

Э. КЮННАП

## СИНТЕЗ РЕЧИ

Синтез речи уже в течение нескольких веков привлекает внимание исследователей. Более серьезные работы в этой области начались в XVIII веке, когда некоторые ученые стали применять синтез для исследования речеобразования [148]. Первый известный синтезатор речи был изготовлен в России профессором Краценштейном [48, 6], который сконструировал трубы различной конфигурации для воспроизведения пяти гласных: *a, e, i, o, u*. Серьезным достижением в 1770—1780 годах была также «говорящая голова», сконструированная Аббе Микала. Большой шаг вперед был сделан Кемпеленом, который в 1791 г. изготовил для этой цели механическое устройство [24, 49, 201]. Его изобретение можно по праву назвать «говорящей машиной». Речь в этом устройстве получалась с помощью воздушной струи, проходящей через резонаторы в виде органичных трубок, поперечные сечения которых изменялись вибрирующим язычком. Для образования воздушной струи использовались воздуходувные мехи. Человек-оператор производил струю воздуха и манипулировал язычком. Устройство могло образовывать 5 гласных, 19 согласных и соединять их в отдельные слова. Качество речевых сигналов в этом синтезаторе во многом зависело от искусности оператора. Кемпелен утверждал, что потребуется примерно три недели, чтобы оператор мог заставить машину «говорить» на латинском, французском, итальянском языках и несколько больше — на немецком из-за сложности получения шипящих согласных и большого количества согласных вообще. Он предполагал, что в следующей изготовленной им машине будет больше язычков и манипуляцию можно будет производить клавиатурой, как на пианино.

Профессор Уитстон по данному Кемпеленом описанию и по тем же принципам построил также говорящую машину, которую он демонстрировал в Дублине в 1835 г. Профессор Фабер из Вены показывал «говорящую машину» на выставке в Лондоне в 1846 г. Посетители выставки могли прослушать отрезки обыкновенной речи, шепотную речь, разговор и пение. Демонстрация заканчивалась исполнением государственного гимна Англии. Эта машина была значительно лучше машины Кемпелена, особенно в том отношении, что в ней основной тон воспроизводился с переменной частотой, позволяющей имитировать пение. Гельмгольц, опубликовавший результаты своих исследований в 1863 г., работал с последовательно соединенными камертонами и резонаторами; Миллер в 1926 г. и Штрумпф в 1927 г. использовали систему свистков. Синтезаторы, но менее удачные, были созданы еще многими авторами [75]. Позже синтезаторы речи на трубах были разработаны в Швеции [7, 60] и из прозрачной акриловой смолы — в Японии [189].

Первый электрический синтезатор речи создан в 1922 г. Стюартом [177], который впервые попытался синтезировать речевые сигналы с помощью только одних электрических цепей. Более современную конструкцию имел водер Дадли [44, 47, 48]. Это было первое электроакустическое устройство, воспроизводившее последовательно звуки и аппроксимирующее динамику речи. Для производства гласных и согласных в этом аппарате имелись электрические механизмы, моделировавшие речь человека. Управлял водером оператор путем нажатия на различные педали. В 1922 году Дадли сконстру-

ровал вокодер с таким же электрическим синтезатором, как и у водера, но управляемый посредством электрических сигналов, получаемых с выхода анализатора самого вокодера. Последний предназначался для передачи в сжатом виде речевых сигналов на дальнее расстояние, т. е. для уплотнения каналов связи. Анализ речи производился посредством системы фильтров; полученные мгновенные электрические величины передавались по линии связи на другой конец линии, где они восстанавливались, и таким образом снова образовывалась речь. При использовании вокодера можно намного уменьшить объем канала связи. Поскольку вокодеры только восстанавливают по электрическим импульсам речевые сигналы в приемном конце связи и имеют мало общего с синтезаторами речи, то мы на них в дальнейшем останавливаться не будем, а напомним лишь о некоторых из них, используемых самостоятельно.

Значение синтезаторов возросло в связи с увеличением объема исследовательских работ по автоматическому распознаванию речевых сигналов, и только в последние 10—15 лет они стали важным техническим средством для исследователей этой области науки [25, 35, 55, 57, 62, 82, 117, 122, 147, 175, 199]. Результаты изучения речевых сигналов и успехи в области электроники позволили создать аппараты искусственной речи, которые синтезировали отрезки речи, звучащие довольно натурально.

Электрический синтезатор речи представляет интерес для многих исследователей. Для инженеров связи он важен тем, что совместно с анализатором речи позволяет сократить объем каналов телефонной или радиосвязи. Когда развитие техники достигнет такого уровня, что аппараты для анализа, синтеза и перевода речи будут иметь размерность и вес, удобные для употребления их людьми на ходу, то синтезатор речи позволит вести разговор на разных языках. Но уже и теперь он помогает в некоторой мере исправить качество произношения, а также способствует устранению дефектов речи [151]. На основе синтезаторов речи можно построить читающие машины для слепых. Изготовлены соответствующие приборы с акустическим выходом [152]. Совместно со спектральными и другими анализаторами синтезатор позволяет значительно эффективнее изучать проблему распознавания речевых сигналов, а решение последней, т. е. проблемы автоматического распознавания речевых сигналов, дает возможность конструировать управляемые голосом машины и устройства, использовать в сложных системах управления команду человека-оператора без промежуточного кодирования, а также устройства для распознавания человека по голосу, имеющие большое значение в криминалистике. Синтезатор является прибором, позволяющим исследователю контролировать достоверность анализа. Спектрограммы речи содержат много избыточной информации. С помощью синтезатора имеется возможность выяснить избыточное и определить, какие характеристики существенны для речевого сигнала, а также определить связь между воспроизводимыми звуками и конфигурациями речевого тракта.

Основываясь на акустических признаках, используемых для распознавания речевых сигналов, можно выработать правила синтеза речи [72, 96, 120, 121, 130, 139]. При построении речи из заранее записанных фонем для получения качественной речи их нужно записывать при разных оттенках произношения. Если же ограничиться только небольшим количеством элементов, речь получится неестественной. Трудно синтезировать только согласные, часть которых существует лишь в динамике. Можно синтезировать по слогам, но тогда память машины должна иметь большой объем. Петерсон и др. [149] предложили использовать диады, т. е. сегменты речи, содержащие участки двух соседних фонем и переходы между ними, или триады, содержащие конец первой фонемы, вторую фонему и начало третьей фонемы. Они нашли, что для воспроизведения английской речи достаточно иметь 8500 диад. Однако это предложение трудно реализовать, поскольку количество диад очень велико. В [29, 119—121, 163] на основе анализа речевых сигналов приводятся данные, показывающие, что при произношении слога образование согласных зависит от характера и скорости изменения частоты, а также от амплитуды формант гласной, особенно второй, следующей за данной согласной, причем начало изменения второй форманты у всех гласных одинаково. Эта начальная точка частоты форманты названа авторами «локусом». Определены характерные «локусы» для всех согласных английского языка [36, 79, 80]. Для нахождения «локуса» использовалась

также модель речевого тракта на трубах [57]. Учет этих характеристик при синтезе позволяет получать речь, звучащую довольно натурально. Аналогичные закономерности синтеза речи рассмотрены также в [42, 165].

Определение инвариантных фонетических факторов речи затрудняется большими различиями в произношении дикторов, а также разнообразным произношением одних и тех же фонем одним и тем же диктором. Для определения влияния формантных частот на качество фонем в зависимости от изменения частоты основного тона был использован синтезатор гласных, состоящий из генератора 100 гармоник, амплитуду которых можно было регулировать в пределах 50 дб. Каждая гармоника была записана на магнитный барабан и затем они суммировались по желанию экспериментатора. Установлено, что смещение частот формант в зависимости от изменения частоты основного тона (но не пропорционально этому изменению) наблюдается главным образом у гласных *æ* и *и* английского языка. У некоторых гласных чем выше был основной тон, тем лучше можно было определить частоту формант, так как в этом случае их перекрывание уменьшалось. Число оборотов барабана регулировали, что позволяло изменять частоту основного тона. Согласные записывались на барабан с необходимыми паузами [184].

Синтезаторы можно подразделить на две группы: на модели артикуляции (аналоги речевого тракта) и на модели акустических волн речевых сигналов [24, 46, 94, 174]. Для этой цели могут быть использованы электрические и физические модели или ЭЦВМ. Особый класс синтезаторов представляют устройства, где речевые сигналы образуются соединением между собой отрезков речи, заранее записанных в память устройства. Такое соединение можно производить также с помощью ЭЦВМ [10, 132, 149]. Правильнее было бы назвать их не синтезаторами, а устройствами для монтажа речи из отдельных участков.

Синтезаторы как модели артикуляции в своем большинстве моделируют полость рта, радиус сужения этой полости, поперечное сечение сужения и отношение площади этой узкой части к расстоянию от нее до губ. В [95] предложено моделировать максимальную высоту передней и задней полостей и сужение между ними, фиксировать траектории этих точек и определять их изменение во времени. Электрические секции динамической модели речевого тракта имеют эквивалентные геометрическим размерам отдельных участков тракта электрические напряжения, при изменении которых воспроизводятся короткие отрезки речи. Для синтеза речи синтезатор должен быть управляемым, и для этой цели наиболее подходят гибридные устройства, состоящие из аналоговой части для получения необходимых электрических параметров и из ЭЦВМ для управления этими параметрами и их обработкой.

Синтезаторы, моделирующие акустические кривые речевых сигналов, синтезируют речь из 3—4 формант, используя в качестве источника возбуждения релаксационные генераторы (источник тона для получения гласных и сонорных согласных) и генератор шума (для получения согласных). В некоторых синтезаторах параметры речи наносятся на пленки и преобразуются в речь под воздействием лучей света или иными способами, например с помощью токопроводящих чернил [61, 99].

Зависимость частот формант от основного тона при произношении изолированных гласных рассматривается также в [5].

