оговорках и ошибках произношения предпочтительные звуки не заменяют неpreferchitelnykh, причем это явление не связано с ошибками домоторной органи-

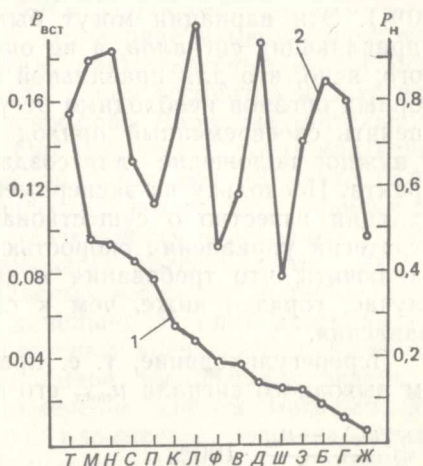


Рис. 5.22. Вероятности встречаемости фонем $P_{вст}$ (1) и субъективной надежности восприятия P_n (2)

зации речи и носит нелингвистический характер [5.118]. Например, наиболее часто встречающийся тип замен — оглушение звонких согласных, особенно в конце слов.

В число критериев качества системы управления входят оценки точности — статической и динамической, величины перерегулирования, длительности переходных процессов. Из § 5.6 следует, что система управления артикуляцией скорее всего обладает астатизмом первого порядка, гарантирующим нулевую статическую ошибку, т. е. разность между требуемым и фактическим положениями артикуляторного органа по прошествии достаточно большого времени после начала переходного процесса. Критерий статической точности играет весьма важную роль при артикуляции фрикативных звуков, где шумообразование возникает лишь при определенном (и весьма малом) расстоянии между языком и нёбом.

Требования к точности тех или иных артикуляционных движений устанавливаются с учетом фонетического состава данного языка. Например, в [5.110] путем сравнения шведского и английского языков было показано, что в шведском языке артикуляция огубленных гласных начинается раньше, имеет большую амплитуду и большую точность, чем в английском языке. В последнем существуют лишь две огубленные гласные, а в шведском — 11. Там же было найдено, что точность артикуляции звуков в осмысленных звуко сочетаниях выше точности их артикуляции в бессмысленных звуко сочетаниях. Точность артикуляции зависит и от разговорного стиля [5.52].

Вопрос о роли динамической точности менее ясен. В процессах речеобразования наблюдаются значительные вариации скорости движения нижней челюсти (до 25%), губ (до 25%) и языка (до 60%). Эти вариации могут быть следствием различий в форме управляющих сигналов, а не ошибок системы управления. Кроме того, ясно, что для правильной координации движений артикуляторных органов необходимо регулировать их скорость, чтобы обеспечить своевременный приход каждого артикуляторного органа в нужное положение для создания требуемой формы речевого тракта. Поскольку из экспериментов с изменением скорости артикуляции известно о существовании различных индивидуальных стратегий управления скоростью артикуляторных органов, можно заключить, что требования к динамической точности, во всяком случае, гораздо ниже, чем к статической точности системы управления.

Перерегулирование, т. е. превышение максимальным значением выходного сигнала u_{max} его стационарного значения u_{∞} ,

$$\sigma = \frac{u_{max} - u_{\infty}}{u_{\infty}} \cdot 100\%$$

играет важную роль в образовании звуков щелевых звуков и тех звуков, которые характеризуются малой площадью поперечного

сечения в отдельных областях речевого тракта. Большие значения перерегулирования могут привести либо к образованию смычки вместо щели, либо к нежелательному изменению акустических характеристик речевого сигнала, например появлению турбулентных шумов. Для таких звуков система управления должна выбирать режим, в котором переходные процессы имеют монотонный (апериодический) характер. Это заставляет либо изменять параметры системы управления, например увеличивать демпфирование, либо уменьшать скорость изменения и амплитуду управляющих сигналов. Взрывные звуки, обладающие стадией смычки, не нуждаются ни в высокой статической точности, ни в уменьшении перерегулирования, поскольку после соприкосновения губ или языка с нёбом дальнейшее движение артикуляторных органов невозможно. При артикуляции согласных допустимы большие скорости изменения управляющих сигналов и соответственно большие скорости движения артикуляторных органов. Уже одно противопоставление требований к системе управления артикуляцией взрывных согласных и фрикативных звуков наглядно демонстрирует необходимость применения различных критериев по точности в различных артикуляционных ситуациях.

В технических системах многие процессы управления требуют реализации минимальной или заданной длительности переходных процессов. Роль первого критерия в речевом образовании пока оценить трудно, хотя известно, что сформулированный на основе минимума длительности переходных процессов принцип максимума Понтрягина [5.26] требует использования управляющих сигналов максимально возможной амплитуды, что входит в конфликт с энергетическими ограничениями, роль которых обсуждалась выше. Более важным для обеспечения координации артикуляционных движений представляется управление длительностью стадий этих движений. Известно, что в ряде языков существуют короткие и длинные гласные, что ударные гласные имеют большую длину, чем безударные, и что длина сегментов играет важную роль в понимании смысла слов и фраз. Вариации длительности коротких фраз составляют примерно 15% [5.41], тогда как вариации длительности отдельных звуков достигают 13—33% [5.49]. По данным наших рентгенографических измерений в зависимости от окружения согласного звука длительность переходных процессов для губ изменяется в наименьших пределах — от 100 до 120 мс, для переднеязычных согласных несколько больше — от 80 до 120 мс, для заднеязычных в довольно широких пределах — от 80 до 180 мс, что объясняется степенью коартикуляционных взаимодействий, увеличивающейся по мере продвижения вглубь речевого тракта. Одновременное измерение электромиограмм, движений артикуляторных органов и параметров речевого сигнала свидетельствует о том, что координация движений в процессе речеобразования осуществляется в значительной мере путем управления длительностью и относительными фазами нейромоторных команд, поэтому временной контроль в управлении артикуляцией,

очевидно, играет важную роль. Методы управления по критерию заданного времени переходных процессов хорошо развиты (так называемое финитное управление) как для сосредоточенных систем, так и для систем с распределенными параметрами, что создает основу для математического моделирования этого аспекта системы управления артикуляцией по мере накопления необходимых экспериментальных данных.

Наряду с критериями, традиционно используемыми для оценки качества систем автоматического управления, в системе управления артикуляцией должны рассматриваться и специфические критерии, связанные с многоуровневым характером процесса речеобразования. Сокращение мышц, которое является непосредственным результатом воздействия нейромоторных команд, служит лишь промежуточным этапом формирования некоторой конфигурации речевого тракта. В свою очередь, конфигурация тракта нужна лишь для того, чтобы обеспечить генерирование акустических колебаний с заданными параметрами. Речевой сигнал служит единственным источником передачи смысловой информации при разговоре людей (при отсутствии визуального контакта), поэтому естественно было бы считать, что конечной целью действий системы управления артикуляцией является создание сигнала с определенными характеристиками. Большое разнообразие значений параметров одного и того же звукосочетания, произнесенного различными дикторами, порождает не только не преодоленные пока трудности автоматического распознавания речи, но и, казалось бы, ставит под сомнение гипотезу акустических целей. В следующем параграфе мы обсудим возможные свойства речевого потока, которые объясняют терпимость аудиторов к значительному разбросу параметров речевого сигнала.

Поскольку исследования единиц восприятия и распознавания речи человеком продвигались слишком медленно, основное внимание исследователей сконцентрировалось на физически измеримых уровнях и задача была сформулирована по-иному: что является целью действий системы управления артикуляцией — формирование заданного вектора акустических характеристик речевого сигнала, заданной площади поперечного сечения речевого тракта или заданных формы и положения артикуляторных органов? Такая постановка задачи породила ряд экспериментальных методов в исследовании процессов речеобразования, основанных на механическом вмешательстве в движениях артикуляторных органов. В некоторых случаях использовались различные протезы типа искусственного нёба [5.40, 5.78], в других — внезапное приложение усилия к нижней челюсти [5.64, 5.65]. Очень важную информацию можно получить, применяя методику так называемой «блокировки нижней челюсти» (*bite—block*), состоящей в помещении твердого предмета между челюстями, так что подъем нижней челюсти ограничивается заранее заданным положением [5.72, 5.105].

В одних случаях изменение формы речевого тракта с помо-

щью протезов приводило к быстрой адаптации (в течение 5 мин) дикторов, нормализовавших акустические характеристики гласных звуков [5.40], тогда как в других адаптация не наступала и после недели ношения протеза [5.78]. Это свидетельствует, с одной стороны, об известной гибкости системы управления артикуляцией, а с другой — об ограниченных возможностях компенсации даже постоянных во времени изменений формы речевого тракта.

Внезапное препятствие подъему нижней челюсти во время артикуляции губного взрывного *П* вызывает дополнительный подъем нижней губы примерно через 10—30 мс, так что губная смычка все-таки образуется. Это явление, во-первых, свидетельствует о том, что система управления артикуляцией имеет в качестве окончательной цели достижение фиксированного для данного звука не положения каждого артикуляторного органа, а формы речевого тракта. Во-вторых, информация о том, что губная смычка не осуществилась вследствие помехи, быстро (по сравнению с самим артикуляторным движением) достигает центров управления и вызывает соответствующие изменения моторных команд к губным мышцам. Наконец, компенсация помех образованию смычки оказывается сравнительно проще для системы артикуляции, чем компенсация искажений формы речевого тракта при генерировании гласных. Очевидно, здесь сказывается пороговый характер формирования смычки «есть — нет», а также снижение требований по точности, которые обсуждались выше.

При использовании блокировки движений нижней челюсти перестройка положений артикуляторных органов (губ и языка) наблюдается уже на первом импульсе голосового источника, т. е. до начала действия акустической обратной связи [5.105]. Было установлено [5.72], что происходит такая перестройка положений артикуляторных органов, которая приводит не к полному сохранению функции площади поперечного сечения речевого тракта, а допускает значительные отклонения на участках с большой площадью поперечного сечения и меньше всего искажает участки с малой площадью. Это наблюдение вновь дает основание для следующих важных выводов: 1. Целью системы управления является создание заданного акустического образа. 2. Цель достигается (когда это возможно) путем компенсации возмущений площади поперечного сечения в наиболее чувствительных к этому возмущению областях. Для достижения этой цели система управления должна быть осведомлена о желаемых акустических характеристиках, фактически имеющихся характеристиках при данном возмущении формы речевого тракта и способах компенсации возмущения без совершения поисковых движений.

Моделирование исходной и возмущенной конфигураций речевого тракта с помощью синтезатора речи привело авторов [5.72] к выводу, что область сужения очень чувствительна к возмущению и именно ее система управления артикуляцией старается сохранить неизменной. В [5.133] сообщается, что наименьший разброс положений языка при артикуляции гласных в направлении,

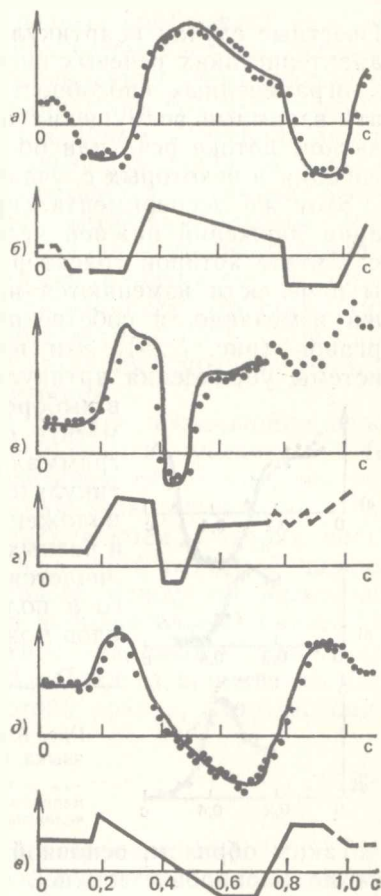
перпендикулярном оси речевого тракта, наблюдается вблизи минимальной площади поперечного сечения тракта, что подкрепляет мнение о наибольшей чувствительности параметров речевого сигнала к изменению этой площади. Обращает на себя внимание то, что относительные различия между нормальной и возмущенной функциями поперечного сечения в области как малых, так и больших значений примерно одинаковы. Поскольку изменения акустических характеристик зависит только от относительного возмущения площади поперечного сечения, можно было бы сказать, что произошла компенсация по всей функции площади поперечного сечения с более тщательной коррекцией в области наибольшего сужения, поскольку именно в ней сильнее всего скажется любое заданное возмущение.

Отсюда следует вывод, что целью подобной компенсации является именно сохранение акустического образа за счет использования компенсаторных возможностей системы управления. Этот вывод согласуется с заключением, сделанным в [5.160] на основе анализа нормированных к одному масштабу рентгенограмм речевых трактов, соответствующих артикуляции близких по акустическим характеристикам звуков речи для 15 языков. Это заключение состоит в том, что положение и конфигурация языка изменяются в зависимости от формы твердого неба диктора с целью производства звуков, принадлежащих к некоторому классу с близкими акустическими характеристиками за счет формирования примерно одинаковых функций площади поперечного сечения. В этой же работе можно найти примеры известной свободы выбора формы речевого тракта для создания близких акустических характеристик — иногда сужение в переднеязычной области сопровождается расширением в фарингиальной области. Как следует из предыдущего параграфа, такое антисимметричное преобразование поперечного сечения не изменяет акустических характеристик сигнала.

Несмотря на существующую свободу выбора формы речевого тракта, не всегда удается достичь желаемого акустического результата. Так, в экспериментах по моделированию управляющих команд с использованием кинорентгенограмм оказалось, что даже в условиях коартикуляции система управления предпринимает меры для наименьшего возмущения акустических характеристик. Эти эксперименты проводились на материале эстонского языка; отдельные слова типа *talapak*, *saņapik*, *annatas*, *panikapa*, *sarrapi* записывались со скоростью 50 кадр./с. На кинорентгенограммах измерялись траектории отдельных точек, расположенных на нижней челюсти, нижней губе, кончике и середине языка (в районе нёбной занавески). В использованной модели движения губы и языка аппроксимировались лишь первыми их модами, а управляющие сигналы выбирались в классе кусочно-линейных функций. Моделирование производилось на аналоговой вычислительной машине, критерием служило сходство измеренных на кинорентгенограммах и вычисленных траекторий точек губ и языка,

Figure 1 consists of six sub-graphs labeled a) through f), each plotting a different parameter against the variable c (ranging from 0 to 1.0).
 - Graph a) shows a parameter that starts at approximately 0.5, rises to a peak of about 0.8 at $c \approx 0.4$, then drops to a minimum of about 0.2 at $c \approx 0.7$, and finally rises to about 0.9 at $c = 1.0$.
 - Graph b) shows a parameter that starts at approximately 0.5, remains relatively constant until $c \approx 0.4$, then drops sharply to about 0.2 at $c \approx 0.7$, and rises back to about 0.5 at $c = 1.0$.
 - Graph c) shows a parameter that starts at approximately 0.2, rises to a peak of about 0.8 at $c \approx 0.4$, then drops to a minimum of about 0.2 at $c \approx 0.7$, and rises back to about 0.8 at $c = 1.0$.
 - Graph d) shows a parameter that starts at approximately 0.5, rises to a peak of about 0.8 at $c \approx 0.4$, then drops to a minimum of about 0.2 at $c \approx 0.7$, and rises back to about 0.8 at $c = 1.0$.
 - Graph e) shows a parameter that starts at approximately 0.2, rises to a peak of about 0.8 at $c \approx 0.4$, then drops to a minimum of about 0.2 at $c \approx 0.7$, and rises back to about 0.8 at $c = 1.0$.
 - Graph f) shows a parameter that starts at approximately 0.5, remains relatively constant until $c \approx 0.4$, then drops sharply to about 0.2 at $c \approx 0.7$, and rises back to about 0.5 at $c = 1.0$.

... — измерение на рентгенограммах;
 — моделирование



... — измерения на кинорентгенограммах: — — — — — моделирование

273

ния динамических спектрограмм речи, в которых отчетливо проявляется влияние второго гласного на акустические характеристики первого гласного в слогах «гласный — согласный — гласный» [5.131]. Этот эффект желательно учитывать и при синтезе речи.

Известные случаи коартикуляции, проявляющейся в акустических характеристиках речевых сигналов, по-видимому, свидетельствуют об ограниченных способностях системы управления компенсировать взаимные возмущения акустических характеристик звуков в слитном потоке речи или об отсутствии потребности в такой компенсации в некоторых случаях.

Этот же экспериментальный материал дает примеры координации движений нижней челюсти и артикуляторных органов, в результате которой траектории совместного движения нижней губы и челюсти изменяются незначительно при существенно большей изменчивости собственных движений этих артикуляторных органов (рис. 5.25). Эти примеры иллюстрируют избыточность системы управления артикуляцией и порождаемую этим свободу

в выборе управляющих команд. Интересно отметить обнаруженную на этих кинорентгенограммах разницу в точности стабилизации артикуляторных органов для разных звуков. Так, положение кончика языка для фрикативного *s* в разных словах *sanapik*, *annatas*, *sarrapi* различается всего на 3,3%, тогда как для гласного *a* положение середины языка для разных слов может отличаться на 30%.

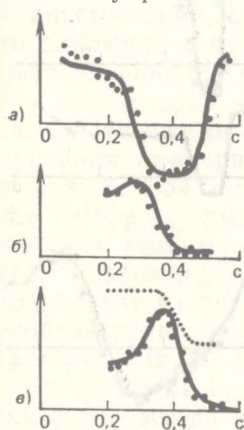


Рис. 5.25. Траектории нижней губы (а), кончика языка (б), середины языка (в).

