

Основы Фонетической теории речи. Фонетическая функция как универсальный природный инструмент кодирования (декодирования) речевой информации любого происхождения

Пирогов А.А.

«Сезам», открывающий дверь, и большие потери в отечественной электронной информатике.

Да, – с незапамятных времён люди мечтали об управлении машинами с помощью голоса. Помните? – «Он набрался храбрости, вдохнул больше воздуха и во весь голос крикнул: – «Сезам, открой дверь!» – И тотчас же дверь раскрылась ...» [1].

1. Что мы слышим, когда распознаём речевые сигналы?

Впервые ответ на этот вопрос был опубликован в 1963 г. – страшно подумать – 1/3 века тому назад ... В монографии [2,с.16] было дано категорическое отрицание «... широко распространённому, неправильному отождествлению спектральных распределений («формантных картинок») и фонетических значений отдельных звуков». Более того, было подчеркнуто: «... можно... уверенно сказать, что каждая фонема отличается характерным для этой фонемы **изменением** спектральных распределений, а не самим спектральным распределением, сопутствующим данной фонеме» (там же, с.19), на совещании специалистов по приборному (инструментальному) преобразованию речи (СПб., 1966г.) было решительно заявлено, что **не существует** такого звука, который бы однозначно воспринимался как фонема – например - (а), см. [3, с.386].

Сказанное поясняется графиками, рис. 1а, рис. 1б, спектральных плотностей $S(f)$ гласных звуков [3, с.12, рис. 1.1в] и амплитудно-частотных характеристик (АЧХ) по звуковому давлению $P_a(f)$ типовых телефонов ТК-47 и ДМК-6А, ТА-4 [4, с.163, рис. 6.11] . Достаточно одного взгляда на эти АЧХ, чтобы понять, что в любом телефонном и просто – речевом канале, спектральные распределения $S(f)$ так искажаются, что ни о каком «соответствии» этих распределений «фонемам» О, А и т.д. говорить не приходится. Так, например, даже при разговоре через стол, когда прямые и отражённые от стола с запаздыванием на 0,6...1 мс сигналы достигают нашего слуха, что на частотах кратных 800...500 Гц определяет полное замирание суммы прямого и отражённого сигналов,- а мы вообще в этих условиях не замечаем никаких искажений!

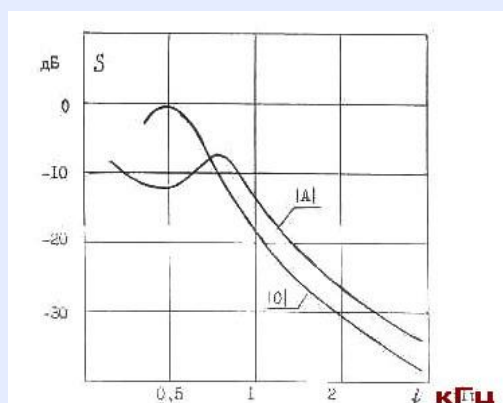


Рис. 1а. График спектральных плотностей $S(f)$ гласных звуков «а» и «о», [3].

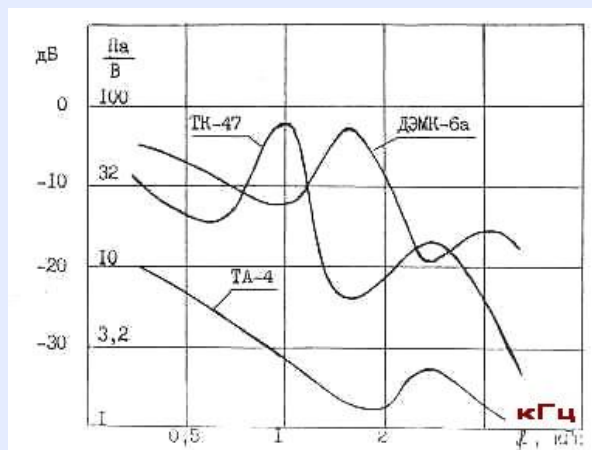


Рис. 1.6. График амплитудно-частотных характеристик (АЧХ) по звуковому давлению $P_a(f)$ типовых отечественных телефонов ТК-47, ДМК-6а, ТА-4, [4].