Влияние спектрального состава звука голосовых связок на качество речи исследовано в [13, 112, 143]. Найдено, что в случае применения в синтезаторе импульсов сложной формы, похожих на импульсы голосовых связок человека, синтезированная речь звучит более натурально, чем при использовании для этой цели импульсов простой и постоянной формы.

Качество синтезированных звуков речи исследовалось многими авторами [26, 28, 33, 67, 74, 76, 80, 83, 88, 125, 133, 134, 137, 138, 142, 143, 154, 170, 181, 196]. Установлено, что синтез из звеньев, меньших, чем слово, не дает удовлетворительного результата, особенно по той причине, что частоты основного тона и формант на месте стыка — разные и их амплитуды не совпадают. Выяснено, что у аналогового синтезатора синхронизация управляемых сигналов его контуров может предотвратить эти недостатки.

## 1. Устройства для монтажа речи из отдельных отрезков

Для исследования элементов речи, несущих информацию в коммуникации, некоторые отрезки речи были записаны на магнитную ленту. В дальнейшем была предпринята попытка образовать из этих отрезков связную речь. Когда элементы синтезированной речи состояли из отдельных гласных и согласных, то речь получалась неестественной и в большинстве случаев непонятной из-за отсутствия переходов, присущих натуральной речи. Например, фонема  $\omega$  в словах *wir* и *wür* имеет разные акустические кривые и если при синтезе речи использовать фонемы, произнесенные изолированно, то такое взаимное влияние не учитывается. Однако имеется ряд согласных, на акустические кривые которых соседние гласные не влияют [93]. При синтезе речи это необходимо принимать во внимание [73]. В [77] описывается образование стандартной речи из блоков. Речь была смонтирована из отрезков, заранее записанных на магнитную ленту. Каждый отрезок имел свой модуль. Модули выбираются из групп отрезков речи, записанных на вращающийся магнитный барабан. При нажатии кнопки возникает электрический сигнал, который усиливается и прижимает соответствующую головку к барабану. Полученные отрезки речи были понятны слушателям в объеме 75%. Аналогичные работы по синтезу речи из сегментов английской и японской речи описаны в [68, 124, 144, 147, 173], а русской речи — в [3, 4].

В [149, 166, 192] описывается синтез речи из сегментов, записанных на магнитную ленту. Сделаны попытки сочетать гармонические составляющие, форманты и огибающие спектра между смежными сегментами речи. Для синтеза необходимо 8500 сегментов. Приведен синтез предложения из 40 таких сегментов. Использовался ручной метод соединения. Предлагается использовать для этой цели ЭЦВМ. Аналогичные исследования описаны в [132, 193].

Разработана аппаратура для распознавания, перевода, синтеза и печатания односложных слов английского, французского, немецкого и испанского языков [145, 146]. Каждое слово произносилось в микрофон, анализировалось, превращалось в слоговый код и затем синтезировалось в речь или было напечатано на одном из упомянутых языков. Синтезатор состоит из слогового декодера, слогового селектора, слогового запоминающего устройства, усилителя и громкоговорителя. Слоговый код подается в слоговое декодирующее устройство. Слоги записаны на магнитный барабан. Каждый слог имеет свой код. Отдельные слог, слова и даже целые фразы образуются включением соответствующей магнитной головки, под которой проходят записанные на барабан слог, причем для получения связной речи между ними выдерживается некоторая пауза.

## 2. Синтезаторы как модели речевого тракта

Если рассматривать совместно спектральный анализ речевых сигналов и артикуляцию, то можно обнаружить, что между ними существует корреляция. Построить модель трехмерного речевого тракта довольно трудно. Да и в этом нет необходимости, поскольку исследователя интересуют только результаты сравнения синтезированной речи с естественной при различных конфигурациях речевого тракта. Модель должна объяснить зависимость между конфигурацией речевого тракта при произношении той или другой фонемы, слога, слова и т. д. и результирующим спектром речи.

Первые попытки синтеза речи посредством электрических цепей были предприняты исходя из статической [50, 163, 176] и динамической [66, 117, 157, 174] аналогии с речевым трактом. Необходимые размеры участка речевого тракта, подлежащие моделированию, были определены по рентгеновским снимкам во время произношения. Модель тракта [163] представляет собой искусственную трансмиссионную линию акустических сигналов, состоящую из последовательно соединенных секций из индуктивностей и емкостей. Электрические характеристики отдельных секций могут быть выбраны так, чтобы напряжение соответствовало геометрическим размерам данного участка тракта и чтобы после суммирования и преобразования напряжений с помощью громкоговорителя возникали волны воздушного давления, воспринимаемые аудитором как отрезки речи. Этот искусственный тракт в комбинации с подходящим источником энергии и громкоговори-

телем может воспроизводить, кроме гласных, и другие речевые сигналы. Электрические величины отдельных секций при синтезе речевых сигналов изменяются аналогично изменению речевого тракта при речеобразовании. Функции каждого артикулятора могут быть изменены и исследованы в отдельности. Модель состоит из 24 последовательно соединенных секций. Акустическим эквивалентом этой линии является цилиндрическая камера с поперечным сечением  $6 \text{ см}^2$  и длиной  $12 \text{ см}$ . Каждая из 24 секций представляет собой участок речевого тракта длиной  $0,5 \text{ см}$ . Фиксированное поперечное сечение является, разумеется, упрощением модели, так как в действительности поперечное сечение речевого тракта не представляет собой постоянной величины. Синтезированные гласные звучали довольно натурально.

В [176] модель речевого тракта состоит из 35 секций. Каждая секция представляет собой участок длиной  $0,5 \text{ см}$ , а модель в целом — поперечное сечение тракта от  $0,17$  до  $17 \text{ см}^2$ . Электрическая линия возбуждается источником либо периодических импульсов, моделирующих выход голосовых связок, либо случайного напряжения, моделирующего шум турбулентного движения воздуха, проходящего через сужения речевого тракта. Аналоговый синтезатор может качественно синтезировать все английские гласные и несколько согласных. Он может быть полезен при исследовании связи между артикуляторными движениями и образующимися звуками [172, 173]. Результаты исследования по синтезу назальных согласных на этом же аналоговом синтезаторе, только усовершенствованном дополнительными секциями, приведены в [92].

Стивенс и Хаус [174] установили довольно простую зависимость между артикуляцией и формантной структурой речи. Ими описано 306 различных комбинаций из трех величин: расстояния от голосовых связок до сужения тракта (обычно от 4 до  $13 \text{ см}$ ), радиуса сужения (обычно от  $0,3$  до  $1,2 \text{ см}$ ) и степени открытия рта, т. е. отношения площади ротового отверстия к его длине, и измерены соответственно частоты формант. Упомянутые авторы разработали также аналоги для синтеза назальных согласных [92, 94]. Аналогичные исследования описаны в [34, 51]. В [95] предлагается моделировать максимальные размеры передней и задней полостей рта и сужения между ними, фиксировать траекторию этих точек и определять их изменение во времени.

Разработана модель речевого тракта, состоящая из 45 секций, каждая из которых соответствует отрезку речевого тракта длиной  $0,5 \text{ см}$ . Параметры каждой секции можно изменять соответственно изменению поперечного сечения тракта, т. е. от  $0,16$  до  $16 \text{ см}^2$  [45, 57, 59]. В этой модели можно имитировать носовые полости и синтезировать гласные и согласные фонемы.

Динамически управляемый электрический аналог речевого тракта [157, 159, 160] позволяет синтезировать последовательные речевые сигналы. Акустическая передающая линия между голосовой щелью и губами выполнена из 11 электрических управляемых переменных и трех фиксированных LC цепей. Аналог можно возбуждать с помощью источников тона и шума как поочередно, так и одновременно. С одинаковой легкостью можно синтезировать все гласные и большинство согласных английского языка. Управляющие напряжения сняты с цепи сопротивлений, в память которых заложены артикуляторные конфигурации. Синтезирующее устройство выбирает моменты возбуждения и артикуляторные конфигурации с подходящей очередностью, а сглаживающие цепи используются для обеспечения естественного перехода от одной конфигурации к другой. Для получения более плавных отрезков речи управление динамическим аналогом речевого тракта производится посредством ЭЦВМ. Управляющая система состоит из ряда генераторов трапециевидной формы импульсов. В качестве поверхности раздела используется 32-канальный аналого-цифровой преобразователь, который может подавать с ЭЦВМ в любую секцию аналога индивидуальные управляющие напряжения, а также управлять амплитудой указанных импульсов, интенсивностью источника шума и соединением секций носового пространства. Количество управляемых параметров 40. Можно определить существенные артикуляторные параметры, необходимые для произношения определенных фонем. Розен [158, 160, 161] предложил использовать несколько упрощенный по сравнению с предложениями других авторов [82, 155, 197, 198] эквивалент речевого тракта.

В [21] описана геометрическая модель для синтеза гласных и согласных, аппроксимирующая типичные функции площади поперечного сечения речевого тракта. Артикуляторные конфигурации полностью детерминированы пятью переменными функциями. Предлагается использовать ЭЦВМ для определения посредством этих функций площади поперечного сечения для расчета формантных частот и ввести в действие аналоговый синтезатор с жесткой программой.

В [85] предложена динамическая артикуляторная модель для синтеза отрезков речи. Из дискретного фонемного входа модель непрерывно генерирует в соответствии с конфигурацией артикуляторного механизма колебания ее формантных частот. Синтезируются только гласные и взрывные согласные, но методика позволяет синтезировать и другие фонемы.