... — измерения; — — моделирование; ... — измерения движений середины языка совместно с нижней челюстью

Таким образом, основной вывод из описываемых экспериментов по блокировке нижней челюсти и моделированию кинорентгенограмм состоит в том, что целью системы управления артикуляцией является создание заданного вектора акустических параметров $V(a_i)$ речевого сигнала. Для реализации этой цели необходимо располагать способами определения фактических значений акустических параметров, например, с помощью акустической обратной связи или внутренней модели. Для разных дикторов векторы $V(a_i)$, соответствующие одному и тому же звуку, могут различаться, и даже один и тот же диктор в зависимости от контекста и условий речевой связи может назначить в качестве цели различные векторы $V(a_i)$ для одного и того же звука, но в любом случае система управления артикуляцией стремится создать такую форму речевого тракта, которая обеспечила бы достижение желаемых акустических характеристик. Как мы видели ранее, не

все возможные векторы $V(a_i)$ могут быть реализованы системой управления артикуляцией, однако любой вектор из потенциально реализуемого подмножества векторов $U(a_i) \subset V(a_i)$ может быть создан за счет использования различных конфигураций речевого тракта, а каждая из реализуемых конфигураций, в свою очередь, может быть создана за счет разных комбинаций возбуждения мышц. Такая избыточность возможностей системы управления обеспечивает надежность функционирования речеобразующего аппарата, компенсируя некоторые нарушения его деятельности, а также позволяет оптимизировать различные критерии, действующие на разных этапах процесса речеобразования. В зависимости от того, насколько плоской оказывается оптимизируемая характеристика в области экстремума, остается большая или меньшая свобода в окончательном выборе управляющих команд. Эта свобода проявляется в индивидуальности тактик систем управления у разных дикторов, например, касающихся реорганизации команд управления при изменении темпа артикуляции—одни дикторы увеличивают скорость движения артикуляторных органов, а другие — уменьшают амплитуду их движений.

Избыточность системы управления создает чрезвычайно важные преимущества — адаптацию к меняющимся условиям (в частности, к изменяющимся характеристикам самой системы управления) и возможность упрощения законов управления. Эти особенности избыточных систем рассматриваются в теории целенаправленных движений [5.18, 5.19], в которой основной подход состоит в описании кинематических связей между обобщенными координатами системы x_i и наложении дополнительных связей и управлений, обеспечивающих выполнение заданного движения или достижение заданных координат X_i . Один из приемов состоит в использовании так называемого способа аркана, записывающегося в следующем виде:

$$\sum_i (x_i - X_i)^2 = \sigma^2(t). \quad (5.29)$$

Если возможно создание такой функции времени $\sigma(t)$, что в некоторый момент $t=t_0$ $\sigma(t_0)=0$, то обеспечивается совпадение фактических обобщенных координат системы x_i и заданных значений X_i . При этом траектории перехода координат $x_i(t)$, вообще говоря, произвольны и могут быть доопределены путем наложения дополнительных ограничений. Методы этой теории, касающиеся сложных систем, в ряде случаев предлагают более эффективные решения, чем традиционные методы автоматического управления, однако необходимо дальнейшее развитие теории целенаправленных движений применительно к системам с распределенными параметрами, каковой является система управления артикуляцией.

Нетрудно видеть, что левый член в (5.29) представляет собой евклидово расстояние в пространстве обобщенных координат x_i . Поэтому метод аркана легко обобщить на случай произвольной метрики как $\rho(x_i, X_i) = \sigma^p(t)$, где ρ — расстояние между

векторами x_i и X_i , или даже на случай неметрических пространств в виде некоторой меры близости $M(x_i, X_i) = \sigma(t)$. Например, учитывая логарифмическую шкалу восприятия частот человеком, меру близости между фактическими f_i и требуемыми резонансными частотами F_i генерируемого речевого сигнала можно было бы определить как $M(f_i, F_i) = \sum_i k_i |\lg f_i - \lg F_i|$ и использовать этот функционал в качестве одного из критериев системы управления артикуляцией на акустическом уровне.

Для системы управления движениями человека характерна кооперация различных мышц или даже органов для достижения заданной цели, как это проявляется, например, в компенсации движениями глаз изменений положения тела. В системе управления речеобразованием такая многоконтурная система управления, по-видимому, действует постоянно с целью уменьшения влияния коартикуляции на акустические параметры речевых сигналов. Эта же схема проявляется и в обсуждавшихся ранее случаях компенсации блокировки нижней челюсти движения языка и губ.

Все методы современной теории автоматического управления, в том числе и теории целей, требуют знания структуры и уравнений, описывающих поведение элементов системы управления для синтеза управляющих сигналов. Самонастраивающиеся системы управления могут работать при первоначальном отсутствии этой информации, но совершаемые ими поисковые движения позволяют в ряде случаев неявно определить передаточную функцию объекта управления. В речеобразовании поисковые движения наблюдаются в экспериментах по имитации [3.9] при различных нарушениях речи, но, по-видимому, нормально эти механизмы у взрослых людей не действуют. Как было показано ранее, пересчет сигналов обратной связи в управляющие команды и синтез управления, оптимального по некоторому критерию, невозможны без знания свойств управляемых объектов — артикуляторных органов и акустических процессов в речевом тракте. В физиологии движений человека ряд аналогичных проблем привел к возникновению гипотезы «внутренней модели» [5.12, 5.39, 5.93, 5.139, 5.141], с помощью которой система управления могла бы предсказать результат предполагаемого управления, не совершая движения фактически.

В системе управления речеобразованием необходимость внутренней модели становится очевидной, как только возникает вопрос о коррекции артикуляторных движений по сигналам акустической обратной связи. Сигналы рассогласования между желаемыми акустическими характеристиками речевого сигнала и его реальными параметрами могут быть пересчитаны в артикуляторное управление только в том случае, когда нервная система располагает и моделью артикуляции, и акустической моделью речеобразования для того, чтобы можно было предсказать акустические последствия того или иного изменения конфигурации речевого тракта.

Гипотеза внутренней модели снимает проблему задержки в цепях обратной связи. Как ранее упоминалось, общая задержка сигналов обратной связи может достигать 20—50 мс, что в некоторых случаях может либо серьезно ухудшить точность системы управления, либо поставить ее на грань устойчивости. Если же предположить, что сигналы обратной связи от механорецепторов мышц и связок используются не только для коррекции медленных движений, но и для настройки внутренней модели, то вырабатываемые этой моделью сигналы предсказания параметров системы речеобразования позволяют не только исключить задержку, но даже ввести опережение, которое действительно иногда наблюдается в γ -каналах еще до совершения какого-либо движения.

Анализируя возможности внутренней модели, мы приходим к чрезвычайно важному выводу: та часть внутренней модели, которая преобразует сигналы акустической обратной связи в нейромоторные команды, может быть использована для распознавания речи других людей. В предыдущих разделах мы видели, что, хотя заданный вектор акустических параметров речевого сигнала может быть порожден множеством конфигураций речевого тракта и множеством управляющих команд еще большей мощности, имеются дополнительные ограничения, которые существенно сужают неоднозначность обратного преобразования — от акустики к конфигурации речевого тракта или нейромоторным командам. Эти ограничения создают регуляризацию, необходимую для решения некорректных (по Адамару) задач [3.8]. Акустический блок внутренней модели должен быть настроен, конечно, на параметры речеобразующего тракта самого диктора, однако воспринимаемая речь других дикторов может либо анализироваться в терминах этой же модели, либо, если удастся выделить информацию, касающуюся персональных характеристик речевого тракта другого диктора (например, его длины), то возможна предварительная нормализация с целью уменьшения ошибок анализа.

МОДЕЛИ ЦЕНТРАЛЬНЫХ МЕХАНИЗМОВ

На протяжении большей части истории развития теории речеобразования основные усилия были направлены на поиски некоего единого жестко детерминированного механизма, управляющего характеристиками речевых сигналов. В частности, главная задача подавляющего большинства теорий состояла в обнаружении инвариантов, т. е. такого описания речевых сигналов, которое непосредственно соотносилось бы с какими-либо дискретными единицами речевого потока, например фонемами, и не зависело бы ни от особенностей произношения дикторов, ни от контекста.

Все теории речеобразования стараются устранить противоречие между способностью человека членить речевой поток на одномерную последовательность дискретных единиц и очевидной не-

прерывностью процессов речеобразования на всех доступных измерению уровнях. Невозможность создания алгоритма, выделяющего из потока речи фонему, привела к усиленному изучению явления коартикуляции — взаимного влияния характеристик звуков. Впервые отчетливые экспериментальные наблюдения за коартикуляцией на спектрально-временном уровне были описаны в [5.131], а вслед за этой работой появились исследования коартикуляции и на артикуляционном, и на нейромоторном уровнях. Были обнаружены эффекты влияния звуков как типа предсказания (справа налево по оси времени), так и типа последствия (так называемый лево-правый эффект). Эти эффекты возникают в разных сочетаниях звуков и зависят от языка [5.58].

Наиболее сильно взаимное влияние звуков оказывается на их ближайших соседях справа и слева, поэтому большой интерес вызывают модели, принимающие в качестве единицы речеобразования слог. В этих моделях фонема появляется как результат длительного процесса обработки речевого сигнала на различных уровнях и служит главным образом для экономной записи речи. Это представление подкрепляется сведениями о том, что время принятия решения о фонеме оказывается большим, чем о слоге (например, для *s* — на 40 мс, для *p* — на 70 мс, а для *x* — даже на 250 мс [5.140]).

Среди работ, посвященных доказательству реальности слога как единицы программы речеобразования, наибольшее влияние на направление исследований оказала [3.9]. В ней приводятся аргументы в пользу того, что единицей речевой программы является открытый слог типа *СГ* (согласный — гласный). Эти аргументы основывались на экспериментальных наблюдениях за огублением согласных в последовательностях *ССС ... Г*, когда гласная *Г* была огублена (например *У*). Позднее этот ряд был представлен в виде *СⁿГ*, где *n* — число согласных, и изучалось влияние параметра *n* [5.44]. Работа [3.9] вызвала много исследований как подтверждающих, так и опровергающих представление о слоге *СГ* как о единице программы речеобразования. Интересно отметить, что слог *СГ*, является полной противоположностью закрытому слогу *ГС*, к которому пришел Бауман (*Bowman*) на основе анализа роли тактильных рецепторов в движениях языка [5.49].

Вместе с тем были получены доказательства о распространении право-левого эффекта коартикуляции через согласный на предшествующий гласный в слогах *СГС* на спектрально-временном уровне [5.131], а также на кинорентгенограммах процесса артикуляции [5.90]. В [5.44] на материале французского языка было обнаружено влияние последнего огубленного гласного на первый гласный в последовательности *istrstry*. В [5.124] описывается движение нёбной занавески на гласном *i* во фразе *frianterio* (*free Ontario*), демонстрирующее влияние назализации через гласный *a*. Аналогичные эффекты влияния назального согласного по крайней мере на три сегмента влево (против хода времени) найдены в языке Хинди [5.59]. Такие факты обнаружены

и в движениях отдельных артикуляторных органов, и в сокращении мышц нижней челюсти и нижней губы [5.115, 3.49]. Все эти данные заставляют усомниться в том, что левая граница открытого слога может быть установлена путем лишь исследования коартикуляции. Скорее всего коартикуляция связана не с какой-либо формой звуко сочетаний, а с самими звуками и критериями управления процессами речеобразования. Наряду с этим открытый слог *СГ* представляется хотя и не единственным, но весьма важным элементом организации речевой программы. В [5.77], однако, высказывается мнение о том, что понятие коартикуляции имеет смысл только в том случае, если различаются фонемы и их аллофоны, т. е. что само это понятие носит терминологический характер, тогда как для определения единиц речевого потока нужно одновременно учитывать артикуляцию, акустику и восприятие.

После того, как было установлено отсутствие инвариантности как на акустическом, так и на артикуляционном уровнях, некоторое время возлагались надежды на обнаружение инвариантности на нейромоторном уровне. Однако интенсивные исследования электрической активности речевых мышц разбили и эти надежды. Более того, как показывают экспериментальные исследования и проведенный в предыдущих параграфах анализ измеримости и управляемости процессов речеобразования, разнообразие параметров растет от акустического уровня к артикуляторному и нейромоторному. Реакцией на крушение идеи инвариантности явилась теория контекстно-чувствительных элементов Викельгрена (*Wickelgren*) [5.159], провозглашающая существование стольких видов нейромоторных команд, сколько существует трехфонемных сочетаний. Эти же сочетания предлагаются и как основа ассоциативного распознавания. Например, последовательность «стол» представляется как $\#c_t - c_o - t_o l \#$, где $\#$ — знак паузы. В такой системе распознавание какого-либо элемента, например « c_o », облегчало бы дальнейшее распознавание, возбуждая по ассоциации, начинающиеся с «о». Согласно этой теории на запоминание всех контекстно-чувствительных единиц потребуется от 10^4 до 10^6 ячеек или нейронов. Само по себе требование такого объема памяти еще не может служить аргументом против этой теории, поскольку нам неизвестны критерии эффективности кодирования нейромоторных команд. Однако по [5.92, 5.115] теория Викальгрена является простой констатацией наблюдаемого многообразия нейромоторных команд и не объясняет эффектов коартикуляции, распространяющихся дальше ближайшего окружения. Более того, эта система не позволяет реорганизовать артикуляторные движения при механических помехах. К тому же значительная часть многообразия нейромоторных команд может быть получена гораздо более экономными средствами за счет использования механизмов обратной связи. Тем не менее представляется правдоподобным, что наиболее часто употребляющиеся артикуляторные движения могут кодироваться или восприниматься как

единое целое. Например, в [5.17] показано, что комбинации менее 200 слогов русского языка покрывают около 80% объема всех встречающихся текстов на русском языке.

Существование центров, управляющих сложными комплексами движений путем задания лишь одного параметра, было эффективно доказано в экспериментах на обездвигнутых животных, у которых были перерезаны все нервные проводящие пути, ведущие от мозга к лапам [5.34]. Были обнаружены такие точки, непосредственное электрическое стимулирование которых приводило к координированному движению лап кошки, имитирующему шаг, бег или галоп, причем смена характера движения происходила лишь за счет изменения амплитуды возбуждающего сигнала. В моторных зонах коры мозга пока не обнаружено таких точек, стимуляция которых вызывала бы специфические речевые движения, но это не служит доказательством их отсутствия.

Идея инвариантности на уровне конфигурации речевого тракта привела к появлению теории «целей», согласно которой существуют идеальные (или предельные) формы тракта, а конкретные команды к их достижению формируются с учетом текущего положения артикуляторных органов с использованием обратной связи, предположительно γ -системы. Первоначально такая идея была высказана в [5.151], а затем развита в [5.114, 5.115]. Согласно [5.114], если имеется последовательность звуков *ABCDEF*, то формирование команд начинается справа налево, т. е. со звука *F*. Команды для звука *E* формируются с учетом необходимости достижения конфигурации речевого тракта, соответствующей *F*, команды для *E* учитывают звук *E* и т. д. Такая модель объясняет различие конфигураций речевого тракта для одного и того же звука в разном окружении простым «недорегулированием», т. е. недостижением заданной идеальной конфигурации. В этой модели влияние отдаленных звуков на предшествующие должно падать, однако пример из [5.44], где в последовательности *istrstry* огублялся первый гласный *i*, противоречит этому. Эта модель также подвергается критике за то, что она устанавливает лишь пространственные цели и пренебрегает динамикой, играющей важную роль, например, для дифтонгов и взрывных.

Исследования синтетической речи, проводившиеся в Хаскинских лабораториях, свидетельствуют о том, что в некоторых случаях информативным признаком для восприятия является не сама конечная цель, которая может быть и не достигнута, а направление движения [5.102].

Модель идеальных целей, но уже на уровне признаков фонем была предложена Хенке (*Henke*) в 1966 г. К сожалению, эта модель описывается только в его неопубликованной диссертации [5.83]. Однако она произвела большое впечатление и вызвала как ряд теоретических комментариев к ней, так и экспериментальных исследований [5.92]. По этой модели значения признаков задаются скачком и сохраняются в течение некоторого промежутка времени, тогда как двигательный аппарат непрерывно отрабаты-

вает заданные цели. Как видно, здесь инвариантное описание отнесено на еще более абстрактный уровень признаков фоном, а промежуточным механизмам, преобразующим эти цели в артикуляционные движения, отводится второстепенная роль.

В модели Хенке, так же как и в модели артикуляционных целей, «недорегулирование» может быть следствием быстрой смены самих целей. Важным преимуществом модели является ее способность объяснить появление некоторого признака от последующего звука, как это происходит с огублением или назализацией. Предполагается, что если для звука в данной позиции часть признаков не определена или безразлична, то признаки заимствуются из последующих звуков. Таким образом, в модели принимается, что речевая программа реализует все различительные признаки, не вступающие в противоречие. Известны, однако, примеры, когда признак от последующего звука противоречит текущему признаку, но тем не менее замещает его. В частности, в [5.44] отмечается, что в 50% случаев в последовательности *istrstry* огубляется и первый гласный звук *i*. Подобный эффект мог бы возникнуть в других языках, где признак огубленности передних гласных избыточен, но во французском языке существуют как огубленные, так и неогубленные гласные, и, следовательно, этот признак играет различительную роль. Наблюдалось также движение небной занавески к закрытию в слогах «назальный — неназальный» вместе с движением языка к позиции назального звука [5.44]. Таким образом, и модель Хенке имеет ограниченную область применения.

Обычно эффект предсказания в коартикуляции распространяется на большее число звуков, чем эффект последствия, и для объяснения этого в [5.124] выдвигается гипотеза о том, что эффект предсказания помогает предвидеть следующий сегмент, т. е. важен для восприятия. Иными словами, предполагается, что прав-левый эффект коартикуляции создается специально для того, чтобы облегчить распознавание речи другими людьми. И хотя эффект предсказания не столь распространен, как эффект последствия, имеются языки, например Хинди, в которых эффект последствия распространяется столь же далеко, как и эффект предсказания [5.115]. Скорее всего эти эффекты коартикуляции имеют разную природу. Эффект предсказания возникает на высших уровнях системы управления для обеспечения надлежащей разборчивости последующих звуков, т. е. создается активно, тогда как эффект последствия является результатом инерционности исполнительных органов и ограничений, наложенных на управляемость артикуляционных процессов [5.45, 5.132].