Здесь необходимо также напомнить классический эксперимент, выполненный по нашему предложению В.И. Кулей и его коллегами [3, с. 376, рис. 8.1]: запись слова «анализатор» была произведена с искажением частотной характеристики по закону $K(f) = S_u(f)/S_a(f)$, после чего в полученную запись вместо фонемы «а» была вставлена **натуральная** фонема «и», которая при этом звучала ещё более отчётливо как ... фонема «а»!

Очень важно: наш слух практически абсолютно **инвариантен** в отношении амплитудно-частотных, фазо-частотных и, в большей степени, – нелинейных искажений, если, конечно, эти искажения не выходят далеко за пределы артикуляторных модуляций и пределов слухового восприятия.

Естественно, возникает **основной, фундаментальный** вопрос: что же мы слышим, что воспринимаем как сигнал, определяющий процесс декодирования речевой информации?

Другими словами: **что конкретно (!)** нужно передавать и принимать, по возможности – без искажений?

2. Ответ на этот – повторяем – **основополагающий** вопрос был дан в 1963 году в нашем докладе Совету по проблеме «Кибернетика» при Президиуме Академии Наук СССР [5].

В качестве сигнала, определяющего процесс декодирования речевой информации, мы воспринимаем открытую нами **«фонетическую функцию»**:

$$P(\omega, t) = \lg (S(\omega, t) / S(\omega, t-\tau)) . \quad (1)$$

Здесь S – спектральная плотность речевого сигнала в моменты времени (t) и $(t - \tau)$ на определяющей (средней) частоте ω соответствующей равноартикуляционной частотной полосе. В стандартной частотной полосе канала ТЧ (тональных частот) $\Pi_{тч} = 0,3...3,4$ кГц. При проектировании синтезаторов речи вокодеров достаточное число таких полос обычно выбирается $N \sim 10$. Интервал времени τ , определяющий соседние спектральные разрезы, обычно составляет: $\tau = 20$ мс, то есть: $B = 50$, отсчётов в секунду.

Заметим, что пресловутая и, зачастую, бездумно используемая теорема Найквиста – Котельникова – Фурье, согласно которой число отсчётов в секунду должно быть $B \geq 2 \cdot F_{\text{высш.}}$, здесь абсолютно не «робастна»: для чёткой передачи «взрывных» согласных время «коммутации спектров» речи должно быть очень кратким, $\tau_k \approx 1$ мс. И поэтому высшую частоту фильтров на выходе анализатора спектра выбирают $F_{\text{высш.}} \approx 70...80$ Гц.

Соответственно, по «теореме» число отсчётов должно быть $V \sim 150$ отсч./сек. Практически применяют $V = 45 \dots 50$ отсч./сек. – V три раза меньше!

Дело в том, что высокие частоты $F_{\text{выш.}}$ в сигналах управления синтезом речи существенно важны для того, чтобы не «размазать» во времени, сохранить крутой фронт взрывного согласного звука, а когда (раньше или позже) этот звук появится, – для нашего слуха, для раскодирования семантических признаков речевого кода, – это совершенно безразлично.

Вернёмся, однако, к вопросу о том, – что мы слышим?

Как следует из (1), общий эффект восприятия ощущений в предположении, что прежние ощущения затухают по экспоненциальному закону [5,6,7].

$$P(\omega, t) = c \cdot \int e^{-\tau/T} \cdot \lg (S(\omega, t) / S(\omega, t-\tau)) \cdot d\tau, \quad (2)$$

где: T – постоянная времени нашего слухового аппарата, пределы интегрирования: $0 \dots \infty$. Функциональная схема соответствующего устройства представлена на рис.2.