Электрический аналог речевого тракта LEA составлен из 45 последовательно соединенных фильтров, состоящих из конденсаторов и катушек, причем каждая секция представляет собой отрезок речевого тракта длиной 0,5 см [56, 94]. Параметры каждого звена можно изменять в соответствии с изменением поперечного сечения тракта в пределах от 0,16 до 16 см<sup>2</sup>. LEA и другие подобные ему синтезаторы [50, 176] используются либо как аналоговые вычислительные машины для преобразования параметров артикуляции, полученных измерением речевого тракта методом рентгеновых лучей, в спектральную форму произнесенных фонем, либо как синтезаторы для получения искусственных стационарных речевых сигналов. С помощью динамического аналога [161] можно было синтезировать и отрезки речи, поскольку изменением параметров секции управляет электрический сервомотор. Результаты фундаментальных исследований зависимости формантных частот от артикуляции приведены в [7, 174].

### 3. Синтезаторы как модели акустических кривых речевых сигналов

Синтез речи моделированием ее акустических кривых производится синтезаторами сосредоточенного или формантного типа, состоящими из нескольких контуров колебаний (обычно 3—4), каждая из которых соответствует колебаниям формантных частот [63, 64, 115, 116, 133, 194, 195]. Выходом синтезатора является электрическое напряжение, изменение которого во времени должно соответствовать изменению акустической волны при произношении определенных фонем, слогов и т. д. Речевой тракт человека возбуждается богатыми гармониками импульсами, возникающими при прохождении воздуха через голосовую щель. Частота импульсов является частотой основного тона. Часть гармоник усиливается, часть подавляется в зависимости от конфигурации речевого тракта, т. е. от произносимой фонемы. Часть фонем, не имеющих основного тона, образуется турбулентным прохождением воздуха через сужения речевого тракта. Некоторые фонемы, например сонорные согласные, образуются одновременно двумя источниками. В результате возбуждения речевого тракта в нем возникают переходные процессы, состоящие из шумовых и затухающих периодических колебаний разных частот, т. е. из формант. Каждая форманта имеет частоту, равную центральной частоте гармоник этой области. Такой процесс моделируется в синтезаторах формантного типа.

В синтезаторе [45] используется электронная схема, для возбуждения которой применяются два источника колебаний: релаксационный генератор (генератор тона), имитирующий звук гортани, и генератор шума. Разработана упрощенная модель речевого тракта, состоящая из трех контуров переменной и одного контура постоянной частоты колебания. Как утверждают авторы, синтезированные фонемы звучали довольно натурально [173].

Генерирование акустических волн речевого сигнала с помощью электрических цепей, имитирующих систему речеобразования у человека, описано в [63]. Приведены схемы и математическое описание синтеза при параллельном и последовательном соединении контуров формантных частот. В [9, 11, 12, 33, 100] предложены синтезаторы такого же типа. В [72] описывается конструкция синтезатора, состоящего только из одного контура резонансных частот. Сделана попытка найти зависимость между шириной и

частотой формант и естественностью синтезированных звуков. Для возбуждения контура использовался выход генератора импульсов. Полученные данные показали, что, изменяя ширину первой форманты и характер перехода второй форманты, можно синтезировать носовые согласные [136]. Синтез изолированных гласных производился также посредством устройства, состоящего из 12 полосовых фильтров, на вход которых подавались напряжения от генератора пилообразных импульсов. При синтезе импульсы направляли только в те фильтры, которые были необходимы для синтеза данной фонемы. Установлено, что звуки *y* и *o* являются одноформантными, звук *a* — двухформантным, а остальные — трехформантными. Синтезированные звуки *y*, *o*, *a* и *u* распознавались в среднем на 90%, а другие — в пределах от 6 до 90%. Синтезатор можно использовать как самостоятельно, так и в системе вокодера [2].

Посредством синтезатора последовательного соединения синтезировано 10 гласных с различной длительностью звучания — от 80 мсек до 1,3 сек. Аудиторы (16 человек) определяли качество синтеза. Выяснено, что для распознавания синтетических гласных продолжительность их звучания имеет меньшее значение, чем для распознавания естественных гласных [196].

Исследовано качество 13 гласнообразных речевых сигналов, полученных на синтезаторе, состоящем из четырех соединенных параллельно формантных контуров. Изменялись частоты только двух формант либо вручную, либо с помощью нанесенных на ацетатную пленку кривых. Качество и распознавание синтезированной речи определялись аудитором [70]. Синтезатор, использующий только шесть изменяющихся параметров, позволил синтезировать отдельные фонемы с разборчивостью до 70% и короткие слоги и слова с разборчивостью до 100%. Этими параметрами являлись: частоты первых трех формант, амплитуда источника согласных (генератора шума), амплитуда и частота источника гласных (генератор тона) [200]. В [137, 138, 141] рассматриваются результаты исследования синтеза гласных и согласных с помощью синтезатора параллельного соединения. Изменялись частота и амплитуда формантных компонентов речи при разной скорости подъема, выдержки и спуска амплитуд генераторов тона и шума и их отношение в течение одновременного возбуждения контуров. Сделана попытка синтезировать некоторые гласные венгерского языка и исследовано влияние ширины формант на качество синтезированной речи. Для этого использовался генератор пилообразных импульсов и фильтры со средней частотой пропускания 470 и 680 гц [182].

В [15] рассматривается возможность синтеза изолированных гласных путем суммирования гармоник, полученных из остронастроенных фильтров. Полосы пропускания фильтров определены по спектральному анализу этих же трех гласных, произнесенных в С-Г-С слогах. Полученные искусственные согласные сравнивались слушателями с естественными и определялась величина ошибок. Найдено, что звуки, полученные при использовании генератора шума, выход которого проходит через указанные фильтры, меньше отличаются от естественных, чем звуки, полученные в случае применения генератора тона, выход которого пропускается через те же фильтры. Разница эта, однако, невелика.

Изготовлен каскадный формантный синтезатор для синтеза гласных и согласных польского языка [100-105]. Синтез осуществляется изменением 12 параметров. Синтезатор состоит из четырех последовательно соединенных контуров для образования четырех формантных частот. В качестве возбудителя использовался генератор пилообразных импульсов, амплитуды которых могут быть изменены от 0 до 40 в, причем время их подъема, выдержки и спуска может быть выбрано от нескольких десятков до нескольких секунд. Выяснилось, что при синтезе польских носовых согласных в слове С-Г необходимо учитывать не только отношение частот формант, но и форму перехода второй форманты. В [154] описывается синтезатор параллельного соединения трех формантных частот для воспроизведения десяти гласнообразных фонем и приводится анализ качества синтезированных гласных. В синтезаторе [140] гласные получаются из основного тона и его гармоник, полученных из генератора синусоидальных колебаний. Можно выбирать как центр частоты формант гласного, так и их интенсивности. Иссле-

довался синтез речи с помощью стационарных сегментов [19]. Варьировались длительность сегментов и время подъема, выдержки и падения амплитуд выбранных частот формант. С помощью 50 RCL фильтров с постоянными полосами пропускания были образованы колебания формантных частот. Первичный сигнал был получен с мультивибратора, причем для образования данной фонемы можно было выбирать любые два фильтра из 50. Выяснилось, что если точно выдерживать продолжительность подъема, выдержки и падения амплитуд выбранных формант, то нет необходимости учитывать все подробности информации о формантах, на что другие исследователи обращают большое внимание.

Метод модуляции выхода генератора шума с помощью выхода релаксационного генератора для получения сонорных согласных, звучащих более естественно, приведен в [129]. Для исследования образования взрывных согласных в комбинации фрикативный — взрывной — фрикативный звук применялся синтезатор формантного типа [185]. Установлено, что перемещение пика энергии в спектре шипения, используемого для произношения фрикативных звуков, может являться намеком на место их артикуляции.

Предложен синтезатор, в котором выходы от генераторов пилообразных сигналов или шума проходят через цепи колебания, а выходы цепей соединены параллельно через усилитель в громкоговорителе [17, 18]. Изготовлен синтезатор для вокодера, который можно использовать также самостоятельно [169]. Вместо полосовых фильтров, обычно применяющихся у вокодеров, используются контуры колебания, фаза, частота и амплитуда которых могут модулироваться синхронно с основным тоном. Установлено, что во время произношения речи как резонансные частоты, так и декременты затухания варьируют. Синтезированная речь имела высокие качественные показатели и звучала натурально. Если из колебаний, образованных мультивибратором, отфильтровывали высокие гармоники и не учитывали указанные выше свойства речи, то в этом случае получалась речь достаточно натуральная, но звучащая грубо.

Построен синтезатор гласных с пятью цепями генерирования формантных частот, в котором частоту, ширину и интенсивность каждой форманты можно регулировать в отдельности [156]. Для изменения частоты формант применялись конденсаторы переменной емкости. Так как ими в данном случае можно пользоваться только при высоких частотах, то употребляли генератор с частотой 15 кГц. Применение конденсаторов переменной емкости позволило синтезировать связную речь. Частоту формант можно было изменять также изменением частоты генератора синусоидальных колебаний. Получены удовлетворительные результаты.

Другой вид опытного синтезатора для вокодера, который употребляется и самостоятельно, описан в [71, 73]. Источниками возбуждения синтезатора являются генераторы тона и шума, а в качестве колебательных контуров, числом девять, используются механические колебания кристаллов, которые одновременно являются и преобразователями механических колебаний в электрооптические с помощью луча света. В вокодере синтезатор управляется голосом человека, говорящего на входном конце связи.

Синтезатор, состоящий из генераторов тона и шума, 30 формантных каналов (фильтров с разной частотой пропускания в пределах от 100 до 12 000 Гц) и устройства автоматического запуска синтезированного отрезка речи, описывается в [1]. Синтезируются короткие отрезки грузинской и русской речи.

Синтез гласных только с двумя формантными частотами посредством синтезатора с управляемыми потенциальными полями, позволяющими выбрать из 20 каналов частот (фильтров с частотой пропускания от 70 до 6300 Гц) 2 канала, необходимые для данной гласной, описан в [184, 185]. Применяются токопроводящая бумага и контактные заклепки, как и в [9, 59, 68]. Для возбуждения системы используются генераторы импульсов трех- или четырехугольной формы и широкополосного шума. Необходимые частоты и их амплитуды выбираются вручную, по растру на токопроводящей бумаге.