Ряд сведений о структуре речевого потока можно извлечь из анализа ошибок, появляющихся при речеобразовании, и особенно нарушений речевой деятельности при различных поражениях центральных и периферических механизмов речеобразования [5.1, 5.6—5.8, 5.31]. Интерпретация речевых нарушений представляет собой значительную трудность: во-первых, одни и те же

внешние признаки могут быть результатом поражения различных механизмов; во-вторых, суждения о работе исправного механизма по наблюдениям за нарушениями могут оказаться ненадежными. Вместе с тем клиническая практика служит почти единственным источником сведений о речевых нарушениях, позволяющим судить о некоторых механизмах высших отделов управления речеобразованием. В частности, эти сведения позволяют оценить роль обратной связи в речеобразовании. Например, в [5.7] описывается больной, который пытался найти правильную артикуляцию, наблюдая в зеркало за положением языка и губ, зажимая нос и, контролируя работу голосового источника, дотрагиваясь до гортани. При этом он произносил немногие доступные ему слова «папа», «мама», «дом» и сравнивал их звучание с положением артикуляторных органов. Другие больные в заданиях по имитации речи или чтении проговаривали вслух или «про себя» заданный текст. Наблюдаются нарушения акустической обратной связи для речи при сохранении восприятия неречевых звуков [5.31]. Известны и обратные случаи — сохранение слуха на речевые звуки при отсутствии восприятия неречевых звуков, таких как кашель, мяуканье кошки, свист, плач ребенка, журчанье воды, тиканье часов и т. д. Иногда речь перестает восприниматься при помехах, в шумах помещениях и при разговоре одновременно нескольких людей [5.1, 5.31].

Аналогично примерам раздельной обработки речевых и неречевых звуков слуховым анализатором имеются сведения о раздельном управлении артикуляторными органами в процессах речеобразования и в рефлексных движениях. Например, при сохранении сокращения мышц мягкого нёба в глотательном рефлексе может наблюдаться их вялость при артикуляции гласных звуков или же способность к нормальным мимическим движениям сочетается с расстройствами произвольных движений губ языка и нижней челюсти [5.8]. Эти данные укладываются в общefизиологическое представление об управлении одними и теми же движениями из разных центров в зависимости от задачи [5.2, 5.10].

Очевидная сложность организации речевой программы находит отражение в различных иерархических моделях. Например, в [5.103, 5.155] рассматриваются уровни слогов, фонем, признаков и мышечных сокращений. Предполагается, что все изменения, связанные со взаимным влиянием звуков, происходят только внутри слога. Кроме того, принимается, что на уровне нейромоторных команд все-таки существует инвариантное представление, а различительные признаки непосредственно связаны с работой отдельных мышц. Из обсужденных ранее экспериментальных данных о коартикуляции и измерений ЭМГ мышечных потенциалов видно, что это далеко не так. Представление об иерархической организации речевой программы весьма популярно, однако зачастую предлагаемые модели ограничиваются простым перечислением уровней без раскрытия взаимоотношений между ними на физиологическом уровне, либо же включают в себя все: фонологию,

морфологию, семантику — и дают настолько общее описание, что оно годится для объяснения любых фактов, но не имеет никакой способности к предсказанию.

Одна из иерархических моделей, описания в [5.92], довольно интересна. Она в известной мере учитывает преимущества и недостатки ранее предложенных моделей и охватывает широкий круг речевых явлений.

При попытке описать общую схему процесса речеобразования прежде всего необходимо исключить идею инвариантности. Эта идея родилась в результате переноса представлений о письменной речи на речь разговорную и кажущейся легкости субъективного членения речевого потока на фонемы. Описание речевого потока в виде последовательности фонем, имеющих постоянные акустические характеристики предлагало четкую и простую схему, реализовать которую как для автоматического распознавания, так и для синтеза речи, казалось, было лишь делом техники. Когда же выяснилось, что акустические характеристики звуков сильно зависят от контекста и диктора, поиски инвариантного описания были перенесены на артикуляторный уровень. Рентгенографические исследования показали отсутствие инвариантности и на этом уровне, и тогда ее стали искать в нейромоторных командах. Там ее тоже не нашли, и, исчерпав, таким образом, все количественно измеримые уровни, перешли к конструированию абстрактных моделей, стремясь во что бы то ни стало сохранить инвариантность в виде либо признаков, либо других единиц типа сегмента, слога и т. д. Ранее было показано, что наряду с ограничениями различного рода система управления артикуляцией обладает и значительной свободой выбора тактики управления и что это не сужает, а, наоборот, расширяет функциональные возможности и надежность речеобразующего аппарата.

5.8. ЭЛЕМЕНТЫ КОДОВОЙ СТРУКТУРЫ РЕЧЕВОГО ПОТОКА

В предыдущих параграфах мы пришли к заключению, что на физически измеримых уровнях конечной целью действий системы управления артикуляцией является создание акустического сигнала с заданными параметрами. Конечная цель достигается через промежуточные цели на уровнях площади поперечного сечения речевого тракта, положений, артикуляторных органов, активности мышц и нейромоторных команд. Мы видели, что по мере продвижения от акустического уровня к нейромоторному растет свобода выбора промежуточных целей, являющаяся результатом избыточности в системе управления и характера протекающих физических процессов. Однако она ограничивается критерием физической реализуемости и дополнительными критериями эффективности. В связи с этим возможности системы управления артикуляцией для компенсации внешних возмущений и раз-

решения внутренних конфликтов коартикуляции ограничены. Возникает вопрос: как сказывается на процессе речевого общения невозможность достижения тех или иных акустических характеристик речевого сигнала? Попытки ответить на него порождают вопрос более общего характера: действительно ли конечная цель системы управления состоит в формировании неизменных для данного звукосочетания положений или траекторий вектора акустических параметров?

Если принять во внимание тот факт, что конечной целью речевого общения на самом деле является передача некоторого мысленного образа, то станет ясно, что акустический сигнал служит лишь переносчиком информации, искажения и потери которой восстанавливаются за счет избыточности речевого потока как кодового сообщения. Эта избыточность включает в себя помимо рассматривающихся ранее физических ограничений также ограничения на семантическом и грамматическом уровнях. Отсюда следует, что механизмы коррекции информации, утерянной в речевом сигнале, должны действовать и в условиях отсутствия помех и искажений в канале передачи речевых сообщений. Корректирующая способность на высших уровнях речевого потока позволяет однозначно восстанавливать речевые сообщения, обладающие большим разнообразием акустических характеристик. Это дает возможность системам управления артикуляцией разных дикторов выбрать в качестве акустических целей разные векторы параметров, да и один и тот же диктор может позволить себе отработку заданных акустических целей с разной степенью точности в зависимости от межличностных отношений, эмоционального и физического состояния, условий речевой связи и т. д.

Таким образом, исследования процессов речеобразования не могут ограничиться лишь нейромоторным, механическим и акустическим уровнями, а в определенный момент смыкаются с исследованиями процессов восприятия речи.

Если речевой поток представить как последовательность символов некоторого кода [5.15, 5.29], то открывается возможность применения методов, развитых в теории кодирования, хотя следует сразу же указать на некоторые отличия речевого кода от обычно рассматриваемых кодов, например в нем не обнаружено никакой алгебраической структуры. В речевом коде слово является наименьшей смысловой единицей, слоги — единицей артикуляционной программы, а фонетические признаки — единицами артикуляторных движений. Фонема также служит элементом речевого кода, но скорее всего появляется уже после анализа слогов. Рассмотрим кодовые свойства каждого из упомянутых уровней по их способности исправлять ошибки, возникающие либо в результате помех в канале речевой связи, либо в результате погрешностей системы управления артикуляцией.

Эксперименты по восприятию речевых и неречевых стимулов [3.9, 5.75] показывают, что пропускная способность мозга близка к 2 бит/символ, а оперативная память — к девяти символам. При

увеличении размерности пространства признаков разрешающая способность по каждому признаку падает до двух-трех градаций. Способность человека анализировать большие массивы информации может объясняться перекодированием [5.22], в процессе которого группе входных символов присваивается символ более высокого уровня, эти символы вновь группируются и т. д. до снижения объема информации до приемлемого уровня. Аналогичная идея многократной группировки символов в технике связи привела к созданию каскадных кодов, отличающихся высокой исправляющей способностью и простотой декодирования.

Указания на то, что в формировании структуры речи сложность процессов генерирования и анализа речевых сигналов играет столь же большую роль, как и надежность речевой связи, приводят к некоторым важным следствиям. Прежде всего это дает основание для анализа структуры речевого кода методами, развитыми в теории кодирования, и позволяет применять в подходящих случаях соответствующие приемы декодирования. Отсюда также следует, что для экономии на каждом уровне речевого потока все имеющиеся связи могут не использоваться и что некоторые уровни могут обладать даже нулевой исправляющей способностью, лишь обнаруживая ошибки.

Одно из практических следствий такого подхода состоит в установлении аналогии между речевым кодом на уровне слов и кодами Шеннона — Фано и Хафмена. Распределение вероятностей, полученное максимизацией информации при ограниченной стоимости и предписывающее слову тем меньшую вероятность появления, чем оно «дороже», как это имеет место в законе Мандельброта, можно рассматривать как обращение принципа, используемого в этих кодах, по которому наиболее вероятным словам присваиваются более короткие коды, а наименее вероятным — длинные.

Как известно, коды Шеннона — Фано и Хафмена являются неприводимыми (или префиксными) кодами, обладающими тем свойством, что никакое кодовое слово не является началом другого, и это свойство позволяет выделять кодовые слова из непрерывной последовательности символов без каких-либо разделяющих сигналов. И хотя неприводимость не является обязательным признаком кодов, построенных по принципу «наиболее вероятные слова — наиболее короткие», можно предположить, что и реальный язык в известной степени подобен неприводимым кодам.

Проверка этого предположения проводилась с использованием словаря русских литературных слов, содержащего 2500 наиболее часто встречающихся слов, записанных в фонетической транскрипции. Слова совмещались по начальным фонемам и попарно сравнивались друг с другом. Подсчет числа слов, которые служат началом других, показал, что таких слов всего 170 из 2500 (6,3%), и, таким образом, можно считать, что на уровне слов речь обладает свойством неприводимости. Общая вероятность появления слов-префиксов равна 0,278 за счет наиболее часто встречающихся

одно- и двухбуквенных слов, без учета которых эта вероятность равна 0,097. Данный результат имеет очень важное практическое значение, поскольку все известные приемы сегментации слитного речевого потока на слова с использованием просодических характеристик (интонации, ударения) или опирающиеся на вероятности начальных и конечных звуко сочетаний не приводят к решению этой проблемы. Подобные косвенные оценки, безусловно, полезны, но они лишь дополняют механизм декодирования, основанный на префиксном свойстве слов. Важно также и то, что алгоритмы декодирования кодов Шеннона—Фано и Хафмена являются самокорректирующимися относительно границ между кодовыми словами, и это позволяет производить декодирование даже в том случае, когда начало слова по тем или иным причинам не было принято.

Реальность существования артикуляционных признаков как различительных признаков фонем удостоверяется в результате анализа процесса речеобразования и некоторых особенностей процесса восприятия. Некоторые артикуляционные признаки отчетливо проявляются в процессе речеобразования: включение голосового источника, положение небной занавески, смычка и т. д. Эти признаки можно выявить и в процессе восприятия. Так, в матрицах переходов фонем (табл. 5.2), полученных при субъективном восприятии слогов ГСГ в условиях маскировки белым шумом, видно, что реализуются далеко не все возможные переходы, а ошибки совершаются за счет маскировки одного-двух артикуляционных признаков.

Известны эксперименты по кратковременному визуальному запоминанию слогов, в которых матрицы переходов букв в некоторой степени аналогичны матрицам переходов фонем так, как если бы происходило перекодирование зрительных образов в слуховые.

Все артикуляционные признаки связаны с мышечными сокращениями в голосовом аппарате, однако проявление их в изменении геометрии голосового тракта и влияние на акустические характеристики речевого сигнала отличаются большим разнообразием и требуют специфических приемов анализа каждого. Некоторые артикуляционные признаки по характеру сопровождающих их физических процессов тяготеют к двоичной градации (включено — выключено). Это признаки голосового источника, шумового источника, назальности, смычки, автоколебаний языка для звука Р. Другие артикуляционные признаки описывают различие форм голосового тракта, причем согласные звуки характеризуются локальным, концентрированным на небольшом участке изменением конфигурации тракта (так называемым «местом артикуляции»), тогда как гласные отличаются более общим характером изменений. Кроме того, имеются артикуляционные признаки, характеризующие движение артикуляторных органов. Все эти признаки имеют более сложную связь с акустическими характеристиками речевого сигнала и проявляются главным образом в положении и траекториях формант.

Принятые звуки. %

	Б	Д	Г	Ж	З	В	Л	М	Н	П	Т	К	Ш	С	Ф	Х	Я	—	Ч	Ц
Б	76,1	13,3	6,5		0,4	0,9	0,7										0,8	1,3		
Д	13	71,2	10,6		0,2	0,7	0,4										2,9	0,7		
Г	14,8	19,4	53,3		2,7	2,5	0,2		0,2								5,9	1,2		
Ж	0,2	2,5	4	45,5	24,5	7,2	0,8										13,7	1,6		
З	2	11,5	8,9	4,2	41,9	10,4	2,3	0,1	0,1	0,1							16,8	1,5		
В	3,9	5,1	12,2	0,7	4,5	48,2	5,7	0,4	0,3			0,2					16,7	2,1		
Л	0,1	0,3	0,3	0,1	0,5	0,6	96,4	0,5	0,3								0,5	0,1		
М	0,1		0,1			0,1	2,4	65,7	0,6	0,1							1	1,3		
Н	0,1	0,1				0,1	2,9	10,8	29,2	0,2							1,8	1,8		
П	0,1	0,2	0,2	0,1		0,3		0,1	82,7	56,3	14,2	6,3	0,7	5,7	11,2	2,6	0,4	1,4	0,2	
Т			0,1		0,1					11,9	60,5	6,2	1	10,7	6,7	0,7	0,2	1,7	0,2	
К								0,1	0,1	11,4	19,7	46,1	1,3	9,4	7,9	2,2	0,1	1,4	0,5	
Ш			0,1					0,2		2,2	12,4	2,3	41,1	30,6	7	1,4	0,2	2	0,5	
С		0,1	0,1					0,1		4,6	19,7	2,8	3,4	56,7	8,2	1,5	0,4	1,4	0,6	0,4
Ф			0,1						0,1	16,4	12,8	7,9	1,6	15,2	37,6	6,5	0,5	1,1	0,3	

Передаваемые звуки

Таким образом, артикуляционные признаки по сложности анализа разбиваются на две группы: для одной достаточно использовать непосредственно акустические характеристики речевого сигнала, а другой — необходим «анализ через синтез» с целью восстановления формы голосового тракта. Соответственно этим группам признаков существует две теории восприятия — акустическая и моторная. Согласно акустической теории качество фонем определяется исключительно их акустическими различительными признаками. Моторная теория предполагает, что в процессе анализа речевого сигнала определяются параметры артикуляции, которые и служат первичным описанием фонетического состава речи. Обе теории подкрепляются экспериментальными данными, но ни одна из них не смогла занять доминирующего положения, что, по-видимому, отражает объективное различие в свойствах артикуляционных признаков.

Сведения о надежности того или иного артикуляционного признака можно получить из психоакустических экспериментов с различными видами маскировки. В табл. 5.3 показана вероятность

Таблица 5.3

Маркировка и искажение	Вероятность правильного приема			
	ГИ	Н	ШИ	МА
Белый шум:				
английский язык	0,97	0,946	0,815	0,707
русский язык	0,987	0,942	0,667	0,680
Импульсная помеха, русский язык	0,859	0,76	0,77	0,674
Полосовое ограничение, Гц:				
200—2500	0,981	1,00	0,893	0,814
200—1200	0,963	0,996	0,864	0,714
Клиппирование, русский язык	0,774—0,964	0,785	0,931	0,699
Средняя вероятность	0,928	0,905	0,823	0,716

правильного приема признаков: голосового источника (ГИ), назальности (Н), шумового источника (ШИ) и места артикуляции (МА) по результатам аудиторных испытаний при маскировке белым шумом для английского и русского языков и импульсной помехой с частотой следования около 120 Гц при полосовом ограничении 200—2500 и 200—1200 Гц с отношением сигнал/шум +12 дБ и при клиппировании в полосе до 7 кГц. На рис. 5.26 показана зависимость вероятности субъективного выделения артикуляционных признаков от отношения сигнал/шум. Из этого рисунка видно, что признаки звонкости и назальности наиболее надежны и мало зависят от уровня помех. В других условиях восприятия, где большую роль играет реверберация, наблюдается заметный процент переходов звонких и глухих звуков друг в друга.

Определив фонему как совокупность артикуляционных признаков, можно подсчитать исправляющую способность речевого

кода на уровне фонем. Если число несовпадающих артикуляционных признаков назвать кодовым расстоянием d , то условие обнаружения ошибок кратности $Q_d \leq d-1$, а условие исправления ошибок кратности $2Q_c \leq d-1$. Составляя таблицу кодовых расстояний фонем и вычисляя по ней спектр взаимных расстояний (рис. 5.27), можно убедиться, что на фонемном уровне потенциально возможно обнаружить около 75% и исправить около 37,2% одиночных ошибок, а число обнаруживаемых двойных ошибок меньше 10%. Стертые артикуляционные признаки можно восстановить на уровне фонем, причем число восстанавливаемых признаков в данном случае равно 75%. Это значит, что исправляющую способность фонемного кода выгоднее использовать для восстановления стертых признаков. Как правило, стертый признак легче восстановить, чем исправить ошибку, потому что место стертого признака точно определено.