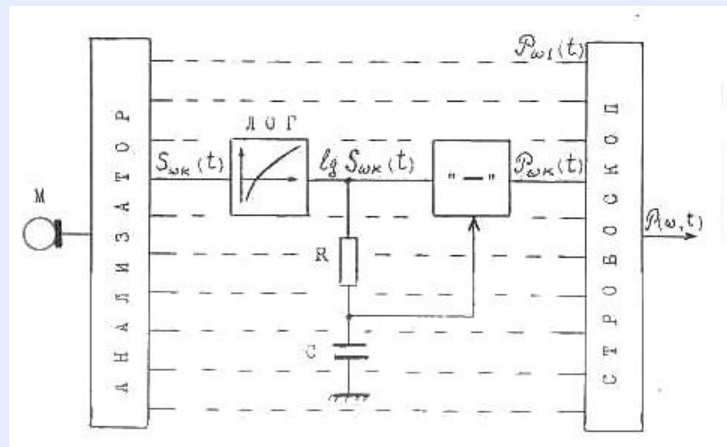


Рис. 2. Функциональная схема устройства, формирующего из речевого сигнала фонетическую функцию $P(\omega, t)$.

Здесь речевой сигнал с микрофона M поступает на спектральный АНАЛИЗАТОР, осуществляющий определение спектральных плотностей $S_{\omega_k}(t)$ сигнала по $k = 1 \dots 10$ равноартикуляционным частотным полосам ω_k . Каждый спектральный отсчёт поступает на логарифмирующее устройство нелинейного преобразования ЛОГ, с выхода которого сигнал $\lg S_{\omega_k}(t)$ подаётся на интегрирующую цепочку R и C с постоянной времени $T = R \cdot C$ (слухового восприятия) и, параллельно, – на вычитающее устройство «-», на выходе которого получается разница:

$$P_{\omega_k}(t) = \lg S_{\omega_k}(t) - \int e^{-\tau/T} \cdot \lg (S(\omega, t) / S(\omega, t-\tau)) \cdot d\tau.$$

Далее все «к» парциальных сигналов поступают на стробоскоп (концентратор) для последовательного группирования информации в двухразмерную (по частоте ω и времени t) фонетическую функцию $P(\omega, t)$ речевого сигнала.

Это и есть то, что мы слышим.

Измените, как хотите (!), частотную характеристику тракта, умножив числитель и знаменатель на произвольный коэффициент $\gamma(f)$ в соотношении (2), – результат не изменится (если $\gamma \neq 0$)!

3. Семантический код речи – корректирующий код с двойной избыточностью

При обработке по закону фонетической функции $P(\omega, t)$ речь автоматически «рассыпается» на фонетические элементы (элементы семантического кода):

– идёт гласный звук, длительностью $T_r > T = RC$, – на выходе устройства вычитания «-» нуль, разрыв функции $P(\omega, t)$;

– идёт согласная «м» – яма – йама – кратчайшая (1...2 мс) пауза – **знак прохождения взрывной согласной (м)**; в момент начала паузы с выхода устройства вычитания «-» – отрицательный разряд, затем снова перезаряд с появлением конечной гласной фонемы «а».

Через 12 с лишним лет после нашей первой публикации [5], а также (после некоторых тяжёлых противостояний) через 8 лет после публикации в журнале «Электросвязь» [6] наша новая **фундаментальная фонетическая теория речи** в точном соответствии с указанными её определениями была блестяще продемонстрирована доктором Ленаром из ф. Лимси, Франция [8].

Ещё через 2 года, через 14 лет после нашей первой публикации [5], было объявлено [9], что французский язык разложен на алфавит семантического кода, включающего 627 элементов – звукосочетаний («дифонов»).

Определяющее значение звукосочетаний следует непосредственно из формулировки фонетической функции, **реагирующей** на **изменение** спектральных распределений. Особенно ярко, экспериментально, это было установлено нами, – В.Н. Куля [10], ещё в 1967 году, на 10 лет (!) раньше, чем появилась французская публикация [9], в упомянутых выше опытах **подмены** одной гласной (а) – «натуральной» другой гласной (и) в условиях **сохранения** фонетической функции звукосочетаний. Иными словами, «...это объясняется замечательным свойством **фонетического корректирующего кода речи**, которое состоит в том, что анализ речи ... ведётся не по фонемам, а по звукосочетаниям...» [3, с. 390]. То есть, – на основе корректирующего кода с двойной избыточностью.