Предложено синтезировать речевые сигналы с помощью управляемого синтезатора формантного типа последовательного соединения. Каждая фонема характеризуется состоянием ее формант. В спецификацию состояния входят: частота трех формант и



описание парциального состояния речевого тракта в течение произношения данной фонемы. Это парциальное состояние охватывает признаки назальности и присутствия голоса. Контрольными сигналами являются частота и ширина трех формант, пара комплекса частот полюс-нуль для назальных и пара комплекса частот полюс-нуль для фрикативных звуков, основной тон, амплитуда выходов генераторов тона и шума. Речь генерируется шаг за шагом соответственно последовательному решению математических уравнений состояний речевого тракта. Считается, что необходимое состояние достигнуто, когда значения формант близки к специализированным эталонным значениям [153].

Частоты формант получены и другими способами. Обратный синтез речевых сигналов (pattern playback) производится по спектрограммам речи, полученным с помощью дивизионного спектрографа [23, 27, 32, 33, 162]. Часть компонентов спектрограммы можно оставить без внимания и, прослушивая фонемы выбранной структуры, судить о значении той или другой частоты и ее амплитуды для воспроизводства данной фонемы. Устройство состоит из прозрачного барабана, на наружной поверхности которого наклеены полоски черной бумаги, представляющие элементы речи. Внутри барабана расположен источник света. Свет проходит через барабан и попадает на 13 фотозащелок, расположенных по вертикали и параллельно модулю барабана. При вращении последнего возникают электрические токи переменной силы, которые суммируются и направляются в громкоговоритель [162]. В этом синтезаторе принимаются во внимание гармоники с частотой не выше 3600 гц. Так, если использовать наименьшую частоту основного тона в 90 гц, то учитываются еще 40-е гармоники, если же применять частоту в 300 гц, то используются гармоники не выше 12-й. Наклеивая на поверхность барабана бумажные полосы с разными характеристиками, можно синтезировать различные слоги и исследовать структуру речи, а также влияние гармоник на ее отдельные компоненты.

Используя метод обратного синтеза речи и применяя 50 гармоник частотой до 6000 гц основного тона 120 гц, авторы [37, 119] определяли образование согласных по скорости и форме изменения частот второй и третьей формант. Вместо барабана применялось колесо с 50 концентрическими лентами из прозрачного материала. Этим же методом исследовалась конструкция согласных английского [25, 69] и французского [52, 53] языков. В [123] описаны основные акустические отличия согласных *ш*, *г*, *е*, *у*, находящихся между гласными и синтезированных методом, изложенным в [23]. Использовались гласные, состоящие из трех формант, каждая длительностью до 0,5 сек. Процесс образования указанных согласных изучался путем изменения находящихся в стационарном состоянии частот формант гласных, их интенсивностей, длительности и формы переходов. Были выбраны следующие частоты гласных: для *i* — 240, 2520 и 3000, для *a* — 780, 1200 и 2520 и для *u* — 240, 720 и 2520 гц. Если, например, изменяли  $F_2$  для *u* от 720 до 840 гц и обратно, то эта комбинация воспринималась слушателями как *ишu*. Такие эксперименты производились и с другими гласными. Синтез согласных *b*, *d*, *g* изменением перехода третьей форманты методом [26] исследован в [37, 79], а метод вокодерного репродуцирования, описанный в [162], использован в [8].

Параметрический искусственный говоритель (РАТ) описан в [115]. Автор этого синтезатора первым успешно применил принцип параметрического кодирования в синтезаторах, включая отдельный генератор функций, где параметры спецификации были введены графически. Для синтеза были использованы шесть параметров: частоты первых трех формант, интенсивность и периодичность сигнала тона и интенсивность сигнала шума. Эти параметры были нанесены в виде кривых на стекло и сканированы с помощью катодных лучей. На синтезаторе, построенном по этому же принципу, но несколько усовершенствованном, был синтезирован отрезок речи продолжительностью 45 сек, со средней скоростью 3,75 слога в секунду [183]. Удовлетворительные результаты были получены также при синтезе комбинаций фрикативных согласных с шипящими согласными английского языка [187]. Исследовано качество отрезка синтетической речи в зависимости от качества исходного источника [113, 143]. Еще более совершенный синтезатор позволил синтезировать речь с помощью восьми [97] и 12 параметров [9, 116].

Синтезировались слоги *si* и *zi* при изменении интенсивности источника тона на четырех уровнях (-5, -10, -15 и -20 дБ) [98]. Синтезатор другого вида описан в [68], где управление параметрами производится кривыми, нанесенными на ленту из пластического материала токопроводящими чернилами.

Сообщается о разработке синтезатора, названного электронным аналогом голоса (EVA). Прибор снабжен программируемым генератором функций и управляемым синтезатором формантного типа. Исходные данные вводятся в устройство в виде 12 кривых, нарисованных токопроводящими чернилами на листе бумаги, вдоль которого перемещается считывающее устройство. Метод задания исходных данных предусматривает их достаточно быстрое изменение в процессе исследования отдельных составляющих речи [99, 191]. В [150] приводятся результаты изучения образования согласных, синтезированных по описанному выше методу.

Синтезатор формантного типа (OVE-II) управляется кривыми, нанесенными токопроводящими чернилами [57, 58, 61]. Эти сигналы, представляющие амплитуды разных частот, принадлежат развертке, полученной с помощью потенциометра со скользящим контактом. Кодированный лист — неподвижный, и цилиндр потенциометра совершает прямолинейное движение вдоль листа. Если синтезатор OVE-I [59] предназначен для синтеза только гласных, то OVE-II [94] позволяет синтезировать различные отрезки речи, состоящие из гласных, согласных и их комбинаций. Синтезатор состоит из соединенных последовательно формантных контуров, через которые пропускаются сигналы генераторов тона и шума с изменяющейся амплитудой. Для воспроизведения отрезка речи длительностью 3 сек авторы оперировали 12 параметрами. Синтезатор разработан на основе теоретических исследований [7, 58]. Он может быть использован также в системе вокодера. В [125, 126] приводятся результаты исследования назализации на синтезаторе OVE-II.

#### 4. Синтез речевых сигналов с помощью ЭЦВМ

В процессе синтеза речи ЭЦВМ могут быть использованы самостоятельно, но в большинстве случаев их применяют в качестве устройств, управляющих синтезаторами аналогового типа в гибридных системах. Они используются в качестве моделирующих устройств электрических цепей анализа и синтеза речи, позволяющих сократить необходимое для конструирования и изготовления цепей время, исчисляемое годами, до нескольких месяцев, затрачиваемых на составление программ. Таким образом, они позволяют исследовать такие процессы речи, на изучение которых другими методами потребовалось бы очень много времени. Составлено большое число программ для синтеза речевых сигналов. Обычно программа предусматривает определение акустических признаков звука посредством использования типичных значений частот формант, частоты основного тона, длительности и интенсивности каждой фонемы. В памяти ЭЦВМ записаны правила, показывающие, каким образом влияет на эти типичные значения контекст. «Говорящие» ЭЦВМ синтезируют довольно разборчивую речь и даже пение, однако далеко не удовлетворительного качества.

Синтез речи на ЭЦВМ рассматривается в [167]. Акустические кривые речи преобразовывали в цифровую форму. Манипулировали спектром речи в цифровой форме и прослушивали результаты, воспроизводимые громкоговорителем. Программа, названная MUSE, составлена для ЭЦВМ JBM-7090. Акустические параметры специфицированы не в промежутке квантования, когда они постоянные, а в промежутке, в течение которого все параметры изменяются линейно. Язык MUSE содержит следующие параметры: номер спектра, время в сентисекундах, умножение амплитуды, частоту основного тона, ширину полос частот резонатора (десять градаций для каждого резонатора), частоту резонанса, видоизменение возбуждения. Было выбрано около 100 резонансных частот. Оба типа аналогов речевого тракта (распределенного и сосредоточенного) можно представить на языке MUSE. Лучшие результаты получались при использовании нескольких спектров. Полученные отрезки речи были разделены на сегменты. Поскольку к сегмен-

там речи предъявлялось требование, чтобы во время их звучания параметры изменялись только линейно, то для образования фонемы выбирали 2—3 сегмента. Выход синтезатора (электрическое напряжение) записывался в двоичном коде на магнитную ленту, с которой через цифро-аналоговый преобразователь и фильтр низкой частоты подавался в громкоговоритель. Программа использовалась только в исследовательских целях.

Разработана программа синтеза речи на ЭЦВМ JBM-7090, предусматривающая последовательную подачу на вход машины фонем в цифровой форме и получение на выходе звуковых волн [83]. При реализации артикуляторной модели синтеза речевой тракт был моделирован соединенными последовательно секциями, числом 21; каждая секция представляла собой участок речевого тракта длиной 0,8 см. Основная трудность составления программы заключалась в сглаживании переходов между фонемами. В [84] описывается программа для расчета спектра звукового сигнала по данным конфигурации речевого тракта. Спектр произнесенного диктором звука можно сравнить со спектром соответствующей полости модели речевого тракта. Параметры модели речевого тракта можно изменять до тех пор, пока генерированный спектр максимально совпадет со спектром произнесенного звука. Посредством ЭЦВМ анализировали и синтезировали в виде сумм ортонормальных функций отрезок речи длительностью 6 сек [126]. Результаты синтеза сравнивались слушателями по звучанию с натуральной речью. Синтезированные кривые речевых сигналов и их спектры сопоставлялись с соответствующими естественными звуками. Найдено, что различие между естественной и синтезированной речью при всех трех методах сравнения обусловлено перерывом синтезированных акустических волн на границе между периодами основного тона.