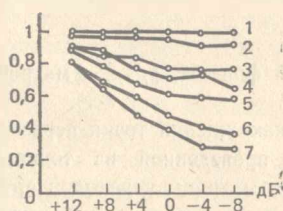


Рис. 5.26. Субъективная вероятность правильного приема артикуляционных признаков:

1 — голосовой источник; 2 — назальность; 3 — переднеязычный; 4 — шумовой источник; 5 — губной; 6 — заднеязычный; 7 — переднеязычно-нёбный

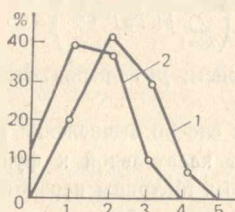


Рис. 5.27. Спектр кодовых расстояний фонем:

1 — полное описание; 2 — без признака места артикуляции

На рис. 5.27 показано, как меняется спектр кодовых расстояний фонем, если признаки места артикуляции станут неразличимыми. При этом 12,4% фонем будут иметь одинаковые коды, обнаружение одиночных ошибок упадет до 47,7%, а их исправление — до 10,5%. Как видно, соотношение между числом восстанавливаемых стертых признаков и числом исправляемых ошибок здесь также в пользу первого.

Вероятность правильного приема артикуляционных признаков в экспериментах по субъективному восприятию при маскировке заметно различается. Это означает, что характеристики артикуляционных признаков при их генерировании и восприятии оказываются зависящими друг от друга. Очевидно, эта зависимость частично является следствием уже обсуждавшегося свойства речевого сигнала — влияния различных артикуляционных признаков на один и тот же акустический признак. Декодирование фонем по зависимым признакам сложнее декодирования по независимым признакам, поэтому нужно оценить влияние зависимости признаков на надежность декодирования слов. С этой целью построим

следующую математическую модель речевого кода на уровне фонем: при формировании слов последовательность фонем выбирается случайной, но зависящей от τ предыдущих фонем и переходы фонем под действием помех независимы. В такой модели можно оценить верхнюю и нижнюю границы надежности декодирования слов при различных способах декодирования фонем с учетом и без учета зависимости артикуляционных признаков.

Используя теорему о кодировании, можно определить потенциальную вероятность ошибки $p(e)$ при различении слов, обладающих кодовыми связями на τ фонем:

$$p(e) \approx e^{-E(R)\tau}, \quad (5.30)$$

где R — количество информации на фонему; $E(R)$ — функции надежности: $E(R) = \min_{0 \leq x \leq \infty} [G(x) - xR]$. Здесь $G(x)$ — функция Галлагера, которая определяется соотношением

$$G(x) = -\ln \sum_{j=1}^I \left(\sum_{k=1}^I p_k p_{kj}^{\frac{1}{1+x}} \right)^{1+x},$$

где I — число фонем; p_k — вероятность появления k -й фонемы; $\|p_{kj}\|$ — матрица переходов фонем.

Графический способ вычисления $E(R)$ состоит в нахождении точки пересечения оси ординат касательной к функции $G(x) - xR$, проведенной из точки, в которой $R = dG/dx$. Матрицы переходов фонем $\|p_{kj}\|$ и матрицы переходов артикуляционных признаков были взяты из экспериментов по восприятию бессмысленных слогов типа гласный — согласный — гласный при разных отношениях сигнал/шум. Матрицы переходов гласных звуков принимались единичными. Для оценки вероятности ошибки $p(e)$ необходимо знать скорость передачи информации в речевом коде R_∞ и глубину кодовых связей. Скорость информации R_∞ принималась лежащей между 0,97 и 0,58 (в натах), а глубина кодовых связей — равной длине слова n (при $n \leq 7$). Поэтому вместо (5.30) для расчетов использовалось

$$p(e) = \sum_{\tau=1}^n p_\tau e^{-E(R)\tau},$$

где p_τ — вероятность появления слова длиной τ , вычисленная по словарю (суммарная вероятность появления слов длиной $n \leq 7$ в словаре близка к 0,74).

На рис. 5.28 показаны нижние границы функции надежности, вычисленные непосредственно по фонемам (как бы с полным учетом зависимости артикуляционных признаков) и по независимым артикуляционным признакам. По этим же графикам можно определить и верхнюю границу ошибок, называемую границей существования: $E(R) = \inf_{0 \leq x \leq 1} [G(x) - xR]$. Эта граница гаранти-

рует существование способа декодирования, обеспечивающего данную надежность. Верхние и нижние оценки вероятности ошибок на словесном уровне для всех длин слов ($1 \leq n \leq 7$) и без одно- и двухфонемных слов ($3 \leq n \leq 7$) приведены в табл. 5.4. Как видно, нижняя граница вероятности ошибок при декодировании

по фонемам столь мала, что выходит за пределы точности исходных экспериментальных данных. На рис. 5.29 приведены расчетные верхние границы вероятностей ошибок в сравнении с субъективными вероятностями ошибок слов, полученными для тех же отношений сигнал/шум при аудиторных испытаниях. Из сравнения следует, что субъективные ошибки находятся ближе к верхней границе ошибок, рассчитанных по независимым артикуляционным признакам, причем разница между вероятностями ошибок при разных способах декодирования фонем и субъективными вероятностями уменьшается по мере снижения уровня шума. Из полученных результатов следует, что при декодировании фонем может быть отдано предпочтение более простому анализу по независимым артикуляционным признакам, в особенности при умеренном уровне шума.

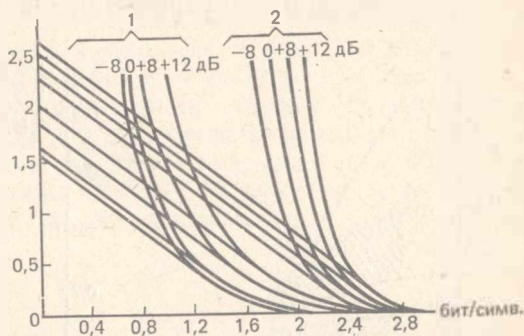


Рис. 5.28. Функция Галлагера и нижние границы функции надежности:

1 — артикуляционные признаки; 2 — фонемы

Учитывая малую исправляющую способность и заметное различие в их помехоустойчивости (рис. 5.30), можно предположить, что основная коррекция ошибок в речевом коде происходит на вышележащих уровнях.

Исправляющая способность слогов не исследована, но известны результаты аудиторных испытаний, связывающие фонемную слоговую и словесную разборчивость. Слоговая разборчивость

Таблица 5.4

Символ	Скорость передачи, БР	Границы оценок вероятности ошибок при декодировании слов и различных уровнях шумов, дБ							
		Верхняя				Нижняя			
		-8	0	+8	+12	-8	0	+8	+12
		1 ≤ n ≤ 7							
Фонема	0,97	0,034	0,029	0,025	0,023	—	—	—	—
	0,58	0,021	0,019	0,016	0,014	—	—	—	—
Признак	0,97	0,113	0,106	0,066	0,051	0,085	0,072	0,034	0,011
	0,58	0,063	0,054	0,039	0,029	0,003	0,002	0,001	0,0001
		3 ≤ n ≤ 7							
Фонема	0,97	3,4 · 10 ⁻³	2,3 · 10 ⁻³	1,5 · 10 ⁻³	1,2 · 10 ⁻³	—	—	—	—
	0,58	1,1 · 10 ⁻³	7,4 · 10 ⁻⁴	5,3 · 10 ⁻⁴	3,8 · 10 ⁻⁴	—	—	—	—
Признак	0,97	4,4 · 10 ⁻²	3,6 · 10 ⁻²	1,5 · 10 ⁻²	8,4 · 10 ⁻³	2,5 · 10 ⁻²	1,8 · 10 ⁻²	3,4 · 10 ⁻³	2 · 10 ⁻⁴
	0,58	1,3 · 10 ⁻²	9,9 · 10 ⁻³	4,7 · 10 ⁻³	2,5 · 10 ⁻³	5,3 · 10 ⁻⁶	4,3 · 10 ⁻⁶	3,4 · 10 ⁻⁶	2,9 · 10 ⁻⁶

всегда хуже фонемной и словесной. Даже если исправляющая способность слогов окажется ничтожно малой, их участок в кодовой структуре будет оправдан возможностью обнаружения физически неосуществимых последовательностей состояний. В каскадных кодах также иногда прибегают к введению каскада, обладающего лишь способностью к обнаружению ошибок.

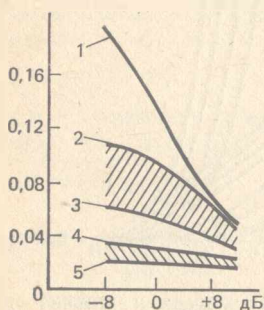


Рис. 5.29. Вероятность неправильного распознавания слов:
1 — субъективная; 2, 3 — декодирование фонем по независимым артикуляционным признакам; 4, 5 — декодирование слов по фонемам (2, 4 — $R_{\infty}=0,97$; 3, 5 — $R_{\infty}=0,58$)

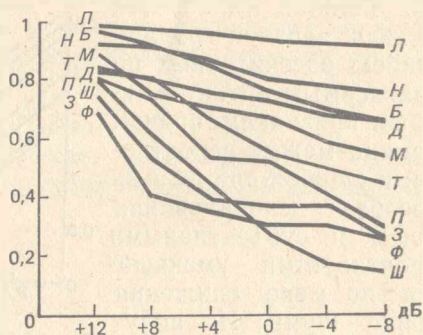


Рис. 5.30. Субъективная вероятность правильного приема фонем при маскировке белым шумом

Фонетическая структура слов отличается значительно большей исправляющей способностью, чем структура фонем и зависит от длины слов. Очевидно, что исправляющая способность однофонемных слов (предлогов и союзов) равна просто исправляющей способности фонем. Но средняя длина слов в русском языке близка к шести фонемам, и это позволяет обнаруживать и исправлять значительную долю ошибок даже в тех случаях, когда для описания фонем используются не все артикуляционные признаки и некоторые фонемы становятся неразличимыми. Оценка спектров кодовых расстояний слов при различных способах вычеркивания артикуляционных признаков (объединения фонем в группы) производилась по словам частотного словаря, записанным в фонетической транскрипции. В этом словаре кодовое расстояние между каждой парой слов определялось как минимальное расстояние по Хеммингу при всевозможных относительных сдвигах этих слов.

На рис. 5.31 условно в виде кривых показаны спектры кодовых расстояний для полных кодов фонем, а также для различных вариантов усеченных кодов. Один эксперимент был проведен при различении лишь трех мест артикуляции вместо пяти и четырех типов гласных (А, Э; И, Е, Ы; О; У); исключены были также признаки автоколебательности и бокового прохода. В этом случае неразличимы, например, следующие звуки: (Б, В), (Д, Р, Л), (Ж, З, И), (С, Ш) и т. д. Другой эксперимент проводился при тех же

условиях относительно согласных, но все гласные считались неразличимыми. Наконец, был полностью исключен признак места артикуляции (при сохранении различения согласных от гласных, а также признаков звонкости, назальности и шумности). Из рис. 5.31 видно, что при использовании полных кодов фонем верхняя грань корректирующей способности слов составляет 99,9% обнаружения и 99,3% исправления одиночных ошибок, 96,3% обнаружения и 88,8% исправления двойных ошибок, 74,9% обнаружения и 57% исправления тройных ошибок и т. д., тогда как исключение признаков места артикуляции, автоколебательности и боковых проходов и полная неразличимость гласных понижают эти величины соответственно до 99% обнаружения и 96,4% исправления одиночных ошибок, 90,1% обнаружения и 80,4% исправления двойных ошибок, 66,7% обнаружения и 50,5% исправления тройных ошибок.

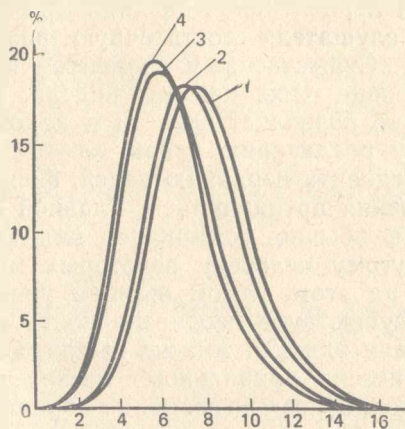


Рис. 5.31. Спектр кодовых расстояний слов:

1 — полные коды фонем; 2 — три места артикуляции, четыре типа гласных, слияние *P* и *L* с переднеязычными; 3 — то же, что и 2, но все гласные неразличимы; 4 — только признаки звонкости, назальности, шумности, «гласный — согласный»

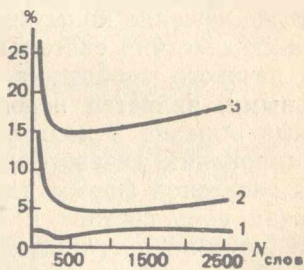


Рис. 5.32. Доля слов с одинаковыми кодами в зависимости от объема словаря:

1 — соответствует 2; 2 — 3; 3 — 4 на рис. 5.31

Вместе с тем для усеченных кодов фонем некоторые слова совпадают. Число слов, обладающих одинаковым кодом, в каждой группе невелико, а их общее число достигает 2,2% для первого способа усечения кодов, 5,6% для второго способа и 17,9% для третьего. На рис. 5.32 показана зависимость числа одинаковых слов от объема словаря. Из этих расчетов следует, что при отсутствии помех подавляющее большинство слов может быть однозначно классифицировано лишь по тем артикуляционным признакам (звонкость, шумность, назальность, признак «гласный — согласный»), которые выделяются непосредственно из акустического сигнала с помощью алгоритмов. По мере роста интенсивности помех

роль признаков места артикуляции должна возрастать из-за необходимости обеспечения большего кодового расстояния между словами.

Кодовые свойства речевого потока действительно позволяют восстановить значительную часть утерянной информации. Так, в экспериментах по распознаванию текстов, состоящих из фраз без пробелов между словами, методом последовательного декодирования было исправлено более 90% искаженных фонем и слов (число искаженных фонем доходило до 23%) [5.15]. Некоторые данные говорят о том, что в действительности процесс декодирования речевого потока происходит не только последовательно во времени, но и параллельно для каждого кодового уровня — признаков, слогов, фонем и слов.

Представление о кодовой избыточности речи позволяет заметить пока не найденные (и вряд ли существующие) инварианты фонетических элементов понятием перцептивной эквивалентности, обеспечивающим с точки зрения слушателя достаточную различимость речевых сигналов [5.106]. Существенным элементом перцептивной эквивалентности на уровне слов, как мы видели, является кодовое расстояние, которое позволяет отнести к какому-либо эталону ближайшую к нему реализацию слова.

В заключение попытаемся определить иерархию целей, формирующих действия системы управления артикуляцией. Главной целью речевого сообщения, которая обычно доминирует над остальными, является передача другому человеку некоторых мысленных образов и понятий. Уже на этом, самом высшем уровне формирования речевого потока, существует неоднозначность выбора словесной формы для передачи одной и той же мысли. Следующая цель состоит в грамматически правильном построении фраз, которая достигается упорядочиванием во времени выбранных слов. Затем следует цель разборчивого произнесения слов, т. е. формирование траектории вектора акустических параметров речевого сигнала с определенными требованиями к его значениям в заданные интервалы времени. Эта цель сохраняется и в том случае, когда речевое сообщение содержит заведомо ложные сведения. Надлежащая разборчивость фраз в условиях помех достигается выбором общеупотребительных и хорошо распознаваемых слов.

Кодовая избыточность речевого потока расширяет область допустимых положений акустических векторов и позволяет не только различным дикторам для одного и того же звуко сочетания формировать несколько различающиеся речевые сигналы, но и одному и тому же диктору несколько изменять артикуляцию от произнесения к произнесению. Отсюда также следует возможность различной организации процессов речеобразования для осмысленных и бессмысленных звуко сочетаний, как это показывают эксперименты по регистрации активности жевательной мышцы и оценки точности артикуляции. Критерий фразовой или словесной разборчивости может также служить для разрешения конфликт-

ных ситуаций, когда текущие цели артикуляции противоречат друг другу.

На последнем этапе программирования артикуляции вступают в действие критерии энергетических затрат и, возможно, алгоритмической сложности реализации программы.

СПИСОК ЛИТЕРАТУРЫ

- B.1. Жинкин Н. И. Механизмы речи. — М.: АПН, 1958. — 370 с.
- B.2. Рамишвили Г. С. Автоматическое опознавание говорящего по голосу. — М.: Радио и связь, 1981. — 224 с.
- B.3. Фант Г. Акустическая теория речеобразования. — М.: Наука, 1964. — 284 с.
- B.4. Фланеган Дж. Анализ, синтез и восприятие речи. — М.: Связь, 1968. — 392 с.
- B.5. Chiba T., Kajiyama M. The vowel — its nature and structure. — Tokyo, 1941.
- B.6. Dallas D. B. Voice programming. A fast, easy way to talk to computers. — Manuf. Eng., 1977, v. 78, N 5, p. 42—43.
- B.7. Davis S. B. Computer evaluation of laryngeal pathology, based on inverse filtering of speech. — SCRL Monograph, 1976, N 13.
- B.8. Dudley H., Tarnozы T. H. The speaking machine of Wolfgang von Kempelen. — JASA, 1950, v. 22, p. 151—166.
- B.9. Egan L. D. Interactive voice systems. — Proc. Soc. Inform. Display, 1978, v. 19, N 4, p. 187—189.
- B.10. Wiring telephone apparatus from computer generated speech/J. L. Flanagan, L. R. Rabiner, R. W. Schafer, J. P. Denman — BSTJ, 1972, v. 51, N 2, p. 391—397.
- B.11. Hirano M., Matsushita H., Kakita Y. Acoustical analysis for voice disorders. — Conf. Rec., IEEE Int. Conf. ASSP, 1976, p. 613—616.
- B.12. Lind J. Newborn infant cry. — Stockholm: Acta Paediat, 1965, Suppl. — 163 p.
- B.13. Auditory identification of pain cry signals of young infants in pathological conditions and its sound spectrographic basis/T. J. Partanen, O. Wasc-Hockert, V. Vuorenkovsky, et al. — Annals Chiz. Gynaec. Fenn., 1967, N 3, p. 330—335.
- B.14. Potter R., Copp G., Green H. Visible speech. — N. Y.: Van Nostrand, 1947.
- B.15. Spectrographic analysis of pain cry in hyperbillirubinemia/O. Wasc-Hocker, M. Koivisto, V. Vuorenkovsky, et all. — Biol. Neonate, 1971, v. 17, p. 260—271.
- 1.1. Боголюбов Н. И., Митропольский Ю. А. Асимптотические методы в теории нелинейных колебаний. — М.: Физматгиз, 1958. — 408 с.
- 1.2. Гурфинкель В. С., Левик Ю. С. Последовательные фазы длительной работы скелетной мышцы человека в режиме зубчатого тетануса. — Биофизика 1979, т. 23, № 4, с. 758.
- 1.3. Дещеревский В. И. Математические модели мышечного сокращения. — М.: Наука, 1977. — 160 с.
- 1.4. Кошляков Н. С., Глинер Э. Б., Смирнов М. М. Уравнения в частных производных математической физики. — М.: Высшая школа, 1970. — 710 с.
- 1.5. Ландау Л. Д., Лифшиц Е. М. Теория упругости. — М.: Наука, 1965. — 202 с.
- 1.6. Пальцев Е. И. Теоретическое изучение некоторых особенностей динамического поведения скелетной мышцы как одномерной вязкоупругой среды: Механика полимеров. — Рига: Зинатне, 1975, с. 618—625.
- 1.7. Пальцев Е. И., Аггашян Р. В. Система стабилизации положения кисти человека при слабых возмущениях: роль вязкоупругости мышц предплечья. — Механика композитных материалов, 1976, № 6, с. 1087—1093.