4. Распознавание гласных и согласных – двухканальная система распознавания речи

В соответствии с изложенным, при указанной абсолютной неопределённости в реальных акустических условиях спектральных характеристик гласных звуков, была установлена и опубликована [6, 10, 11, 12] система распознавания гласных звуков по их **спектральным переходам**.

Идея такого способа определения гласных звуков поясняется рис.3.

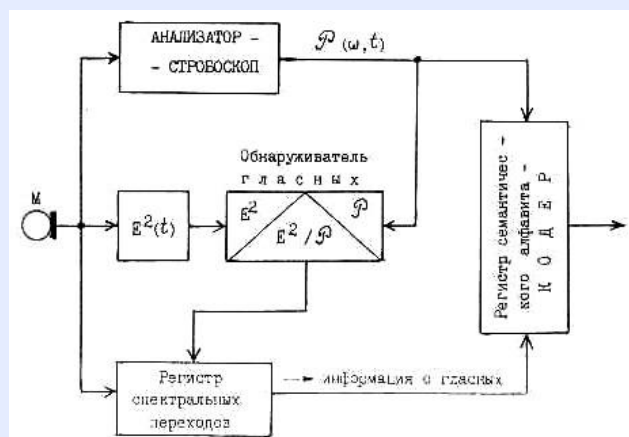


Рис. 3. Функциональная схема устройства распознавания гласных звуков по их спектральным переходам.

Речевой сигнал из телефонного сигнала или из микрофона поступает на анализатор – стробоскоп, вырабатывающий фонетическую функцию $P(\omega, t)$, см. рис. 2. Далее (рис. 3) он подаётся на обнаружитель гласных. Здесь, по соотношению E^2/P в момент замирания функции $P(\omega, t)$, но при наличии детектируемой в устройстве E^2/P сигнала о наличии энергии E , появляется сигнал о прохождении гласной, подаваемый на **регистр спектральных переходов**. Если же энергия функции $P(\omega, t)$ спала (как результат паузы, определяемой прохождением согласной, см. выше), то на регистр спектральных переходов (РСП) никакой сигнал не поступает.

С получением в РСП сигнала о прохождении гласной из памяти регистра извлекается спектр $S_{n-1}(\omega)$ предыдущей и $S_n(\omega)$ текущей гласной, определяется спектральный переход и, при этом, одновременно:

- подтверждается корректность ранее определённой гласной (если это ранее было сделано);
- определяется фонетическое значение текущей гласной.

На рис.4 представлены примеры таких спектральных переходов от гласной фонемы «а» к пяти другим гласным фонемам [3].

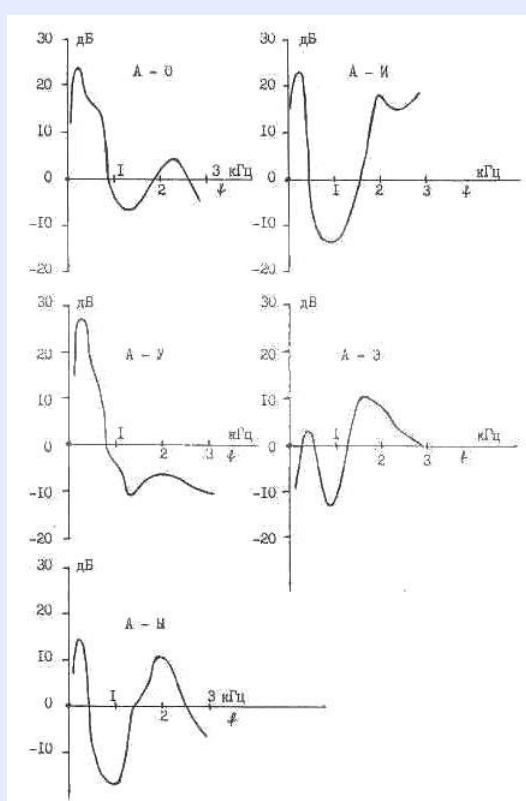


Рис. 4. Графики спектральных переходов гласных (а-о), (а-и), (а-у), (а-э), (а-ы): $20 \lg (S_f(o,и,у,э,ы) / S_f(a))$.