Синтезирована речь из цифрового кода. Цель исследования — выбрать минимальный код, еще обеспечивающий распознаваемость речи. Частоту специфицируют шесть битами, длительность звука — двумя (20; 60; 120 и 240 мсек). Каждый бит из шести включает один из контуров колебания (220; 400; 700; 1300; 2700 и 4000 гц). Выходы этих контуров смешиваются и направляются в громкоговоритель. Синтезированная речь звучала, однако, неестественно [16].

Составлена программа для синтеза гласных с целью: 1) выделения параметров сигналов речи с использованием автокорреляционной функции, 2) исследования восприятия слушателями изменения во времени параметров речи и 3) выработки искусственного источника речи для оценки качества передаточной функции речевого тракта. В программе предусмотрено определение девяти параметров речи, таких, как частота и ширина трех первых формант и др. Синтезированный разговорный речевой сигнал, образованный пятью японскими гласными, звучал довольно разборчиво и натурально [81].

Синтезирован акустический выход ЭЦВМ JBM-7770 или 7772. Потребитель обращается к машинам с помощью телефонного диска, набирая номер интересующего его блока информации. После этого машина также обращается к этому же блоку информации, обрабатывает входные данные, преобразует их и передает потребителю в виде устной речи. Эта система находит применение в банковских и страховых операциях, в производстве и торговле [190]. Аналогичная система описана в [110].

Составлены программы для моделирования на ЭЦВМ: процесса сегментации и синтеза речи, происходящего в нервной системе человека [135]; процесса взаимосвязи между источником голоса и речевым трактом и синтеза речи по данным анализа натуральной речи, проведенного синхронно основному тону [65].

В [108] описывается цифровая вычислительная техника, используемая для синтеза речи данных, полученных при анализе основного тона и синхронизации между формантами, записанных в память ЭЦВМ. Вся система позволяет исследовать роль основного тона и длительность звучания гласного в ударной речи. О синтезе речи с помощью ЭЦВМ JBM-7090 сообщено в [168]. Записанная на магнитную ленту информация преобразуется в колебания электрического напряжения и суммируется. Предлагается метод управления вычислительными операциями посредством произносимых в микрофон отдельных слов, образуемых семью согласными и пятью гласными [171]. Составлены

бессмысленные слоги из 2—3 звуков. Каждая последовательность соответствует определенной цифре в пределах от 0 до 5. Акустические колебания преобразуются в электрические сигналы, которые опознаются с помощью селекторных схем. Составлена на ЭЦВМ программа для синтеза речи как по данным фонем, записанным на магнитную ленту, так и по кривым речевых сигналов, а также по данным их спектрального анализа и др. [14, 43, 109, 111, 114]. Синтез выполнен главным образом для исследования речи. Приводятся результаты классификации синтезированных звуков. В [117] сообщается о моделировании на ЭЦВМ цифрового синтезатора формантного типа.

## 5. Синтезаторы, управляемые вычислительными машинами

Кроме упомянутых в предыдущих разделах синтезаторов этого типа, разработан еще ряд других устройств.

Разработан синтезатор, который имеет три последовательно соединенных контура колебания, генераторы тона и шума и специальное устройство для синхронизации акустических волн фонем. Синтезатор работает совместно с ЭЦВМ JBM-7090, но вместо нее можно применять и более простую машину. Синтезатор управляется девятью сигналами: интенсивностью генераторов тона и шума, частотой основного тона и трех формант и шириной этих формант. Управляющие сигналы вырабатываются ЭЦВМ по вводимым в нее кодовым сигналам [106].

Управляющие сигналы синтезатора каскадного типа получены по нарисованным на пленке из прозрачного материала кривым после их квантования и фотоумножения. Кривые видоизменялись до тех пор, пока не были получены удовлетворительные результаты, после чего они запоминались на перфокартах, которые в совокупности образовали библиотеку сигналов управления. Предложено применять в качестве сегментов речи не меньше чем слово [54, 109].

Управление синтезатором речи SCOFIS-I формантного типа параллельного соединения производится электрическим напряжением, полученным при прохождении импульса прямоугольной формы через активную обостряющую цепь. Синтезатор состоит из источников тона и шума, управляемых усилителями низкой частоты, из модуляторов, демодуляторов и операционных суммирующих усилителей. Управлению подлежат частота и амплитуда основного тона, частоты и амплитуды первых трех формант и амплитуда источника шума. Составлена соответствующая программа для ЭЦВМ Burroughs 220 [30, 31].

В [107] поставлена задача синтезировать речь из входных данных, состоящих из названий фонем (или аллофонов) при минимальном количестве информации об интонации, флексии и синхронизации между фонемами. В качестве синтезаторов использовались формантная и аналоговая модели речевого тракта. В обоих случаях был разработан ряд правил для генерирования управляющих сигналов из фонетической информации. Аналог речевого тракта состоял из 21 секции, каждая из которых моделировала 0,8 см его длины, начиная с точки возбуждения. Авторы утверждают, что синтез речи аналогом речевого тракта имеет некоторое преимущество перед продуцированием речи синтезатором формантного типа, но в обоих случаях натуральное звучание речи не было достигнуто.

На ЭЦВМ моделирован синтезатор формантного типа, где применяются значения частот полюсов и нулей передаточной функции речевого тракта. Результаты вычисления преобразуются и подаются в громкоговоритель. Передаточные функции гласных и сонорных согласных описаны четырьмя полюсами, носовые согласные — пятью полюсами и одним нулем. Эти фонемы генерированы источником тона по одной ветви синтезатора. По другой ветви источником шума генерированы согласные, которые описаны одним нулем и либо одним, либо двумя полюсами. Частоты и ширины полюсов и нулей, а также амплитуды источников управляемые [66].

Управление моделью речевого тракта (DAVO) [63, 157] производится как 24 напряжениями, определяющими площадь поперечного сечения тракта в виде функции рас-

стояния от его начала, так и тремя напряжениями, управляющими амплитудами голового и шумового возбуждения, а также связью между моделями полости носа и рта. В [40, 180] входные данные преобразовываются в ряд чисел, представляющих изменения управляющих напряжений во времени. Эти числа запоминаются в ЭЦВМ TX-0, затем считываются и по составленной программе производится синтез речи. Составлены отрезки речи длительностью до одной секунды.

Синтез речи по «правилам» с помощью системы, состоящей из программы ЭЦВМ и электронного аналогового синтезатора, описан в [89, 90]. Входом системы синтезатора являются фонема, слог, слово и т. д. Вычислительная машина определяет приближенные значения девяти параметров для каждых 10 мс и аналоговый синтезатор продуцирует звуки, специфицированные по очередности значений параметров. Келли и др. [106, 107] впервые синтезировали речи «по правилам» с помощью ЭЦВМ и использовали такую же программу без аналогового синтезатора. Но использование аналогового синтезатора позволяет применять менее мощную ЭЦВМ. Разработан специальный язык для ЭЦВМ, позволяющий управлять синтезатором аналогового типа в реальном масштабе времени [41].

Разработан синтезатор формантного типа, управляемый ЭЦВМ DDP-24 [20]. Комбинация синтезатор—вычислительная машина позволяет синтезировать речь в реальном масштабе времени. Время обращения к ЭЦВМ 5 мсек. Синтезатор может работать в режиме как последовательного, так и параллельного соединения. Частота  $F_3$  может быть изменена в широких пределах — до 20:1 по отношению к частоте  $F_1$ . Ширина каждой форманты также регулируется в широких пределах. Амплитуду можно регулировать в пределах до 60 дб. Все эти изменения частоты, ширины и амплитуды формант производятся независимо друг от друга. Программирование относительно простое. В [185] описывается синтезатор формантного типа, комбинированный с ЭЦВМ небольшого размера, где коэффициенты цепи определяются цифровым умножением. Серии команд записаны на консоле пишущей машины. Согласно такому методу, синтезируемые выражения подаются с подготовленной бумажной ленты в ЭЦВМ, которая вычисляет указанные выше коэффициенты.

Предложено образовывать звуковые сигналы синтезатором, состоящим из релаксационного генератора, трех модуляторов спектра, четырех формантных цепей гласных, пяти формантных цепей носовых фонем, двух формантных цепей согласных, генератора белого шума, трех антиформантных цепей, управляемых посредством соответствующего сигнала ЭЦВМ [87]. Изготовлена аналоговая модель речевого тракта, значения параметров которого вычисляет ЭЦВМ. Для нахождения необходимой формы речевого тракта при разных произношениях он исследуется с помощью рентгеновых лучей [22].

Синтезатор, управляемый аналоговой вычислительной машиной, описан в [91]. Речевой тракт аппроксимирован десятью секциями. Полученные результаты позволили приступить к разработке нового устройства, управляемого ЭЦВМ, аналоговая часть которого состояла из пятнадцати секций. Управляющие сигналы, в количестве 25, выдаются из ЭЦВМ JBM-7090 через каждые 5 мсек. Данные для вычисления управляющих сигналов были записаны на магнитную ленту в виде бинарных чисел, и синтезатор управлялся в реальном масштабе времени [129].

Для исследования речеобразования и определения значения характерных параметров синтезируемые предложения были записаны фонетически в память ЭЦВМ DDP-24. Выход синтезатора аналогового типа был подключен к входу осциллографа. Синтезатор имел 12 входов, управляемых ЭЦВМ в течение каждых 10 мсек. Экспериментатор прослушивал синтезированную фразу и в то же время наблюдал за ее траекторией на экране осциллографа. Все изменения параметров формант, выполненные вручную с помощью кнопок, перепрограммируются сразу для ЭЦВМ и отражаются также на экране осциллографа. Одновременно изменяется и звук, выходящий из громкоговорителя. Задача эксперимента — определить влияние длительности, интенсивности и частоты основного тона на ударение и флексию в предложении и синтезировать речь, звучащую более естественно [38, 39]. В [179] ЭЦВМ была использована для анализа и синтеза речи. Выходы из 100 фильтров, охватывающих частоты от 20 до 3980 гц,

выпрямляются и преобразуются в цифровую форму. Полученные данные, поданные в ЭЦВМ, используются для управления синтезатором формантного типа параллельного соединения. Синтезированы гласные и согласные.