- 1.8. Самуль В. И. Основы теории упругости и пластичности. — М.: Высшая школа, 1970. — 288 с.
- 1.9. Скучик Е. Простые и сложные колебательные системы. — М.: Мир, 1971. — 553 с.
- 1.10. Сорокин В. Н. Влияние скорости артикуляции на нейромоторные процессы лицевых мышц. — Физиология человека, 1981, т. 7, № 1, с. 40—45.
- 1.11. Сорокин В. Н., Гей Т., Эван У. Идентификация системы управления движениями нижней челюсти. — Физиология человека, 1981, т. 7, № 4, с. 743—751.
- 1.12. Тимошенко С. П., Гудьер Дж. Теория упругости. — М.: Наука, 1975. — 575 с.
- 1.13. Фельдман А. Г. Центральные и рефлекторные механизмы управления движениями. — М.: Наука, 1979. — 181 с.
- 1.14. Фын Ян Чен. Математическое представление реологических свойств сердечной мышцы. — В кн.: Гидродинамика кровообращения. — М.: Мир, 1971, с. 132—161.
- 1.15. Хилл А. Механика мышечного сокращения. — М.: Мир, 1972. — 183 с.
- 1.16. Abbs J. H., Eilenberg G. R. Peripheral mechanisms of speech motor control. — In Contemporary issues in experimental phonetics. — N. Y.: Academic Press, 1976, p. 139—168.
- 1.17. Berg van Denj. Microphonic effect of the larynx. — Nature, 1957, v. 179, p. 625—626.
- 1.18. Buchtal F., Kaiser E. Factor determining tension development in skeletal muscle. — Acta Physiol. Scand., 1944, N 8, p. 38—74.
- 1.19. Fujisaki H. Dynamic characteristics of voice fundamental frequency in speech and singing. — STL-QPSR 1/1981, p. 1—20.
- 1.20. Hanson J., Huxley H. E. The structural basis of contraction in striated muscle. — Sympos. Soc. Experiment. Biol., Cambridge Univ. Press, 1955, N 9, p. 228—264.
- 1.21. Hill A. V. The heat of shortenig and the dynamic constants of muscle. — Proc. Roy. Soc. — London, 1938, B126, p. 136—195.
- 1.22. Hirano M. Phonosurgery. Basic and clinical investigations. — Otologia, Fukuoka, 1975, Suppl. 1, v. 21.
- 1.23. Husson R. Etude des phenomenes physiologiques et acoustiques fondamentaux de la voix chantee. — Paris: These, 1950.
- 1.24. Huxley H. E. Muscle structure and theories of contraction. — Progr. Biophys. and Biophys. Chem., 1957, N 7, p. 255—318.
- 1.25. Huxley H. E. Electron microscope studies on the structure of natural and synthetic protein filaments from striated muscle. — J. Mat. Biol., 1963, v. 7, N 3, p. 281—308.
- 1.26. Kakita Y. Physical properties of the vocal fold tissue: measerements on excited larynx. — Vocal fold physiology conference, Kurume, 1980, p. 6b-1 — 6b-10.
- 1.27. Response characteristics of the velar movement to the activity of the levator palatini muscle/S. Kiritani, H. Hirose, V. Sorokin, et all. — Ann. Bull. RILP, Univ. of Tokyo, 1980, N 14, p. 79—84.
- 1.28. Levin A., Wyman J. The viscous elastic properties of muscle. — Proc. Roy. Soc. London, 1927, B101, N 709, p. 218—243.
- 1.29. Significance of perioral biomechanics to lip movements during speech/E. Muller, J. Abbs, J. Kennedy, C. Larson — Chicago; Paper presented to the ASHA, 1977.
- 1.30. Oestreicher H. Field and impedance of an oscillatory sphere in a viscoelastic medium with an application to byophisics. — JASA, 1951, v. 23, N 6, p. 707—714.
- 1.31. Sandow A. A theory of active state mechanisms in isometric muscular contraction. — Science, 1958, N 127, p. 760—762.
- 2.1. Александр П. Биомеханика. — М.: Мир, 1970. — 339 с.
- 2.2. Камке Э. Справочник по обыкновенным дифференциальным уравнениям. — М.: Наука, 1971. — 576 с.
- 2.3. Константинов Б. П. Гидродинамическое звукообразование и распространение звука в ограниченной среде. — Л.: Наука, 1974. — 142 с.

- 2.4. Лойцянский Л. Г. Механика жидкости и газа. — М.: Наука, 1970. — 904 с.
- 2.5. Ржевкин С. Н. Курс лекций по теории звука. — М.: МГУ, 1960. — 335 с.
- 2.6. Сапожков М. А. Речевой сигнал в кибернетике и связи. — М.: Связьиздат, 1963. — 450 с.
- 2.7. Сорокин В. Н. Некоторые вопросы построения общей модели речеобразования. — Труды АРГО-8. — Львов, 1974, с. 97—100.
- 2.8. Сорокин В. Н. Голосовой источник как система с распределенными параметрами. — Акустический журнал, 1981, т. 27, вып. 3, с. 434—440.
- 2.9. Стретт Дж. Теория звука. Т. 1. — М.: ГИТТЛ, 1955. — 503 с.
- 2.10. Тимошенко С. П. Колебания в инженерном деле. — М.: ГИФМЛ, 1959. — 436 с.
- 2.11. Человек. Медико-биологические данные. — М.: Медицина, 1977.
- 2.12. Ямке Е., Эмде Ф., Лёш Ф. Специальные функции. — М.: Наука, 1968. — 342 с.
- 2.13. Baer T. Investigation of phonation using excised larynx. — Ph. D. Thesis, MIT, Cambridge, 1975.
- 2.14. Baer T. Investigation of the phonatory mechanisms. — ASHA Rep., 1981, № 11, p. 38—47.
- 2.15. Behringer S. Die anordnung der muskulatur in der menschlichen stimmlippe und im gebiet des conus elasticus. — Z. Anant. Entwicklung, 1955, v. 118, p. 324—342.
- 2.16. Bell-Berty F. Control of pharyngeal cavity size for English voiced and voiceless stops. — JASA, 1975, v. 57, p. 456—461.
- 2.17. Berg van Den J., Moll J. Zur anatomie des menschlichen musculus vocalis. — Z. Anant. Entwicklung, 1955, v. 118, p. 465—470.
- 2.18. Berg van Den J. On the role of the laryngeal ventricle in voice production — Folia Phoniatica, 1955, v. 7, № 2, p. 53—60.
- 2.19. Berg van Den J. Direct and indirect determination of the mean subglottic pressure. — Folia Phoniatica, 1956, v. 8, № 1, p. 1—24.
- 2.20. Berg van Den J. Subglottic pressures and vibrations of the vocal folds. — Folia Phoniatica, 1957, v. 9, № 2, p. 65—71.
- 2.21. Berg van Den J., Zantema J. T., Doornenbal J. On the air resistance and the Bernoulli effect of the human larynx. — JASA, 1957, v. 27, № 5, p. 626—631.
- 2.22. Berg van Den J. Myoelastic-aerodynamic theory of voice production. — J. Speech Hear. Res., 1958, № 1, p. 227—244.
- 2.23. Berg van Den J., Moolenaar-Bijl A. J., Damste P. H. Oesophageal speech. — Folia Phoniatica, 1958, v. 10, № 2, p. 66—83.
- 2.24. Booth J. R., Childers D. G. Automated analysis of ultra-high speed laryngeal films. — IEEE Trans. on Biomedical Engineering, 1979, v. BME-26, № 4, p. 185—192.
- 2.25. Broad D. J. Some physiological parameters for prosodic description. — SCRL Monograph, 1968, № 3.
- 2.26. Cavagna G. A., Margaria R. An analysis of the mechanics of phonation. — J. Physiology, 1965, v. 20, p. 301—307.
- 2.27. Campbell E. The respiratory muscles. — Ann. of the New York Academy of Sciences, 1968, v. 155, p. 135—140.
- 2.28. Davis S. B. Computer evaluation of laryngeal pathology based on inverse filtering of speech. — SCRL Monograph, 1976, № 13.
- 2.29. Dudgeon D. E. Two-mass model of the vocal cords. — JASA, 1970, v. 48, p. 118A.
- 2.30. Erickson D., Atkinson J. E. The function of the strap muscles in speech. — SRSR, Haskins Lab., 1976, 45/46, p. 205—210.
- 2.31. Farnsworth D. W. High-speed motion pictures of the human vocal cords. — Bell Lab. Rec., 1940, v. 18, № 7, p. 203—208.
- 2.32. Ferry J. D. Viscoelastic properties of polymers. — N. Y.: J. Wiley & Sons, 1960.
- 2.33. Flanagan J. L., Landgraf L. L. Self oscillating source for vocal tract synthesizers. — IEEE Trans. Audio and Electroac., v. AU-16, part 1, 1968, p. 57—64.
- 2.34. Fujimura O., Sawashima M. Consonant sequences and laryngeal control. — Ann. Bull. RILP, 1971, № 5, p. 1—6.

- 2.35. Fujimura O. Stereo fiberscope. — Seminar on research of the dynamic aspects of speech production, Tokyo Univ. Press, 1976.
- 2.36. Guérin B., Boe L. J. Etude de l'influence du couplage acoustique source-conduit vocal sur F_0 des voyeles orales. — *Phonetica*, 1980, № 37, p. 169—192.
- 2.37. Hamlet S. L. Ultrasonic measurements of larynx height and vocal fold vibratory pattern. — *JASA*, 1980, v. 68, № 1, p. 121—126.
- 2.38. Hirano M., Ohala J., Venard W. The function of laryngeal muscles in regulating fundamental frequency and intensity of phonation. — *JSHR*, 1969, № 12, p. 616—628.
- 2.39. Hirano M., Venard W., Ohala J. Regulation of register, pitch and intensity of voice. — *Folia Phoniatica*, 1970, v. 22, p. 1—20.
- 2.40. Hirano M. Structure and vibratory behavior of the vocal folds. — Seminar on research of dynamic aspects of speech production, Tokyo: Univ. Press, 1976, p. 13—27.
- 2.41. Hirano M. Data from high-speed motion picture studies. — Vocal fold physiology conference, Kurume, 1980, p. 4a-1 — 4a-6.
- 2.42. Hirose H. The posterior cricoarytenoid as a speech muscle. — *Ann. Bull. RILP*, 1975, № 9, p. 47—66.
- 2.43. Hiroto J. Introductory remarks. — Vocal fold physiology conference, Kurume, 1980, p. IR2-1—IR2-6.
- 2.44. Hollien H. Vocal pitch variation related to changes in vocal folds length. — *JSR*, 1960, v. 3, p. 150—156.
- 2.45. Hollien H., Moore P. Measurements of the vocal folds during changes in pitch. — *JSHR*, 1960, v. 3, p. 157—165.
- 2.46. Hollien H. Vocal fold thickness and fundamental frequency of phonation. — *JSHR*, 1962, v. 5, p. 237—243.
- 2.47. Hollien H., Curtis J. F. Elevation and tilting of vocal folds as a function of vocal pitch. — *Folia Phoniatica*, 1962, v. 15, p. 23—36.
- 2.48. Holmes J. N. Speech synthesis. — London: Mill and Boon, 1972.
- 2.49. Holmes J. N. Formant excitation before and after glottal closure. — Philadelphia: IEEE Int. Conf. on ASSP, 1976, p. 39—42.
- 2.50. Holmes J. N. Synthesis of natural sounding speech using a formant synthesizer. — *Frontiers of speech communication research*. — London: Academic Press, 1979, p. 275—286.
- 2.51. Ishiki N. Regulatory mechanism of the pitch and volume of voice. — Kyoto: Oto-Rino-Laryngology Clinic, 1959, v. 52, p. 1065—1094.
- 2.52. Ishiki N. Regulatory mechanism of voice intensity variation. — *JSHR*, 1964, № 7, p. 17—29.
- 2.53. Ishiki N., Ringel R. Airflow during the production of selected consonants. — *JSHR*, 1967, v. 7, p. 233—244.
- 2.54. Ishizaka K., Kaneko T. On equivalent mechanical constants of the vocal cords. — *JASA*, 1968, v. 24, № 5, p. 312, 313.
- 2.55. Ishizaka K., Matsudaira M. Fluid mechanical considerations of vocal cords vibration. — *SCRL Monograph*, 1972, № 8.
- 2.56. Ishizaka K., Flanagan J. L. Synthesis of voiced sounds from a two-mass model of the vocal cords. — *BSTJ*, 1972, v. 51, № 6, p. 1233—1268.
- 2.57. Ishizaka K., Flanagan J. L. Acoustic properties of longitudinal displacement in vocal cord vibration. — *BSTJ*, 1977, v. 56, № 6, p. 889—918.
- 2.58. Kakita Y., Hiki S. A study of laryngeal control for pitch change by use of anatomical structure model. — *IEEE Int. Conf. on ASSP*, 1976, p. 43—48.
- 2.59. Kakita Y., Hiki S. Investigation of laryngeal control by use of thyrometer. — *JASA*, 1976, v. 59, № 3, p. 669—674.
- 2.60. Mechanical properties of the vocal folds: measurements in vivo/T. Kaneko, K. Uchida, H. Suzuki, et al. — Vocal fold physiology conference, Kurume, 1980, p. 6a-1 — 6a-10.
- 2.61. Kent R. D., Moll K. L. Vocal tract characteristics of the stop cognants. — *JASA*, 1969, v. 49, p. 1549—1555.
- 2.62. Koike Y., Hirano M. Glottal-area time function and subglottal pressure variation. — *JASA*, 1973, v.54, № 6, p. 1618—1627.
- 2.63. Koike Y. Sub- and supraglottal pressure variation during phonation. — Vocal fold physiology conference, Kurume, 1980, p. 3d-1 — 3d-10.

- 2.64. Ladefoged P. Some physiological parameters in speech. — *Language and Speech*, 1963, № 6, p. 109—119.
- 2.65. Ladefoged P. Three areas in experimental phonetics. — London: Oxford Univ. Press, 1967.
- 2.66. Ladefoged P. Preliminaries to linguistic phonetics. — Chicago: The Univ. of Chicago Press, 1971.
- 2.67. Liberman P., Knudson R., Mead J. Determination of the rate of change of fundamental frequency with respect to subglottal air pressure during sustained phonation. — *JASA*, 1969, v. 45, p. 1537—1543.
- 2.68. Lindqvist J., Sundberg J. Pharyngeal constrictions. — *STL-QPSR*, 1971, № 4, p. 26—31.
- 2.69. Lindqvist J. A descriptive model of laryngeal articulation in speech. — *STL-QPSR*, 1972, № 2—3, p. 1—27.
- 2.70. Transillumination of the larynx in running speech/L. Lisker, A. Abramson, F. S., Cooper, M. Schvey — *JASA*, 1969, v. 45, p. 1544—1546.
- 2.71. Löfqvist A., Baer T., Joshioka H. Scaling of glottal opening. — *Phonetica*, 1981, v. 38, № 5, 6, p. 266—276.
- 2.72. Martony J. On the correction of the voice pitch level for severely hard of hearing subjects. — *Am. Ann. of the Deaf*, 1968, v. 113, p. 195—202.
- 2.73. Martenson A. Some functional and histochemical properties of the intrinsic laryngeal muscles. — *Speech Motor Control*, Pergamon Press, 1982, v. 36, p. 119—127.
- 2.74. Mead J., Bouhygs A., Procter D. F. Mechanisms generating subglottic pressure. — *Sound production in man. Ann. of the New York Academy of Science*, 1968, v. 155, p. 177—181.
- 2.75. Nadal-Saris M., Childers D. G. Comparison of natural speech with glottal area form synthetic speech. — *IEEE Int. Conf. on ASSP*, 1977, p. 37—40.
- 2.76. Oates J. M., Kirkby R. J. Acoustic investigations of abnormal voice quality. — *Aust. J. of Human Comm. Dis.*, 1979, v. 7, № 1, p. 4—15.
- 2.77. Ohala J., Ewan W. G. Speed of pitch change. — *JASA*, 1973, v. 59, p. 345(A).
- 2.78. Ohala J. Aspects of the control and production of speech. — *UCLA Working Papers in Phonetics*, 1975, № 15, p.
- 2.79. Perkell J. A. Physiology of speech production: Results and implications of a quantitative cineradiographic study. — Cambridge: MIT Press, 1969.
- 2.80. Peterson G., Barney H. Control methods used in a study of the vowels. — *JASA*, 1952, v. 24, p. 175—184.
- 2.81. Rothenberg M. The breath-stream dynamics of simple-released-plosive production. — *Biblioteke Phonetica*, 1968, № 6, p. 1—117.
- 2.82. Rothenberg M. Glottal noise during speech. — *STL-QPSR*, 1974, № 2, p. 1—10.
- 2.83. Rozsypal A. T., Millar B. F. Perception of jitter and shimmer in synthetic vowels. — *J. of Phonetics*, 1979, № 7, p. 343—355.
- 2.84. X-ray stroboscopy/S. Saito, H. Fukuda, Y. Isodai, H. Ono. — *Vocal fold physiology conference*, Kurume, 1980, p. 4b-1 — 4b-9.
- 2.85. Sawashima M. Movements of the larynx in articulation of Japanese consonants. — *Ann. Bull. RILP*, 1968, № 2, p. 11—20.
- 2.86. Sawashima M., Hirose H. New laryngoscopic technique by the use of fiber optics. — *JASA*, 1968, v. 43, p. 168, 169.
- 2.87. Sawashima M., Niimi S. Laryngeal conditions in articulations of Japanese voiceless consonants. — *Ann. Bul. RILP*, 1974, № 8, p. 13—18.
- 2.88. Sawashima M., Hirose H., Yoshioka H. Abductor (PCA) and adductor (INT) muscles of the larynx in voiceless sounds production. — *Ann. Bull. RILP*, 1978, № 12, p. 53—60.
- 2.89. Sawashima M. Laryngeal control for voicing distinctions: a review of recent work. — *Ann. Bul. RILP*, 1979, № 13, p. 23—26.
- 2.90. Sears T. A., Newsom D. The control of respiratory muscles during voluntary breathing. — *Ann. of the New York Academy of Science*, 1968, v. 155, p. 183—190.
- 2.91. Scherer R. C., Titze J. R., Curtis J. F. Pressure flow relationships in two model of the larynx having rectangular glottal shapes. — *JASA*, 1983, v. 73, № 2, p. 668—676.