Разработан управляемый ЭЦВМ синтезатор речи, предназначенный для экспериментальных исследований. Синтезатор формантного типа параллельного соединения связан с ЭЦВМ посредством цифро-аналогового преобразователя. Для возбуждения синтезатора используются, как обычно, генераторы тона и шума. Входная информация для программы записана на 8-дорожковой бумажной ленте и представляет собой последовательность синтезируемых звуков или слов [164].

В статье рассмотрены лишь наиболее типичные синтезаторы речи и приведено их условное разделение по принципиальным схемам и способам управления. Большинство из них еще не вышли из стен лаборатории. Но все возрастающая необходимость в замене человека в сложных системах управления машинами позволяет предполагать, что в недалеком будущем синтезаторы наряду с устройствами автоматического распознавания речи займут довольно важное место в системах управления и связи.

## ЛИТЕРАТУРА

1. Белозеров Л. Б., Экспериментальная установка для синтеза отрезков речи. Тр. Ин-та электроники, автоматики и телемеханики АН ГрузССР, 4 (1963).
2. Богданов Б. В., Синтез звуко сочетаний при помощи стационарных по спектру отрезков. Док. на Всесоюз. совещ. секция речи Комиссии по акустике АН СССР, М., 1959.
3. Дукельский Н. И., Методы пересадки звуков речи в фонетике. Вопросы языкознания, № 1 (1958).
4. Дукельский Н. И., Принципы сегментации речевого потока, М.—Л., 1962.
5. Кюннап Э. Ю., Математический анализ гласных эстонского языка. Докл. на Всесоюз. акустической конференции, М., 1968.
6. Мясников Л. Л., Физические исследования звуков русской речи, Изв. АН СССР, Сер. физ., 13, № 6 (1949).
7. Фант Г., Акустическая теория речеобразования, М., 1964.
8. Abramson A. S., Lisker L., Voice onset time in stop consonants: Acoustic Analysis and Synthesis. 5-e Congrès International d'Acoustique, Liège, 7—14 septembre, 1965.
9. Anthony J., Lawrence W., A resonance analogue speech synthesizer, Fourth International Congress on Acoustics, Copenhagen, 21—28 August, 1962.
10. Arnold G. F., Denes P., Gimson A. C., O'Connor J. D., Trim J. L. M., The synthesis of english vowels. Language and speech, 1, part 2, April—June (1958).
11. Ayers E. W., New ideas in the transmission of speech, Post Office Telecommunication Journal, 8, No. 4 (1956).
12. Vadoyanis G., Molony D., Fich S., Slade J. J., Electronically tunable selective amplifier for formant synthesis, JASA, 28, No. 4 (1956).
13. Berg van der J., Imitation of Dutch vowels and words by a hemilaryngectomized subject using a throat loudspeaker as a pseudolarynx, JASA, 27, No. 1 (1955).
14. Borst J. M., Use of spectrograms for speech analysis and synthesis. J. Audio Engng Soc., 4, No. 1 (1956).
15. Carpenter A., Discrimination of vowel sounds synthesized by harmonic addition, Language and Speech, 5, part 3, July-September (1962).
16. Clapper G. L., Speech synthesis system, USA patent No. 3102165, patented 27 August, 1963.
17. Clapper G. L., Synthetic Speech Sound Generator, USA patent No. 3146309, patented 25 August, 1964.
18. Clapper G. L., Synthetic speech sound generator, JASA, 38, No. 4 (1965).
19. Cohen A., Hart J., Speech synthesis of steady-state segments, Proceedings of the Speech Communication Seminar, Stockholm, II, August 29—September 1, 1962.
20. Coker C. H., Cummiskey P., On-line computer control of a formant synthesizer, JASA, 38, No. 5 (1965).
21. Coker C. H., Fujimura O., Model for specification of the vocal tract area function, JASA, 40, No. 5 (1966).
22. Computer mimics vocal tract, Canad. Electron. Engng, 11, No. 1 (1967).
23. Cooper F. S., Spectrum analysis, JASA, 22, No. 6 (1950).
24. Cooper F. S., Speech Synthesizers, Proceedings of the Fourth International Congress of Phonetic Sciences, Helsinki, September 4—9, 1961.
25. Cooper F. S., Instrumental methods for research in phonetics, Proceedings of the Fifth International Congress of Phonetic Sciences, Münster, August 16—22, 1967.
26. Cooper F. S., Delattre P. C., Liberman A. M., Borst J. M., Gerstman L. J., Some experiments on the perception of synthetic speech sounds, JASA, 24, No. 6 (1952).
27. Cooper F. S., Liberman A. M., Borst J. M., Preliminary studies of speech produced by a pattern playback, JASA, 22, October (1950).
28. Cooper F. S., Liberman A. M., Borst J. M., Synthetic speech: A study of the auditory perception of complex sounds, Science, 112, October 13 (1950).
29. Cooper F. S., Liberman A. M., Lisker L., Gaitenby J. H., Speech synthesis by rules, Proceedings of the Speech Communication Seminar, Stockholm, II, August 29—September 1, 1962.
30. DeClerk J. L., Phylfe D. L., Fisch R. A., Electronically controlled formant synthesizer, NEREM Record, 4 (1962).
31. DeClerk J. L., Phylfe D. L., Fisch R. A., Formant synthesizer electronically controlled, Proceedings of the Speech Communication Seminar, Stockholm, II, August 29—September 1, 1962.

32. Delattre P., The use of the pattern playback in studies of vowel color by synthesis, *JASA*, 22, October (1950).
33. Delattre P. C., Voyelles synthétiques à deux formants et voyelles cardinales, *Maître Phonétique*, 96, July—December (1951).
34. Delattre P., Les attributs acoustiques de la nasalité, *Studia Linguistica*, 8, n° 2 (1954).
35. Delattre P., Liberman A. M., Cooper F. S., Speech Synthesis as a Research Technique, Proceedings of the VI International Congress of Linguistics, London, 1952.
36. Delattre P., Liberman A., Cooper F., Acoustic loci and transitional cues for consonants, *JASA*, 27, No. 4 (1955).
37. Delattre P., Liberman A. M., Cooper F. S., Gerstman L. J., An experimental study of the acoustic determinants of vowel colour, *Word*, 8, No. 3 (1952).
38. Denes P. B., On line computing in speech research, 5-e Congrès International d'Acoustique Liège, 7—14 sept., 1965.
39. Denes P. B., The elusive process of speech, *Bell Labs Rec.*, 44, No. 8 (1966).
40. Dennis J. B., Computer control of an analog vocal tract., Proceedings of the Speech Communication Seminar, Stockholm, 11, August 29—September 1, 1962.
41. Dennis J. B., Control language for experimental speech synthesis, *JASA*, 36, No. 5 (1964).
42. Dersch W. C., A voice analog, *JASA*, 31, No. 11 (1959).
43. Digital computer synthesizes human speech, *Bell Labs Rec.*, 40, No. 6 (1962).
44. Dudley H., Remaking speech, *JASA*, 11, No. 2 (1939).
45. Dudley H. W., Synthetic speech, *Bell System Techn. J.*, No. 19 (1940).
46. Dudley H., Fundamentals of speech synthesis, *J. Audio Engng Soc.*, 3, No. 4 (1955).
47. Dudley H., Phonetic pattern recognition vocoder for narrow-band speech transmission, *JASA*, 30, No. 8 (1958).
48. Dudley H., Riesz R. R., Watkins S. S. A., A synthetic speaker, *J. Franklin Inst.*, 227, 738 (1939).
49. Dudley H., Tarnoczy T. H., The speaking machine of Wolfgang von Kempelen, *JASA*, 22, March (1950).
50. Dunn H. K., The calculation of vowel resonances and an electrical vocal tract, *JASA*, 22, No. 11 (1950).
51. Durand M., De la formation des voyelles nasales, *Studia Linguistica*, 7, n° 1 (1953).
52. Durand M., La perception des consonnes occlusives, *Studia Linguistica*, Année VIII, n° 2 (1954).
53. Durand M., De la perception des consonnes occlusives, questions de sonorité, *Word*, 12, No. 1 (1956).
54. Esters S. E., Kerby H. R., Maxey H. D., Walker R. M., Speech synthesis from stored data, *JASA*, 34, No. 12 (1962).
55. Fant G., Modern instruments and methods for acoustic studies of speech. The Royal Institute of Technology, The Speech Transmission Laboratory, Report, No. 8, June 11 (1957).
56. Fant G., Modern instruments and methods for acoustic studies of speech, *Acta Polytechn. Scand. Phys. Including Nucleonics Series*, No. 1 (1958).
57. Fant C. G. M., Modern instruments and methods for acoustic studies of speech, Proceedings of the VIII International Congress of Linguistics, Oslo, 1958.
58. Fant C. G. M., Acoustic analysis and synthesis of speech with application to Swedish, *Ericsson Technics*, 15, No. 1 (1959).
59. Fant C. G. M., Speech analysis and synthesis, Royal Institute of Technology, Speech Transmission Laboratory, Report, No. 26, June (1962).
60. Fant C. G. M., Formants and cavities, Proceedings of the Fifth International Congress of Phonetic Sciences, Münster, August 16—22, 1964.
61. Fant G., Mårthony J., Rengman U., Risberg A., OVE-II synthesis strategy, Proceedings of the Speech Communication Seminar, Stockholm, 11, August 29—September 1, 1962.
62. Fischer-Jørgensen E., What can the new techniques of acoustic phonetics contribute to linguistics? Proceedings of VIII International Congress of Linguistics, 1, Oslo, August 5—9, 1957.
63. Flanagan J. L., Note on the design of "terminal-analog" speech synthesizers, *JASA*, 29, No. 2 (1957).
64. Flanagan J. L., *Speech Analysis, Synthesis and Perception*, — Springer Verlag, New York, 1964.
65. Flanagan J. L., Recent Studies in Speech Research at Bell Telephone Laboratories (II), 5-e Congrès International d'Acoustique, Liège, septembre 7—14, 1965.
66. Flanagan J. L., Coker C. H., Computer simulation of a formant-vocoder synthesizer, *JASA*, 34, No. 12 (1962).
67. Flanagan J. L., Saslow M. G., Pitch discrimination for synthetic vowels, *JASA*, 30, No. 5 (1958).
68. Fourcin A., A potential dividing function generator for the control of speech synthesis, *JASA*, 32, No. 11 (1960).
69. Fry D. B., Experiments in the perception of stress, *Language and Speech*, 1, Part 2, April—June (1958).
70. Fry D. B., Abramson A. S., Eimas P. D., Liberman A. M., The identification and discrimination of synthetic vowels, *Language and Speech*, 5, No. 3 (1962).
71. Fujimura O., Sound synthesizer with optical control, *JASA*, 30, No. 1 (1958).
72. Fujimura O., Nakata K., Speech synthesis by rule, *J. Acoust. Soc. Japan (JASJ)*, 22, No. 1 (1966).
73. Fujimura O., Ogawa T., Hiki S., On the speech sound synthesizer by use of ADP piezo-optic resonators, *JASJ*, 14, No. 2 (1958).
74. Fukumura T., Some basic properties of quality response to the synthesized complex sounds having formant-like spectra, *JASJ*, 14, No. 2 (1958).
75. Halle M., *Sound Pattern of Russian*, s'Gravenhage, 1959.
76. Halle M., Stevens K., Speech recognition: A model and a program for research, *IRE Trans. Inform. Theory*, IT-8, February (1962).
77. Harris C. M., A speech synthesizer, *JASA*, 25, No. 5 (1953).
78. Harris C. M., A study of the building blocks in speech, *JASA*, 25, No. 5 (1953).
79. Harris K., Hoffman H., Liberman A., Delattre P., Cooper F., Effect of third-formant transition on the perception of the voiced stop consonants, *JASA*, 30, No. 2 (1958).
80. Harris K. S., Liberman A. M., Hoffman H., Griffith B. C., Differential sensitivity to synthetic speech sounds within and between phoneme categories, *JASA*, 28, No. 4 (1956).
81. Hasimoto S., Matsumoto M., Synthesis of vowels by a digital computer, *JASJ*, 20, No. 6 (1964).