- 2.92. Shin T. The function of the extrinsic laryngeal muscles. — Vocal fold physiology conference, Kurume, 1980, p. 3c-1—3c-8.
- 2.93. Shipp T., Haller R. M. Vertical larynx movement. — JASA, 1972, v. 52, p. 124(A).
- 2.94. Shipp T. Aspects of voice production and motor control. — Speech Motor Control, Pergamon Press, 1982, v. 36, p. 105—112.
- 2.95. Shimizu K. Experimental studies on movements of the vocal cords during phonation by high voltage radiograph motion pictures. — Studia phonologica, 1961, № 1, p. 111—116.
- 2.96. Smith S. Remarks on the physiology of the vibrations of the vocal cords. — Folia Phoniatrica, 1954, v. 6, № 3, p. 166—172.
- 2.97. Smith T. S. A phonetic study of the function of the extrinsic tongue muscles. — UCLA Working Papers in Phonetics, 1971, № 18.
- 2.98. Sonesson B. Vocal fold kinesiology. — Speech Motor Control, Pergamon Press, 1982, v. 37, p. 113—118.
- 2.99. Sonesson B. On the anatomy and vibratory patterns of the vocal folds. — Acta Oto-Laryng., 1960, Suppl. 156.
- 2.100. Stetson R. H. Motor phonetics. Amsterdam, 1951.
- 2.101. Titze J. R. The human vocal cords: a mathematical model. Part 1. — Phonetica, 1973, v. 28, p. 129—170.
- 2.102. Titze J. R. The human vocal cords: a mathematical model. Part 2. — Phonetica, 1974, v. 29, p. 1—21.
- 2.103. Titze J. R., Strong W. J. Normal modes in vocal cord tissues. — JASA, 1975, v. 57, N 3, p. 736—744.
- 2.104. Titze J. R. On the mechanisms of vocal-fold vibration. — JASA, 1976, v. 60, № 6, p. 1366—1380.
- 2.105. Titze J. R., Talkin D. T. A theoretical study of the effects of various laryngeal configurations on the acoustics of phonation. — JASA, 1979, v. 66, № 1, p. 60—74.
- 2.106. Warren D. Aerodynamics of speech production. — Contemporary issues in experimental phonetics. — New York: Academic Press, 1976, p. 105—136.
- 2.107. Wegel R. L. Theory of vibration of the larynx. — BSTJ, 1930, v. 9, № 1, p. 207—227.
- 2.108. Wustrow F. Bau und function des menschlichen musculus vocalis. — Z. Anat. Entwicklung, 1952, v. 116, p. 506—522.
- 3.1. Сорокин В. Н. Динамические модели артикуляции. — Биофизика, 1972, т. 27, № 7, с. 674—680.
- 3.2. Сорокин В. Н. Математические модели артикуляции гласных. — Биофизика, 1973, т. 28, № 6, с. 1100—1104.
- 3.3. Сорокин В. Н. Анализ системы управления речевыми движениями. — Биофизика, 1975, т. 20, № 4, с. 715—719.
- 3.4. Сорокин В. Н. К теории речеобразования. — В кн.: Речевое общение в автоматизированных системах. — М.: Наука, 1975, с. 103—127.
- 3.5. Сорокин В. Н. Механика движений языка. — В кн.: Описание и распознавание объектов в системах искусственного интеллекта. — М.: Наука, 1980, с. 42—71.
- 3.6. Сорокин В. Н., Гей Т., Эван В. Идентификация периферической системы управления движениями нижней челюсти. — Физиология человека, 1981, № 4, с. 743—751.
- 3.7. Теория автоматического регулирования. Кн. 1. — М.: Машиностроение, 1967. — 768 с.
- 3.8. Тихонов А. Н., Арсенин В. Я. Методы решения некорректных задач. — М.: Наука, 1974. — 233 с.
- 3.9. Чистович Л. А., Кожевников В. А. и др. Речь. Артикуляция и восприятие. — М.: Наука, 1965. — 240 с.
- 3.10. Abbs J. H. The influence of the gamma motor system on jaw movements during speech: a theoretical framework and some preliminary observations. — JSHR, 1973, № 16, p. 175—200.
- 3.11. Abbs J. H., Eilenberg G. R. Peripheral mechanisms of speech motor control. — Contemporary issues in experimental phonetics. — New York: Academic Press, 1976, p. 139—168.

- 3.12. Bawa P., Mannard A., Stein R. B. Effects of elastic loads on the contraction of cat muscles. — *Biol. Cybernetics*, 1976, N 22, p. 129—137.
- 3.13. Bawa P., Mannard A., Stein R. B. Prediction and experimental tests of a visco-elastic muscle model using elastic and inertial loads. — *Biol. Cybernetics*, 1976, № 22, p. 139—145.
- 3.14. Bell-Berty F. An electromyographic study of velopharyngeal function in speech. — *JSHR*, 1976, № 19, p. 225—240.
- 3.15. Bell-Berty F. Velopharyngeal function: a spatial-temporal model. — *SRSR*, 1980, № 63—64, p. 41—66.
- 3.16. Electromyographic study of the velum in French/A.-P. Benguerel, H. Hirose, M. Sawashima, T. Ushijima. — *Ann. Bul. RILP*, 1975, № 9, p. 79—90.
- 3.17. Bjork L. Velopharyngeal function in connected speech. — *Acta Radiologica*, 1961, Suppl. 202.
- 3.18. Bosman F. Control of jaw movements. Utrecht, 1972.
- 3.19. Biomechanical considerations in the surgical correction of mandibular deficiency/R. A. Finn, G. S. Throckmorton, W. H. Bell, H. L. Legan. — *J. Oral Surgery*, 1980, v. 38 p. 257—264.
- 3.20. Folkins J. W., Abbs J. H. Lip and jaw motor control during speech: Responses to resistive loading of the jaw. — *JSHR*, 1975, v. 18, № 1, p. 207.
- 3.21. Folkins J. W., Abbs J. H. Additional observations on responses to resistive loading of the jaw. — *JSHR*, 1976, v. 19, p. 820—821.
- 3.22. Fritzell B. The velopharyngeal muscles in speech: an electromyographic and cineradiographic study. — *Acta Oto-laryngologica*, 1969, Suppl. 250.
- 3.23. Fujimura O., Kakita Y. Remarks on quantitative description of the lingual articulation. — *Frontiers in Speech Communication Research*. Academic Press, 1979, p. 17—24.
- 3.24. Effect of speaking rate on labial consonant—vowel articulation/T. Gay, Y. Ushijima, H. Hirose, F. S. Cooper. — *J. of Phonetics*, 1974, v. 2, № 1, p. 47—63.
- 3.25. Hardcastle W. J. Physiology of speech production. — London: Academic Press, 1976.
- 3.26. Harshman R., Ladefoged P., Goldstein L. Factor analysis of tongue shapes. — *JASA*, 1977, v. 62, № 3, p. 693—707.
- 3.27. Hiki S., Nijada K., Oizumi J. Speech synthesis by control neurophysiological parameters. — *Int. Conf. on Speech Communication*. — N. Y.: U. S. Air Force Cambridge Res. Lab., 1972, Rep. № 131.
- 3.28. Hiki S., Imaizumi S., Observation of symmetry of tongue movement by use of dynamic palatography. — *Ann. Bul. RILP*, 1974, № 8, p. 69—74.
- 3.29. House A. S., Stevens K. N. Analog studies of the nasalization of vowels. — *J. Speech and Hearing Disorders*, 1956, № 21, p. 218—232.
- 3.30. Isshiki N., Honjow J., Morimoto M. Effects of velopharyngeal incompetence upon speech. — *Cleft Palate J.*, 1968, № 5, p. 218—232.
- 3.31. Ionemoto K., Kiritani S., Hirose H. A correlation analysis of velar movement and the EMG activity of the levator palatini muscle during speech. — *Ann. Bul. RILP*, 1978, № 12, p. 21—27.
- 3.32. Kent R. D., Carney P. J., Severeid L. R. Velar movement and timing: evaluation of a model for binary control. — *JSHR*, 1974, v. 17, № 3, p. 470—488.
- 3.33. Kiritani S., Itoh K., Fujimura O. Tongue-pellet tracking by a computer-controlled X-ray microbeam system. — *JASA*, 1975, v. 57, p. 1516—1520.
- 3.34. Kiritani S., Mijanaki K., Fujimura O. A computational model of the tongue. — *Ann. Bull. RILP*, 1976, № 10, p. 243—252.
- 3.35. Lanfer A., Condax J. D. The epiglottis as an articulator. — *UCLA Working papers in phonetics*, 1979, p. 60—83.
- 3.36. Lindblom B. E. F., Sundberg J. E. F. Acoustical consequences of lip, tongue, jaw and larynx movement. — *JASA*, 1971, v. 50, N4 (Part 2), p. 1166—1179.
- 3.37. Lubker J. F. Normal velopharyngeal function in speech. — *Clinics in plastic surgery*, 1975, v. 2, № 2, p. 249—259.
- 3.38. Maeda S. Un modele articulatoire de la langue avec des composantes lineaires. — *Grenoble 10-eme Journee d'etude sur la parole*, 1979, p. 1—9.
- 3.39. McNeilage P. F., Sholes G. N. An electromyographic study of the tongue during vowel production. — *JSHR*, 1964, № 7, p. 209—232.

- 3.40. Mijawaki K. A study on the musculature of the human tongue. — Ann. Bull. RILP, 1974, № 8, p. 23—50.
- 3.41. A preliminary report on the electromyographic study of the activity of lingual muscles/K. Mijawaki, H. Hirose, T. Ushijima, M. Sawashima. — Ann. Bul. RILP, 1975, № 9, p. 91—106.
- 3.42. Moll K. L. Velopharyngeal closure on vowels. — JSHR, 1962, № 5, p. 30—77.
- 3.43. Moll K. L., Daniloff R. G. Investigation of the timing of velar movements during speech. — JASA, 1971, v. 50, p. 678—684.
- 3.44. Ohman S. E. G. Numerical model of coarticulation. — JASA, 1967, v. 41, № 2, p. 310—320.
- 3.45. Perkell J. A physiologically-oriented model of tongue activity in speech production. — Ph. D. Thesis MIT, 1974.
- 3.46. Riper Van C., Irwin J. Voice and articulation. N. Jersey: Prentice-Hall Inc., 1950.
- 3.47. Shirai K., Honda M. An articulatory model and the estimation of articulatory parameters by nonlinear regression method. — Electr. and Commun. in Japan, 1976, v. 59-A, № 8, p. 35—43.
- 3.48. Sorokin V. N., Gay T., Ewan W. G. Some biomechanical correlates of jaw movements. — JASA, 1980, v. 68, Suppl. 1, p. S32.
- 3.49. Sussman H. M., McNeilage P. F., Hanson R. J. Labial and mandibular movement dynamics during the production of bilabial stop consonants: preliminary observations. — JSHR, 1973, v. 16, № 3, p. 397—420.
- 3.50. Warren D. Nasal emission of air and velopharyngeal function. — Cleft Palate J., 1967, № 4, p. 148—156.
- 3.51. Warren D. Aerodynamics of speech production. — Contemporary issues in experimental phonetics. — New York: Academic Press, 1976, p. 104—137.
- 3.52. Zemlin W. R. Speech and hearing science. N. J. Prentice-Hall Inc., 1968.
- 4.1. Белицкая Е. Е., Власов Е. В., Сорокин В. Н. Назальные звуки речи. — Акустический журнал, 1982, т. 28, № 4, с. 449—456.
- 4.2. Власов Е. В. Модификация метода Галеркина для расчета частотных параметров речевых сигналов. — В кн.: Проблемы построения систем понимания речи. — М.: Наука, 1980, с. 136—142.
- 4.3. Власов Е. В., Сорокин В. Н. Исследование акустических эффектов вертикальных колебаний голосовых складок. — Акустический журнал, 1983, т. 29, № 1, с. 11—18.
- 4.4. Водопьянов В. Г., Римский-Корсаков А. В. О затухании звука, вызванном местным сужением потока в трубе. — В кн.: Аэроакустика. — М.: Наука, 1980, с. 78—86.
- 4.5. Вейбель Э. Р. Морфометрия легких человека. — М.: Медицина, 1970.
- 4.6. Галунов В. И., Коваль С. Л., Тампель И. Б. Проблемы акустической теории речеобразования. — В кн.: Анализ и синтез речи в системах управления. М.: Научный совет по комплексной проблеме «Кибернетика», 1981, с. 60—74.
- 4.7. Исакович М. А. Общая акустика. — М.: Наука, 1973. — 495 с.
- 4.8. Канторович Л. В., Крылов В. И. Приближенные методы высшего анализа. — М.—Л.: Физматгиз, 1962.
- 4.9. Лэмб Г. Динамическая теория звука. — М.: ГИТТЛ, 1960. — 372 с.
- 4.10. Маделунг Э. Математический аппарат физики. — М.: Физматгиз, 1961. — 618 с.
- 4.11. Маркелл Д. Д., Грей А. Х. Линейное предсказание речи. — М.: Связь, 1980. — 307 с.
- 4.12. Маркюс Ж. Дискретизация и квантование. — М.: Энергия, 1969. — 143 с.
- 4.13. Морз Ф. Колебания и звук. — М.—Л.: ГИТТЛ, 1949. — 496 с.
- 4.14. Скучик Е. Основы акустики. Т. 1. — М.: ИЛ, 1958. — 617 с.
- 4.15. Скучик Е. Основы акустики. Т. 2. — М.: ИЛ, 1959. — 565 с.
- 4.16. Смирнов В. И. Курс высшей математики. Т. 2. — М.: ГИТТЛ, 1957. — 628 с.
- 4.17. Сорокин В. Н. Потери в речевом тракте. — Акустический журнал, 1977, т. 23, № 6, с. 939—946.

- 4.18. Сорокин В. Н. О роли подглоточной области в процессе речеобразования. — В кн.: Проблемы построения систем понимания речи. — М.: Наука, 1980, с. 125—135.
- 4.19. Сорокин В. Н. Шумы турбулизации в речеобразовании. — Киев: Труды АРСО-12, 1982, с. 432—435.
- 4.20. Фурдуев В. В. Электроакустика. — М.—Л.: ГИТТЛ, 1948. — 515 с.
- 4.21. Berg van Den J. W., Zantema J. T., Doornenbal P. On the air resistance and the Bernoulli effect of the human larynx. — JASA, 1957, v. 29, p. 626—631.
- 4.22. Blumstein S. E., Stevens K. N., Nigro G. N. Property detectors for bursts and transitions in speech perception. — JASA, 1977, v. 61, № 5, p. 1301—1313.
- 4.23. Blumstein S. E., Stevens K. N. Acoustic invariance in speech production: Evidence from measurements of the spectral characteristics of stop consonants. — JASA, 1979, v. 66, № 4, p. 1001—1017.
- 4.24. Blumstein S. E., Stevens K. N. Perceptual invariance and onset spectra for stop consonants in different vowel environments. — JASA, 1980, v. 67, № 2, p. 648—662.
- 4.25. Dadson R. S., Robinson D. W., Greig R. G. R. The mechanical impedance of human mastoid process. — British J. Appl. Phys., 1954, v. 5, № 2, p. 435—442.
- 4.26. Edwards T. J. Multiple feature analysis of intervocalic English plosives. — JASA, 1981, v. 69, № 2, p. 535—547.
- 4.27. Speech analysis and speech production. Subglottal formants/G. Fant, K. Ishizaka, J. Lindqvist, J. Sundberg. — STL QPSR, 1972, № 1, p. 1—12.
- 4.28. Fant G., Pauli S. Spatial characteristics of vocal tract resonance model. — Stockholm: Preprints of the Speech Commun. Seminar, 1972, v. 2, p. 121—132.
- 4.29. Fant G. Vocal tract wall effects, losses and resonance bandwidth. — STL QPSR, 1973, № 2—3, p. 28—52.
- 4.30. Fant G., Nord L., Branderud P. A note on the vocal tract wall impedance. — STL QPSR, 1976, № 4, p. 13—20.
- 4.31. Fant G., Wakita H. Toward a better vocal tract model. — STL QPSR, 1978, № 1, p. 9—29.
- 4.32. Flanagan J. L., Ishizaka K. Automatic generation of voiceless excitation in a vocal cord-vocal tract speech synthesizer. — IEEE on ASSP, 1976, v. ASSP-24, № 2, p. 163—170.
- 4.33. Franke E. K. The mechanical impedance of the surface of the human body. — J. Appl. Phys., 1951, v. 3, № 10, p. 582—590.
- 4.34. Franke E. K. The impedance of the human mastoid. — JASA, 1952, v. 24, № 4, p. 410—411.
- 4.35. Fujimura O., Lindqvist J. Sweep-tone measurements of vocal tract characteristics. — JASA, 1971, v. 49, № 2, p. 541—558.
- 4.36. Heinz J. Perturbation function for the determination of vocal tract area function from vocal tract eigenvalues. — STL QPSR, 1967, № 1, p. 1—14.
- 4.37. Holmes J. N. Synthesis of natural-sounding speech using a formant synthesizer. — Frontiers of Speech Commun. Res. London, Academic Press, 1979, p. 275—286.
- 4.38. Ishizaka K., Matsudaira M. Fluid mechanical consideration of vocal cord vibration. — SCRL Monograph, 1972, № 8.
- 4.39. Ishizaka K., French J. C., Flanagan J. L. Direct determination of vocal tract wall impedance. — IEEE on ASSP, 1975, v. ASSP-23, № 4, p. 370—373.
- 4.40. Kelly J. L., Lochbaum C. C. Speech synthesis. — Proc. 4 Int. Congr. Acoust., 1962, Paper G42, p. 1—4.
- 4.41. Kewly-Port D. Time-varying features as correlates of place of articulation in stop consonants. — JASA, 1983, v. 73, № 1, p. 322—335.
- 4.42. Liljencrants J., Fant G. Computer program for VT-resonance frequency. — STL QPSR, 1976, № 4, p. 15—20.
- 4.43. Ludvig G. D. The velocity of sound through tissues and acoustical impedance of tissues. — JASA, 1950, v. 22, p. 862—866.
- 4.44. Maeda S. On a simulation method by dynamically varying vocal tract: reconsideration of Kelly-Lochbaum model. — Articulatory modeling and phonetics. — Grenoble: 1977, p. 281—288.

- 4.45. Maeda S. A digital simulation method of the vocal-tract system. — *Speech Communication*, 1982, № 1, p. 199—229.
- 4.46. Mermelstein P. Determination of the vocal-tract shape from measured formant frequencies. — *JASA*, 1967, v. 41, № 5, p. 1283—1289.
- 4.47. Meyer-Eppler W. Zum Erzeugungsmechanismus der Geräuschaute. — *Z. für Phonetik*, 1953, v. 7, № 3, 4, p. 191—212.
- 4.48. Müller E. M., Brown W. S. Variations in the supraglottal air pressure waveform and their articulatory interpretation. — *Speech and Language: Advances in Basic Research and Practice*. — New York: Academic Press, 1980, p. 317—389.
- 4.49. Nguyen D. T., Guerin B. Effects of nasal coupling on the vowels. — Paper presented on the 99-th Meeting of the Acoustical Society of America, Atlanta, 1980.
- 4.50. Ohde R. N., Stevens K. N. Effect of burst amplitude on the perception of place of articulation for stops. — *Working papers, MIT Speech Communication Group*, 1982, v. 1, p. 61—70.
- 4.51. Rice L., Ohman S. On the relationship between formant bandwidth and vocal tract shape features. — *UCLA Working papers in phonetics*, 1976, p. 27—31.
- 4.52. Shirai K., Fujisawa H., Koyama Y. Modeling of the generation process of the fricative consonants. — *Trans. IEECE*, 1975—1976, v. 58-A, № 6, p. 345—352.
- 4.53. Shroeder M. R. Determination of geometry of the human vocal tract by acoustic measurements. — *JASA*, 1967, v. 41, № 4, p. 1002—1010.
- 4.54. Stevens K. N. Airflow and turbulence noise for fricative and stop consonants: static considerations. — *JASA*, 1971, v. 50, № 4, Part 2, p. 1180—1192.
- 4.55. Stevens K. N., Blumstein S. E. Invariant cues for place of articulation in stop consonant. — *JASA*, 1978, v. 64, № 5, p. 1301—1313.
- 4.56. Stevens K. N. Acoustic correlates of some phonetic categories. — *JASA*, 1980, v. 68, № 3, p. 836—842.
- 4.57. Ungeheuer G. Elemente einer akustischen theorie der vokalartikulation. — Berlin: Springer, 1962.
- 4.58. Walley A. C., Carrel T. P. Onset spectra and formant transitions in the adult's perception of place of articulation in stop consonants. — *JASA*, 1983, v. 73, № 3, p. 1011—1022.
- 4.59. Westbury J. R. Enlargement of the supraglottal cavity and its relation to stop consonant voicing. — *JASA*, 1983, v. 73, N 4, p. 1322—1336.
- 5.1. Бару А. В., Карасева Т. А. Мозг и слух. — М.: Изд. МГУ, 1971. — 105 с.
- 5.2. Бернштейн Н. А. О построении движений. — М.: Медгиз, 1947.
- 5.3. Бондарко Л. В. Звуковой строй современного русского языка. — М.: Просвещение, 1977. — 175 с.
- 5.4. Бриллюэн Л. Наука и теория информации. — М.: Физматгиз, 1960. — 392 с.
- 5.5. Бутковский А. Г. Структурная теория распределенных систем. — М.: Наука, 1977. — 320 с.
- 5.6. Винарская Е. Н., Шур С. Н. Нарушения звуковой системы речи при бульбарной дизартрии взрослых. — В кн.: Очерки по патологии речи и голоса. — М.: Просвещение, 1967, с. 128—135.
- 5.7. Винарская Е. Н. Клинические проблемы афазии. — М.: Медицина, 1971.
- 5.8. Винарская Е. Н., Пулатов А. М. Дизартрия и ее топико-диагностическое значение в клинике очаговых поражений мозга. — Ташкент: Медицина, 1973. — 142 с.
- 5.9. Винер Н. Нелинейные задачи в теории случайных процессов. — М.: ИЛ, 1961. — 158 с.
- 5.10. Гельфанд И. М., Гурфинкель В. С., Цетлин М. Л. О тактиках управления сложными системами в связи с физиологией. — В кн.: Биологические аспекты кибернетики. — М.: Изд. АН СССР, 1962, с. 61—73.
- 5.11. Гранит Р. Основы регуляции движений. — М.: Мир, 1973. — 275 с.
- 5.12. Гурфинкель В. С., Левик Ю. С. Сенсорные комплексы и сенсорная интеграция. — *Физиология человека*, 1979, № 3, с. 399—414.
- 5.13. Дейч С. Модели нервной системы. — М.: Наука, 1970. — 324 с.

- 5.14. Дёх Т. Руководство к практическому применению речеобразования Лапласа и z-преобразования. — М.: Наука, 1971. — 288 с.
- 5.15. Зигангиров К. Ш., Сорокин В. Н. Применение последовательного декодирования к распознаванию слитной речи. — Проблемы передачи информации, 1977, № 4, с. 81—88.
- 5.16. Исследование процессов управления мышечной активностью. — М.: Наука, 1970. — 204 с.
- 5.17. Книппер А. В., Махонин В. А. К описанию речевых сигналов. — В кн.: Речевое общение в автоматизированных системах. — М.: Наука, 1976, с. 46—59.
- 5.18. Корнев Г. В. Цель и приспособляемость движения. — М.: Наука, 1974. — 528 с.
- 5.19. Корнев Г. В. Введение в механику человека. — М.: Наука, 1977. — 263 с.
- 5.20. Либкинд М. С. Моделирование интерференционной биоэлектрической активности. Ч. 2. — Биофизика, 1968, т. 13, вып. 4, с. 685—693.
- 5.21. Либкинд М. С. Моделирование интерференционной биоэлектрической активности. Ч. 4. — Биофизика, 1972, т. 17, вып. 1, с. 124—130.
- 5.22. Миллер Д. А. Магическое число семь плюс или минус два: О некоторых пределах нашей способности перерабатывать информацию. — В кн.: Инженерная психология. — М.: Прогресс, 1964.
- 5.23. Пенфильд В., Робертс Л. Речь и мозговые механизмы. — Л.: Медицина, 1964. — 264 с.
- 5.24. Персон Р. С., Кудина Л. П. Ритмика мотонейронов человека при произвольном напряжении мышцы. — Нейрофизиология, 1972, т. 3, № 6, с. 609—619.
- 5.25. Персон Р. С. Двигательные единицы и мотонейронный пул. — В кн.: Физиология движений. — М.: Наука, 1976, с. 69—101.
- 5.26. Математическая теория оптимальных процессов/Л. С. Понтрягин, В. Г. Болтянский, Р. В. Гамкрелидзе, Е. Ф. Мищенко — М.: ГИФМЛ, 1961. — 388 с.
- 5.27. Северин Ф. В. Мышечные веретена и эфферентная регуляция их деятельности. — В кн.: Физиология движений. — Л.: Наука, 1976, с. 102—123.
- 5.28. Сорокин В. Н. К теории речеобразования. — В кн.: Речевое общение в автоматизированных системах. — М.: Наука, 1975, с. 103—127.
- 5.29. Сорокин В. Н. Элементы кодовой структуры речи. — В кн.: Распознавание образов (Теория и приложения). — М.: Наука, 1977, с. 42—60.
- 5.30. Теория автоматического регулирования. Кн. 3. Ч. 1. — М.: Машиностроение, 1969. — 607 с.
- 5.31. Трауготт Н. Н., Кайданова С. Н. Нарушения слуха при сенсорной алалии и афазии. — Л.: Наука, 1975.
- 5.32. Физиология сенсорных систем. Ч. 3. Физиология механорецепторов. — Л.: Наука, 1975.
- 5.33. Фицнер Л. Н. Биологические поисковые системы. — М.: Наука, 1977. — 135 с.
- 5.34. Шик М. Л. Управление наземной локомоцией млекопитающих животных. — В кн.: Физиология движений. — Л.: Наука, 1976, с. 233—275.
- 5.35. Abbs J. H. The influence of the gamma motor system on jaw movements during speech: a theoretical framework and some preliminary observations. — JSHR, 1973, № 16, p. 175—200.
- 5.36. Abbs J. H., Eilenberg G. R. Peripheral mechanisms of speech motor control. — Contemporary issues in experimental phonetics. — New York: Academic Press, 1976, p. 139—168.
- 5.37. Abbs J. H., Cole K. J. Consideration of bulbar and suprabulbar afferent influences upon speech motor coordination and programming. — Speech Motor Control. — London: Pergamon Press, 1982, p. 159—186.
- 5.38. Abbs J. H., Gracco V. L. Evidence for speech muscle functional compartmentalization: Theoretical and methodological implications. — JASA, 1982, v. 71, Suppl. 1, p. 33.
- 5.39. Adams J. A. Issues for a closed-loop theory of motor learning. — Motor Control. Issues and Trends. — London: Academic Press, 1976, p. 87—107.

- 5.40. Amerman J. D., Daniloff R. G. Articulation patterns resulting from modification of oral cavity size. — ASHA Reports, № 13, p. 559.
- 5.41. Allen G. D. Segmental timing in speech production. — J. of Phonetics, 1973, v. 1, № 3, p. 219—237.
- 5.42. Aritzumi H., Shigenaga M. Speech synthesis by articulatory movements. — Rep. at Group of Speech, Acoust. Soc. of Japan, 1979, № 9, p. 291—297.
- 5.43. Barr M. L. The human nervous system. An anatomical viewpoint. Maryland, Hasper and Row, 1974.
- 5.44. Benguerel A.-P., Cowan H. A. Coarticulation of upper lip protrusion in French. — *Phonetica*, 1974, № 30, p. 41—55.
- 5.45. Bell-Berty F., Harris K. S. Some aspects of coarticulation. — SRSR Haskins Lab., 1976, № 45, 46, p. 197—204.
- 5.46. Dichotic signs of the recognition of speech elements in normals, temporal lobectomies, and hemispherectomies/C. I., Berlin, R. J. Porter, S. S. Lowe-Bell, et al IEEE Trans. and Audio Electroacoust., 1973, AU-21, p. 189—195.
- 5.47. Borden G. J., Harris K. S., Oliver W. Oral feedback. 1. Variability of the effect of nerve block anesthesia upon speech. — J. of Phonetics, 1973, v. 1, № 4, p. 289—295.
- 5.48. Borden G. J., Harris K. S., Catena L. Oral feedback. 2. An electromyographic study of speech under nerve-block anesthesia. — J. of Phonetics, 1973, v. 1, № 4, p. 297—308.
- 5.49. Bowman J. R. The muscle spindle and neural control of the tongue. Implications for speech. — Springfield: Illinois, 1971.
- 5.50. Broman H., Lindström L. A model describing the power spectrum of myoelectric signals. Part 2: Motor unit signals. — Göteborg: Rep. of Chalmers Univ. of Technol., 1974, № 8.
- 5.51. Carlson R., Granström B., Klatt D. H. Some notes on the perception of temporal patterns in speech. — *Frontiers of Speech Communication Research*. — N. Y.: Academic Press, 1979, p. 233—244.
- 5.52. Speaking clearly: acoustic characteristics and intelligibility of stop consonants/F. R. Chen, V. W. Zue, M. A. Picheny. — Speech Commun. Group Working Papers, MIT, 1983, p. 1—8.
- 5.53. Cleveland S., Ross H. G. Dynamic properties of Renshaw cells: frequency response characteristics. — *Biol. Cybernet.*, 1977, v. 27, p. 175.
- 5.54. Cooper S. Muscle spindles in the intrinsic muscles of the human tongue. — J. of Physiology, 1953, N 22, p. 193—202.
- 5.55. Cooper W. E. Articulatory effects on speech perception: a preliminary report — J. of Phonetics, 1975, v. 3, № 2, p. 87—98.
- 5.56. Cooper W. E., Billings D. Articulatory effects on speech perception: a second report. — J. of Phonetics, 1976, v. 4, № 3, p. 219—232.
- 5.57. Cooper D. S., Folkins J. W. The temporal sampling problem in electromyographic studies of speech musculature. — JASA, 1982, v. 71, Suppl. 1, p. S33.
- 5.58. Daniloff R. G., Hammaberg R. E. On defining coarticulation. — J. of Phonetics, 1973, v. 1, N 3, p. 239—248.
- 5.59. Dixit R. D., MacNeilage P. F. Coarticulation of nasality: evidence from Hindi. — Oral paper presented at the 83-d Meeting of the Acoust. Soc. of America, 1972.
- 5.60. Dorman M. F., Freeman F. J., Borden G. J. The effect of delayed auditory feedback on phonation: an electromyographic study. — 8 Int. Congr. of Phonetic Sciences, England, 1975, p. 187—196.
- 5.61. Euler C. V. Proprioceptive control in respiration. — Muscular afferents and motor control. — New York: John Wiley, 1966.
- 5.62. Evarts E. V. Analogies between central motor programs for speech and for limbs movements. — London: Speech Motor Control, Pergamon Press, 1982, p. 19—41.
- 5.63. Fairbanks G. A theory of the speech mechanism as a servosystem. — J. of Speech and Hearing Disorders, 1954, N 19, p. 133—139.
- 5.64. Folkins J. W., Abbs J. H. Lip and jaw motor control during speech: Responses to resistive loading of the jaw. — JSHR, 1975, N 18, p. 207—220.
- 5.65. Folkins J. W., Abbs J. H. Additional observations on responses to resistive loading of the jaw. — JSHR, 1976, № 19, p. 820—831.

- 5.66. **Freeman F. J., Ushijima T.** The stuttering larynx: EMG, fiber optic study of laryngeal activity the moment of stuttering. — SRSR, Haskins Lab., 1975, N 41, p. 217—228.
- 5.67. **Freeman F. J., Ushijima T.** Laryngeal muscles in stuttering. — SRSR Haskins Lab., 1976, № 45—46, p. 211—236.
- 5.68. **Fromkin V. A.** The non-anomalous nature of anomalous utterances. — *Language*, 1971, N 47, p. 27—52.
- 5.69. **Gay T., Hirose H.** Effect of speaking rate on labial consonant production. — *Phonetica*, 1973, N 27, p. 44—56.
- 5.70. **Effect of speaking rate on labial and consonant-vowel articulation/T. Gay, T., Ushijima, H. Hirose, F. Cooper.** — *J. of Phonetics*, 1974, v. 2, № 1, p. 47—63.
- 5.71. **Gay T.** The motor control of speech rate. Honolulu: Paper presented at 96-th Meeting of Acoust. Soc. of America, 1978.
- 5.72. **Gay T., Lindblom B., Lubker J.** Production of bite-block vowels: Acoustic equivalence by selective compensation. — *JASA*, 1981, v. 69, № 3, p. 802—810.
- 5.73. **Goertler K.** Die anordnung histologie und histogenese der quergestreiften muskulatur im menschlichen stimmband. — *Z. Anat. Entro-Gesh*, 1950, № 115, p. 352—401.
- 5.74. **Granit R.** The functional role of the muscle spindles — facts and hypothesis. — *Brain*, 1975, v. 98, Part 4, p. 531—536.
- 5.75. **Hages J. R. M.** Memory span for several vocabularies as a function in a complex vocabulary size. — *Quart. Progr. Rep., Acoust. Lab. MIT*, 1952.
- 5.76. **Hardcastle W. J.** Some aspects of speech production under controlled conditions of oral anesthesia and auditory masking. — *J. of Phonetics*, 1975, № 3, p. 197—214.
- 5.77. **Hammaberg R.** On redefining coarticulation. — *J. of Phonetics*, 1982, v. 10, № 2, p. 123—137.
- 5.78. **Hamlet S. L., Stone M.** Compensatory vowel characteristics resulting from the presence of different types of experimental dental prostheses. — *J. of Phonetics*, 1976, N 4, p. 196—218.
- 5.79. **Some stress effects on electromyographic measures of consonant articulations/K. S. Harris, T. Gay, G. N. Scholes, P. Liberman.** — SRSR, Haskins Lab., 1968, N 13—14, p. 137—152.
- 5.80. **Harris K. S.** Vowel stress and articulatory reorganization. — SRSR, Haskins Lab., 1971, N 28, p. 167—178.
- 5.81. **Hatze H.** Neuromusculoskeletal control systems modeling — a critical review of recent developments. — *IEEE Trans. on Automatic Control*, 1980, v. AC—25, N 3, p. 375—385.
- 5.82. **House A. S.** Discussion paper on acoustics of speech. — *Speech and Cortical Functioning*. — New York: Academic Press, 1972, p. 167—176.
- 5.83. **Henke L. D.** Dynamic articulatory model of speech production used computer simulation. Doctoral Thesis, 1966.
- 5.84. **Huggins A. W. F.** Delayed auditory feedback and the temporal properties of the speech material. — *Zeitschrift fur Phonetick*, 1968, Band 21, Heft 1/2.
- 5.85. **Hutchinson J. M., Putnam A. H. B.** Aerodynamic aspect of sensory deprived speech. — *JASA*, 1974, N 56, p. 1612—1617.
- 5.86. **Imaizumi S., Hiki S.** Extraction of motor command from electromyography by means of a pulse coding method. — *Japan Journ. of Medical Electronics and Biological Engineering*, 1977, v. 15, N 6, p. 13—20.
- 5.87. **Kadanoff D.** Die sensiblen nervendigungen in der mimischen muskulatur des menschen. — *Z. Mikz.-anat. Forsch.*, 1956, N 62, p. 1—5.
- 5.88. **Keene M. F. L.** Muscle spindles in human laryngeal muscles. — *J. of Anatomy* 1961, N 95, p. 25—29.
- 5.89. **Keele S. T., Summers J. J.** The structure of motor programs. — *Motor Control*. — N. Y.: Academic Press, 1976, p. 109—142.
- 5.90. **Kent R. D., Moll K. L.** Tongue body articulation during vowel and diphthong gestures. — *Folia Phoniatica*, 1972, N 24, p. 286—300.
- 5.91. **Kent R. D.** Models of speech production. — *Contemporary issues in experimental phonetics*. — N. Y.: Academic Press, 1976, p. 79—104.