82. Hearne H., Electronic production of percussive sounds, *J. Audio Engng Soc.*, 9, No. 4, October 1961.
83. Hecker M. H. L., Studies of nasal consonants with an articulatory speech synthesizer, *JASA*, 34, No. 2 (1962).
84. Heinz J. M., Analysis of speech spectra in terms of a model of articulation, *JASA*, 35, No. 7 (1963).
85. Henke W. L., Dynamic articulatory model of speech production, *JASA*, 40, No. 5 (1966).
86. Hillix W., Use of two nonacoustic measures in computer recognition of spoken digits, *JASA*, 35, No. 12 (1963).
87. Hirasaki T., Date H., Computer controlled terminal analog speech synthesizer, *JASJ*, 22, No. 6 (1966).
88. Hirsch I. J., Reynolds E. G., The recognition of synthetic and natural vowels, *JASA*, 25, No. 4 (1953).
89. Holmes J. N., Mattingly J. G., Shearme J. N., Speech synthesis by rule, *Language and Speech*, 7, Part 3, July—September (1964).
90. Holmes J. N., Shearme J. N., Speech synthesis by rule controlled by a small low-speed digital computer, *JASA*, 35, No. 11 (1963).
91. Honda T., Ogawa Y., Speech synthesis by means of hybrid computer system, *JASJ*, 23, No. 1 (1967).
92. House A. S., Analog studies of nasal consonants, *J. of Speech and Hearing Disorders*, 22, No. 2 (1957).
93. House A. S., Fairbanks G., The influence of consonant environment upon the secondary acoustical characteristics of vowels, *JASA*, 25, No. 1 (1953).
94. House A. S., Stevens K. N., Analog studies of nasalization of vowels, *J. of Speech and Hearing Disorders*, 21, 218 (1956).
95. Ichikawa A., Nakano Y., Nakata K., Control rule of vocal-tract configuration, *JASA*, 42, No. 5 (1967).
96. Ingemann F., Speech synthesis by rule, *JASA*, 29, No. 12 (1957).
97. Ingemann F., Eight-parameter speech synthesis, *JASA*, 32, No. 11 (1960).
98. Ingemann F., Formants as a voicing cue in the perception of [z], *JASA*, 32, No. 11 (1960).
99. Johnson W., System to generate speech from written pattern shown, *Electronic News*, 8, No. 392, August 26 (1963).
100. Kacprowski J., An approach to the synthesis of Polish nasal consonants by means of the "terminal analog" speech synthesizer, *Proceedings of Vibration Problems*, 4, No. 3 (1963).
101. Kacprowski J., Formant synthesis of Polish vowels and nasal consonants (theoretical fundamentals), *Arch. Elektrotechnik*, 13, No. 3 (1964).
102. Kacprowski J., Mikiel W., Synthesis of Polish C-V syllables by means of the terminal-analog speech, IV International Congress of Acoustics, Copenhagen, 1962.
103. Kacprowski J., Mikiel W., Preliminary synthesis of Polish vowels by means of recurrently impulsive formant filters, *Proc. Vibration Problems*, 4, No. 1 (1963).
104. Kacprowski J., Mikiel W., Synthesis of Polish C-V syllables by means of the terminal-analog speech synthesizer, 5-e Congrès International d'Acoustique, Liège, Septembre 7—14, 1965.
105. Kacprowski J., Mikiel W., Simplified rules for parametric synthesis of nasal and stop consonants in C-V syllables by means of the "terminal analog" speech synthesizer, *Acustica*, 16, No. 6 (1965—1966).
106. Kelly J. L., Gerstman L. J., An artificial talker driven from a phonetic input, *JASA*, 33, No. 6, 1961.
107. Kelly J. L., Lochbaum C., Speech synthesis, *Proceedings of the speech communication seminar, Stockholm*, II, August 29—September 1, 1962.
108. Kerby H. R., Enke G. G., Introduction of fundamental frequency and timing information into speech synthesized from stored data, *JASA*, 35, No. 11 (1963).
109. Kerby H. R., Maxey H. D., Walker R. M., Speech synthesis from stored data, *JASA*, 34, No. 12 (1962).
110. Knauft G., Lamparter H., Spruth W. G., Some new methods for digital encoding of voice signals and for voice code translation, *IBM J. Research and Development*, 10, No. 3 (1966).
111. Kùpfmùller K., Warns O., Sprachsynthese aus Lauten, *Nachrichtentechnische Fachberichte*, 3, January (1956).
112. Ladefoged P., Some possibilities in speech synthesis, *Language and Speech*, 7, part 4, October—December (1964).
113. Laver J. D., Lawrence W., Synthesis of components of voice quality, *JASA*, 37, No. 6 (1965).
114. Lavington S. H., Rosenthal L. E., Some facilities for speech processing by computer, *Comput. J.*, No. 4 (1967).
115. Lawrence W., The synthesis of speech from signals which has a low information rate, *Communication Theory*, ed. W. Jackson, 1953.
116. Lawrence W., Parametric artificial talking device, *JASA*, 32, No. 11 (1960).
117. Lawrence W., Role of synthetic speech in speech research, *JASA*, 36, No. 5 (1964).
118. Lawrence R., Digital formant synthesizer for speech synthesis studies, *JASA*, 42, No. 5 (1967).
119. Liberman A. M., Delattre P. C., Gerstman L. J., Cooper F. S., Tempo of frequency change as a cue for distinguishing classes of speech sounds, *J. Experim. Psychology*, 52, No. 2 (1956).
120. Liberman A. M., Ingemann F., Lisker L., Cooper F. S., Minimal rules for synthesizing speech, *JASA*, 31, No. 1 (1959).
121. Liberman A. M., Ingemann F., Lisker L., Delattre P., Cooper F. S., Minimal rules for synthesizing speech, *JASA*, 31, No. 11 (1959).
122. Lindgren N., Machine recognition of human language, Part II. Theoretical models of speech perception and language, *IEEE Spectrum*, April (1965).
123. Lisker L., Minimal cues for separating [w, r, l, y] in intervocalic position, *Word*, 13, No. 2 (1957).
124. Lisker L., Linguistic segments, acoustic segments, and synthetic speech, *Linguistic Society of America, Language*, 33, July—September (1957).
125. Mabuchi K., Uagi M., Endo K., Oizumi J., On the study of the regions of voiceless consonants by synthesizing, *JASJ*, 14, June (1958).
126. Manley H. J., Analysis-synthesis of connected speech in terms of orthogonalized exponentially damped sinusoids, *JASA*, 35, No. 4 (1963).
127. Mártony J., Cederlund C., Liljencrants J., Lindblom B., On the analysis and synthesis of vowels and fricatives, *Proceedings of the Fourth International Congress of Phonetic Sciences, Helsinki* September 4—9, 1961.
128. Mártony J., Fant G., Information bearing aspects of formant amplitude, *Proceedings of 5th International Congress of Phonetic Science, Münster*, 1964.



129. Matsui E., Suzuki T., Umeda N., Omura H., On the programming system for synthesis of Japanese alphabet, JASJ, 22, No. 6 (1966).
130. Mattingly I. G., Speech synthesis by rule as a research technique, JASA, 41, No. 6 (1967).
131. Maxey H. D., Terminal-analog synthesis of voiced fricatives, JASA, 35, No. 11 (1963).
132. Meocks W. W., Borst J. M., Cooper F. S., Syllable synthesizer for research on speech, JASA, 26, No. 1 (1954).
133. Meyer-Eppler W., Synthetische Sprache, Physikalische Zeitschrift, 29, April (1950).
134. Miller R. L., Auditory tests with synthetic vowels, JASA, 25, No. 1 (1953).
135. Mundie J. R., Bionics approach to speech segmentation and synthesis, JASA, 38, No. 5 (1965).
136. Nakata K., Synthesis of nasal consonants by a terminal analog synthesizer, J. Inst. Electr. Communic. Engrs Jap., 42, No. 5 (1959).
137. Nakata K., Synthesis and perception of Japanese fricative sounds, J. Radio Res. Labs, 7, July (1960).
138. Nakata K., Kadokawa Y., Synthesis and perception of /h/ and /k/ sounds, J. Radio Res. Labs, 7, No. 34 (1960).
139. Nakata K., Mitsuoka T., Phonemic transformation and control aspects of synthesis of connected speech, J. Radio Res. Labs, 12, No. 61, May (1965).
140. Nakata K., Obara K., Vowel synthesizer, USA patent No. 3087989, patented 30 April, 1963.
141. Nakata K., Suzuki J., Terminal analog speech synthesizer, J. Inst. Electr. Communic. Engrs of Jap., 42, September (1959).
142. Nakata K., Suzuki J., Synthesis and perception of Japanese vowels and vowel-like sounds, J. Radio Res. Labs (Tokyo), 6, No. 28 (1959).
143. Nakayama T., Ichikawa A., Nakata K., Control rule of voice source to improve the naturalness of synthetic speech, JASA, 42, No. 5 (1967).
144. Oizumi J., Kubo E., Synthesizing speech, JASJ, 10, September (1954).
145. Olson H. F., Speech processing system, IEEE Spectrum, February (1964).
146. Olson H. F., Belar H., Sobrino R., Demonstration of a speech processing system consisting of a speech analyzer, translator, typer and synthesizer, JASA, 34, No. 10 (1962).
147. Peterson G. E., Fundamental problems in speech analysis and synthesis, Proceedings of VIII International Congress of Linguistics, Oslo, August 5-9, II, 1957.
148. Peterson G. E., Sivertsen E., Objectives and techniques of speech synthesis, Language and Speech, 3, part 2, April-June (1960).
149. Peterson G. E., Wang W. S.-Y., Sivertsen E., Segmentation techniques in speech synthesis, JASA, 30, July (1958).
150. Pickett J. M., Some acoustic cues for synthesis of the [n-d] distinction, JASA, 38, No. 3 (1965).
151. Pimonov L., La parole synthétique et son application dans la correction auditive, Annales des Télécommunication, Cahiers d'Acoustiques, n° 134, 135 (1965).
152. Poppe C. W., Suhr P. Y., An automatic voice readout system, Proceedings of Eastern Joint Computer Conference, 1957.
153. Rabiner L. R., Speech synthesis by state simulation, JASA, 40, No. 5 (1966).
154. Redlich von J., Wendahl R. W., On the relative loudness of synthesized vowels, JASA, 32, No. 7 (1960).
155. Rindfleisch H., Ein einfaches Verfahren zur Erzeugung von synthetischer Sprache. Technische Hausmitteilung des Nordwestdeutschen Rundfunks, 3 (1951).
156. Rischel J., Instrumentation for vowel synthesis, Annual Report of the Institute of Phonetics University of Copenhagen, 1957.
157. Rosen G., Dynamic analog speech synthesizer, JASA, 30, No. 2 (1958).
158. Rosen G., Fricative production by a dynamic analog synthesizer, JASA, 31, No. 11 (1959).
159. Rosen G., Dynamic analog speech synthesizer, USA patent No. 3042748, patented 3 July, 1962.
160. Rosen G., Stevens K. N., Apparatus for the generation of psycho-acoustic stimuli varying in complexity from noise burst to syllables, JASA, 31, No. 1 (1959).
161. Rosen G., Stevens K. N., Heinz J. M., Dynamic analog of the vocal tract, JASA, 28, No. 4 (1956).
162. Schott L. O., A playback for visible speech, Bell Telephone Lab. Record, 26, August (1948).
163. Schott L. O., An electrical vocal system, Bell Labs Record, December (1950).
164. Scott R. S., Grace D. A., Mattingly J. G., A computer-controlled on-line speech synthesizer system, IEEE Internat. Communic. Conference, New York, 2 (1966).
165. Sholes G. N., Synthesis of final (z) without voicing, JASA, 31, No. 11 (1959).
166. Sivertsen E., Segment Inventories for speech synthesis, Language and Speech, 4, part I (1961).
167. Slawson A. W., MUSE — a sound synthesizer, Proceedings of IFIP Congress, Munich, August 27-September 1, 1962.
168. Slawson A. W., Valentine J. W., Kish A., Cheap talk from a computer, JASA, 35, No. 5 (1963).
169. Smith C. P., Vocal response synthesizer, JASA, 37, No. 1 (1965).
170. Speaks V., Jerger J., Jerger S., Performance-intensity characteristics of synthetic sentences, J. Speech and Hearing Res. 9, No. 2 (1966).
171. Stenby W. H., Speech-controlled apparatus and method for operating speech-controlled apparatus, USA patent No. 3215821, patented 2 November, 1965.
172. Stevens K. N., Synthesis of speech by electrical analog devices, J. Audio Engng Soc., 2, No. 1 (1956).
173. Stevens K. N., Bastide R. P., Electrical synthesizer of continuous speech, JASA, 27, No. 1 (1955).
174. Stevens K. N., House A. S., Development of a quantitative description of vowel articulation, JASA, 27, No. 3 (1955).
175. Stevens K. N., House A. S., Studies of formant transition using a vocal tract analog, JASA, 28, No. 4 (1956).
176. Stevens K. N., Kasovski S., Fant G. C., An electrical analog of the vocal tract, JASA, 25, No. 4 (1953).
177. Stewart J. Q., An electrical analogue of the vocal organs, Nature, 110, 311 (1922).
178. Stowe A. N., Hamton D. B., Speech synthesis with prerecorded syllables and word, JASA, 33, No. 6 (1961).
179. Strong W. J., Machine-aided formant determination for speech synthesis, JASA, 40, No. 5 (1966).
180. Sussex J. R., Dennis J. B., A computer program for controlling the dynamic vocal tract analog (DAVO), Quarterly Report Research Laboratory of Electronics, Massachusetts Institute of Technology, No. 68 (1963).
181. Szési L., Künstliche Erzeugung von Vokalen, Frequenz, 15, No. 8 (1961).
182. Tarnóczy T. H., Vowel formant bandwidths and synthetic vowels, JASA, 34, No. 6 (1962).
183. Tomlinson R. S., SPASS — an improved terminal-analog speech synthesizer, JASA, 38, No. 5 (1965).

184. Tschescher W., Über ein Steuersystem zur Sprachsynthese. 5-e Congrès International d'Acoustique, Liège, Septembre 7-14, 1965.
185. Tscheschner W., Fuchs S., Ein Steuersystem zur Sprachsynthese, Nachrichtentechnik, 16, No. 3 (1966).
186. Uldall E. T., Transitions in fricative noise, JASA, 35, No. 11 (1963).
187. Uldall E. T., Transition in fricative noise, Language and Speech, 7, Part 1, January-March (1964).
188. Uldall E., Anthony J. K., The synthesis of a long piece of connected speech on PAT. Proceedings of the Speech Communication Seminar, Stockholm, 11, August 29-September 1, 1962.
189. Umeda N., Teranishi R., Phonemic feature and vocal features synthesis of speech sounds, using an acoustic model of tract, JASJ, 22, No. 4 (1966).
190. Urquhart A. B., Voice Output from IBM System/360, AFJPS Conference Proc., 27, Part I, 1965.
191. Vocal analog synthesized speech, Electronics, 36, No. 41 (1963).
192. Wang W. S.-Y., Peterson G. E., Segment inventory for speech synthesis, JASA, 30, No. 8 (1958).
193. Warns O., The intelligibility of synthetic speech, Frequenz, 11, June (1957).
194. Weibel E. S., Vowel synthesis using resonant circuits to represent formants, JASA, 26, September (1954).
195. Weibel E. S., Vowel synthesis by means of resonant circuits, JASA, 27, No. 5 (1955).
196. Wendhal R. W., Case A. L., Stream R. W., Effect of duration upon the intelligibility of synthetic vowels, JASA, 32, No. 7 (1960).
197. Winckel F., Elektrische Abbildung von Sprachlauten, Electron, 5 (1951).
198. Winckel F., Die menschliche Stimme in elektrischen Modellvorstellungen. Elektrotechnische Zeitschrift, Ausgabe A, 73, H. 22, 11. November (1952).
199. Winckel F., Analyse und Synthese der Sprachlaute, Funk und Ton, 8, Nr. 1, 2 (1954).
200. Woolans D. J., Gill A. M. R., A simple speech synthesizer, Electronic Technology, 37, No. 10 (1960).
201. Zemanek H., Geschichte des Automaten: Mechanische Automaten des 18. Jahrhunderts von Wolfgang von Kempelen, Elektronische Rechenanlagen, Jg. 8, H. 1 (1966).

*Институт кибернетики  
Академии наук Эстонской ССР*

Поступила в редакцию  
12/X 1968

E. KÜNNAP

### KÕNE SÜNTEES

Kõne süntesaatorit saab kasutada kõne standardse generaatorina lingvistiliste ja kõnesignaalide automaatse äratundmise probleemide uurimisel, sidekanalite tihendamisel ja väljundseadmena ühendustes «inimene-masin». Artiklis tutvustatakse kõne sünteesi alal valminud uurimistööid ja nende tulemusi.

E. KÜNNAP

### SPEECH SYNTHESIS

The speech synthesizer can be used as a standard sound generator for linguistic studies and for solving the problem of automatic recognition of speech sounds, in a band-width compression system, as a computer output device for "man-machine" communication. The article presents a review of researches in the sphere of speech synthesis and of some results obtained.