- 5.92. Kent R. D., Minifie F. D. Coarticulation in recent speech production models. — *J. of Phonetics*, 1977, N 5, p. 115—133.
- 5.93. Kelso S. J. A., Stelmach G. E. Central and peripheral mechanisms in motor control. — *Motor Control. Issues and Trends*. — N. Y.: Academic Press, 1976, p. 3—40.
- 5.94. Kimura D., Folb S. Neural processing of backwards speech sounds. — *Science*, 1968, p. 396.
- 5.95. Kimura D. The neural basis of language qua gesture. — *Studies in Neuro-linguistics*. V. 2. — New York: Academic Press, 1976, p. 145—156.
- 5.96. Klatt D. H. Interaction between two factors that influence vowel duration. — *JASA*, 1973, № 54, p. 1102—1104.
- 5.97. Klatt D. H. Linguistic uses of segmental duration in English: acoustic and peripheral evidence. — *JASA*, 1976, N 59, p. 1208—1221.
- 5.98. Klatt D. H. Synthesis by rule of segmental durations in English sentences. — *Frontiers of Speech Communication Research*. — New York: Academic Press, 1979, p. 287—300.
- 5.99. Landgren S., Olson K. A. Oral mechanoreceptors. — *Speech Motor Control*. — London: Pergamon Press, 1982, p. 129—139.
- 5.100. Lee B. S. Effects of delayed speech feedback. — *JASA*, 1950, № 22, p. 824.
- 5.101. Lehiste J. Suprasegmentals. — Cambridge: MIT Press, 1970.
- 5.102. Minimal rules for synthesizing speech/A. M. Liberman, F. Ingemann, L. Lisker. — *JASA*, 1959, v. 31, N 11, p. 1490—1499.
- 5.103. Liberman A. M. The grammars of speech and language. — *Cognitive Psychology*, 1970, N 1, p. 301—323.
- 5.104. Lindblom B., Lyberg B., Holmgren K. Durational patterns of Swedish phonology: do they reflect short-term motor memory processes? Rep. of Institute of Linguistics, Univ. of Stockholm, 1976.
- 5.105. Linblom B. E. F., Lubker J., Gay T. Formant frequencies of some fixed-mandible vowels and a model of speech motor programming by predictive simulation. — *J. of Phonetics*, 1979, N 7, p. 147—161.
- 5.106. Lindblom B. The interdisciplinary challenge of speech motor control. — *Speech Motor Control*. — London: Pergamon Press, 1982, p. 3—18.
- 5.107. Lindström L. A model describing the power spectrum of myoelectric signals. Part 1: Single fiber signal. Rep. of Chalmers Univ. of Technol., Göteborg: 1974, № 8.
- 5.108. Lindström L., Broman H. A model describing the power spectrum of myoelectric signals. Part 3: Summation of motor unit signals Rep. of Chalmers Univ. of Technol., Göteborg: 1974, № 9.
- 5.109. Lindström L., Kådefors R. A model describing the power spectrum of myoelectric signals. Part 4: Total power. Rep. of Chalmers Univ. of Technol., Göteborg: 1974, № 10.
- 5.110. Lubker J., Gay T. Spatial—temporal goals: maturational and cross—linguistic variables. — *Speech Motor Control*. London: Pergamon Press, 1982, p. 205—216.
- 5.111. De Luca C. J., Forrest W. J. Some properties of motor unit potential trains recorded during constant force isometric contraction in man. — *Kybernetik*, 1973, N 12, p. 160—168.
- 5.112. De Luca C. J. Towards understanding the EMG signal. — In: *Muscle alive*. Baltimore, Williams & Wilkins, 1978, p. 53—78.
- 5.113. Sensitivity of the human perioral reflex to parameters of mechanical stretch/C. R. Lurson, J. W. Folkins, M. D. McClean, E. M. Muller. — *Brain Res.*, 1978, N 146, p. 159—164.
- 5.114. McNeilage P. F. Motor control of serial ordering in speech. — *Psychological Review*, 1970, v. 77, p. 182—196.
- 5.115. McNeilage P. F. Speech physiology. — In: *Speech and Cortical Functioning*. N. Y., Academic Press, 1972, p. 1—72.
- 5.116. Mechanical properties of single motor units in speech musculature/P. F. MacNeilage, H. M. Sussman, J. R. Westbury, R. K. Powers. — *JASA*, 1979, v. 65, N 4, p. 1047—1052.
- 5.117. McNeilage P. F. Speech production. — *Proc. of the 9 Int. Congr. of Phonetic Sciences*, Copenhagen: 1979, p. 11—39.

- 5.118. McNeilage P. F. Speech production mechanisms in aphasia. — *Speech Motor Control*. — London: Pergamon Press, 1982, p. 43—60.
- 5.119. Mandelbrot B. Contribution à la théorie mathématique des jeux de communication. — *Publ. de l'Inst. de statistique de l'Univ. de Paris*, 1952, v. 2, № 1—2, p. 80—102.
- 5.120. Martin J. G. Rhythmic (hierarchicall) versus serial structure in speech and other behavior. — *Psychological Review*, 1972, v. 79, p. 487—509.
- 5.121. Matthews P. B. C. Mammalian muscle receptors and their central actions. — London: Arnold, 1972.
- 5.122. Matthews P. B. C. The advances of the last decade of animal experimentation upon muscle spindles. — *New developments in electromyography and clinical neurophysiology*, 1973, v. 3, p. 95—125.
- 5.123. Momsen R. B. The production of English stop consonants in the speech of deaf children. — *J. of Phonetics*, 1976, № 4, p. 29—41.
- 5.124. Moll K. L., Daniloff R. G. Investigation of the timing of velar movements during speech. — *JASA*, 1971, № 50, p. 678—684.
- 5.125. Moore A. D. Synthesized EMG waves and their implications. — *Am. J. Phys. Med.*, 1967, v. 46, p. 1302.
- 5.126. Mortimer J. T., Magnusson R., Petersen J. Conduction velocity in ischemic muscle: effect of EMG frequency spectrum. — *Am. J. Phys.*, 1970, v. 219, N 5, p. 1324—1329.
- 5.127. Nelson W. L., Westbury J. L. Jaw movements during increasingly rapid articulation of single syllables: Implication on speech motor control. — *JASA*, 1982, v. 71, Suppl. 1, p. S33.
- 5.128. Netsell R. A developing framework for research in speech production. — *Progress Rep. of Speech Res. Lab., Univ. of Wisconsin*, 1971, N 1, p. 10—28.
- 5.129. Ohala J. J. The temporal regulation of speech. — *Auditory analysis and perception of speech*. — New York: Academic Press, 1975, p. 431—454.
- 5.130. Oller D. K. The effect of position in utterance on speech segment duration in English. — *JASA*, 1973, v. 54, p. 1235—1247.
- 5.131. Öhman S. E. G. Coarticulation in VCV utterances: spectrographic measurements. — *JASA*, 1966, v. 39, p. 151—168.
- 5.132. Parker F. The coarticulation of vowels and stop consonants. — *J. of Phonetics*, 1974, v. 2, N 3, p. 211—221.
- 5.133. Perkell J., Nelson W. L. Articulatory targets and speech motor control: a study of vowel production. — *Speech Motor Control*. London: Pergamon Press, 1982, p. 187—204.
- 5.134. Person R. S., Libkind M. S. Simulation of electromyograms showing interference patterns. — *Electroenceph. and Neurophys.*, 1970, N 28, p. 625—632.
- 5.135. Peters R. W. The effect of changes in side-tone delay and level upon rate of oral reading of normal speakers. — *J. Speech and Hearing Disorders*, 1954, v. 19, p. 483—490.
- 5.136. Poppele R. Systems approach to the study of muscle spindles. — In: *Control of Posture and Locomotion*. — New York: Plenum Press, 1973, p. 127—146.
- 5.137. Porter B., Crosley T. R. Modal control. — London: Academic Press, 1972.
- 5.138. Ringel R. L. Oral sensation and perception: a selective review. — *Proc. of Conf. Speech and dento-facial complex*. New Orleans: 1970, p. 12—14.
- 5.139. Russel D. G. Spatial location cues and movement production. — *Motor Control*. — N. Y.: Academic Press, 1967, p. 67—85.
- 5.140. Savin H. B., Bever T. G. The nonperceptual reality of the phoneme. — *J. of Verbal Learning and Verbal Behavior*, 1970, № 9, p. 295—302.
- 5.141. Schmidt R. A. The schema as a solution to some persistent problems in motor learning theory. — *Motor Control*. — N. Y.: Academic Press, 1976, p. 41—66.
- 5.142. Sear T. A., Davis J. N. The control of respiratory muscles during voluntary breathing. — *Ann. of the New York Academy of Sciences*, 1968, v. 115, p. 183—190.
- 5.143. Shigenaga M., Ariizumi H. Articulatory movements by rule. — *Articulatory modeling and phonetics*, Grenoble: 1977, p. 193—202.

- 5.144. Shirai K., Honda M., Tsuruta R. A model of control mechanism of articulatory motion. — J. of Acoust. Soc. of Japan., 1976, v. 32, N 1, p. 24—25.
- 5.145. Shirai K., Honda M., Tsuruta R. Modeling and simulation of speech production processes. — Rep. at Group of Speech, Acoust. Soc. of Japan, 1975, N 3, p. 1—S.
- 5.146. Smith T. S. A phonetic study of the function of the extrinsic tongue muscles. — UCLA Working papers in phonetics, 1971, N 8.
- 5.147. Smith T. S., Lee C. J. Peripheral feedback mechanisms in speech production models. — 7 Int. Congr. of Phonetic Sciences. Montreal; 1971, p.
- 5.148. Stålberg E. Propagation velocity in human muscle fibers in situ. — Acta Physiologica Scandinavica, 1966, v. 70, Suppl. 287.
- 5.149. Stein R. B., Ogustörel M. N. Tremor and other oscillations in neuromuscular systems. — Biol. Cybernetics, 1976, № 22, p. 147—157.
- 5.150. Stein R. B. Nerve and Muscle. — New York: Plenum Press, 1980.
- 5.151. Stevens K. N., House A. S. Perturbations of vowel articulation by consonant context: an acoustical study. — JSHR, 1962, v. 6, p. 111—128.
- 5.152. Stevens K. The quantal nature of speech. — evidence from articulatory-acoustic data. — Human Communication: A Unified View. — New York: Medran Hill, 1972, p. 51—66.
- 5.153. Studdert-Kennedy M. Speech perception. — Contemporary issues in experimental phonetics. — New York, Academic Press, 1976, p. 243—293.
- 5.154. Tatham M. A. A. The control of muscles in speech. — Occasional Papers. Univ. of Essex Language Center, 1969, N 3, p. 23—40.
- 5.155. Tatham M. A. A. A speech production model for synthesis-by-rule. — Working Papers in Linguistics, Computer and Information Sciences Center, Ohio State Univ., 1970, N 6.
- 5.156. Tokizane T., Shimazu H. Functional differentiation of human skeletal muscles. — Tokyo: Univ. of Tokyo Press, 1964.
- 5.157. Verillo R. T. Specificity of a cutaneous receptor. — Perception and Psychophysics, 1966, v. 1, p. 149—153.
- 5.158. Verillo R. T. Vibrotactile sensitivity and the frequency response on the Paccinian corpuscle. — Psychol. Sci., 1966, v. 4, p. 135—136.
- 5.159. Wickelgren W. A. Context-sensitive coding, associate memory, and serial order in (speech) behavior. — Psychological Review, 1969, v. 76, p. 1—15.
- 5.160. Wood S. The weakness of the tongue — arching model of vowel articulation. — Working Papers, Phonetic Lab., Lund Univ., 1975, å 11, p. 55—108.

- Анестезия 222
 Аркана метод 275
 Бернулли эффект 52
 Вебстера уравнение 153
 Веретена 214, 254
 Внутренняя модель 276
 Возбуждение парциальное 201
 Воздушного потока уравнение 54
 Возмущения метод 158
 Волновое уравнение 144
 Галеркина метод 161, 194
 Галлагера функция 291
 Гистограммы импульсов 229
 Голосовой щели импеданс 166
 Голосовые складки:
 вертикальные смещения 49, 78
 жесткость 46
 собственные функции 66, 74
 схлопывание 70
 уравнение 65, 73
 Гольджи органы 212
 Губы:
 амплитудно-частотные характери-
 стики 109
 смещение 110
 скорость 110
 уравнение 113
 Гука закон 16, 17
 Гюгонно уравнение 53
 Давление подвязочное 58
 Давления перепад 50
 Двигательные единицы 210
 Длительности вариации 269
 Задержки время 83, 224, 226
 Запирания явление 151
 Затухание эквивалентное 189
 Излучения импеданс 166
 Коартикуляция 278
 Кодовых расстойный спектр 289, 293
 Коэффициент усиления 238
 Келли—Локбаума схема 161
 Лапласа преобразование 236
 Ли эффект 225
 Модальное управление 254
 Модуляция амплитудно-импульсная 231
 Модуль сдвига 46
 Мышцы эффективность 82
 Наблюдаемость 253
 Нозализации эффект 98
 Неразрывности условия 53
 Небная занавеска:
 амплитудно-частотная характери-
 стика 102, 104
 высота 97
 коартикуляция 106
 Нижняя челюсть:
 амплитудно-частотная характери-
 стика 88
 блокировка 270
 вязкость 95
 жесткость 91
 момент инерции 87
 переходные процессы 30, 94
 собственная частота 89, 91
 точность управления 85
 Окончания первичные и вторичные 214
 Оперение мышечной активности 262
 Пачини тельца 212
 Перерегулирование 268
 Пограничный слой 185
 Префиксность 285
 Пуазейля закон 174
 Пуассона коэффициент 16
 Резонанс дополнительный 37, 196, 208
 параметрический 37
 радиальный 145
 Рейнольдса число 52, 197
 Релаксация 34
 Рецепторы тактильные 213
 Системы астатические 238
 статические 238
 Смычка 207
 Сопротивление динамическое 53, 199
 Стенок импеданс 149
 податливость 146
 Точность артикуляции 268, 274
 Трение капиллярное 186
 Условия граничные мышцы 23
 Устойчивости область 242
 Фурье метод 23
 Хилла уравнение 10, 36
 Шредингера метод 246
 Шум турбулентный 200
 Эйлера метод 56
 Электромиограммы (ЭМГ) 226
 Энергетический критерий 258, 267
 Эри функция 20
 Юнга модуль 16
 Юссона теория 37, 196, 208
 Язык:
 амплитудно-частотная характери-
 стика 123
 переходный процесс 136
 поворот 119
 смещение 119
 собственные функции 127, 131, 136
 уравнение 126, 130

ОГЛАВЛЕНИЕ

	Стр.
Предисловие	7
Введение	8
Глава 1. Мышца	10
1.1. Микромеханика мышечного сокращения	10
1.2. Упругие свойства мышечной ткани	13
1.3. Одномерная модель мышцы	21
1.4. Параметрические явления	33
Глава 2. Голосовой источник	39
2.1. Строение и кинематика гортани	39
2.2. Механические свойства тканей голосовых складок	44
2.3. Колебательные процессы в голосовых складках	48
2.4. Аэродинамика голосового источника	50
2.5. Механика колебаний голосовых складок	59
2.6. Одномерная модель голосовых складок	63
2.7. Двухмерная модель	72
2.8. Модель вертикальных движений	78
Глава 3. Механика артикуляции	81
3.1. Нижняя челюсть	81
3.2. Небная занавеска	96
3.3. Губы	107
3.4. Язык	117
Глава 4. Акустика речеобразования	143
4.1. Уравнение речевого тракта	143
4.2. Излучение речевой волны	165
4.3. Голосовая щель и подсвязочная область	170
4.4. Потери в речевом тракте	183
4.5. Назализация	191
4.6. Турбулентные процессы	197
4.7. Взрывные процессы	202
Глава 5. Система управления артикуляцией	210
5.1. Элементы структуры системы управления	210
5.2. Рецепторы	211
5.3. Скорость распространения возбуждения	220
5.4. Обратная связь	221
5.5. Электрическое поле потенциалов мышечного волокна	226
5.6. Автономный контур управления	234
5.7. Пространственно-временная организация системы управления	244
5.8. Элементы кодовой структуры речевого потока	283
Список литературы	295
Предметный указатель	311