Например, при переходе от «а» к «э» низкочастотные гармоники ОТ в области «первой форманты» (~1000 Гц) сильно подавляются, в области «второй форманты» (~ 2 кГц) – очень усиливаются. Совершенно ясно, что при изменениях частотной характеристики (см. рис.1б) эти переходы, определяющие признаки гласных, несколько не изменяются.

С выхода регистра спектральных переходов информация о гласных подаётся на КОДЕР – **регистр семантического алфавита** для регистрации и выдачи в канал связи соответствующей кодовой комбинации о прошедшем **гласном** звуке.

Подчёркиваем: паузы в передаче фонетической функции $P(\omega, t)$ соответствуют прохождению согласных «фонем». Это характеризуется специфическими модуляциями (коммутацией) фонетической функции, предшествующими и последующими коммутации спектров речи в момент (1...2 мс !) передачи согласного звука. Так вот, – при паузах в передаче фонетической функции

$P(\omega, t)$ – в регистре семантического алфавита, в КОДЕРЕ, отбираются кодовые комбинации, определяющие модуляции спектров речи в моменты «до» и «после» прохождения согласного звука.

Далее. Соответствующая информация о прохождении согласного звука также подаётся на КОДЕР для выдачи в канал связи соответствующей кодовой комбинации о прохождении **согласного** звука (более высокоэнтропийного, чем гласная!).

Нужно отметить, что двухканальная модель опознавания и регистрации гласных звуков была незамедлительно оценена иностранными учёными; в частности – ещё до публикации работ Ленара, доктор Аугустин Дитл опубликовал очень представительную работу «Адаптивное распознавание гласных», с честными ссылками на наши публикации [13]. Нужно сказать, что и к другим нашим работам по созданию нового класса ортогональных вокодеров (Б.Е. Муравьёв, С.П. Баронин, Б.И. Куля, А.Ю. Лев, Ю.К. Трофимов, Н.К. Игнатъев, В.Н. Соболев, В.Г. Ходатай, Ю.С. Мельников и др.) иностранные учёные – J.L. Flanagan, M.R. Schroeder, E.E. David, G. Fant, J.L. Melsa и др. также были очень внимательны к нашим работам, отмечая их в своих монографиях и публикациях и даже добиваясь (в тех условиях!) встречи с нами ...

5. Об «узнаваемости речи»

Изложенная выше фундаментальная фонетическая теория речи, определяющая её семантический код, инвариантный к акустическим характеристикам тракта приёма-передачи речевых сигналов, наряду с огромными практическими возможностями, позволяет:

- реализовать системы управления машинами (ЭВМ) с помощью речевых команд;
- осуществлять речевую связь по телеграфным каналам ёмкостью $B \approx 150$ бит/с [14] (в 430 раз меньше, чем стандарт ИКМ-64, то есть 64000 бит/с.);
- обеспечить надёжнейшим образом защиту от несанкционированного доступа (подслушивания разговоров);
- на порядки уменьшить энергозатраты и повысить надёжность телефонной связи.

В то же время, по мнению некоторых специалистов, фонетическая теория речи обладает, якобы, вот каким «недостатком»: речь, синтезируемая у абонента на приёмной стороне, не характеризуется, якобы, «узнаваемостью» говорящего; синтезатор говорит, якобы, стандартизованным голосом радио диктора.

Хотя такой «недостаток» не вполне обоснован, ибо в условиях телефонного диалога на логических уровнях анализа обычно можно распознать корреспондента, наличие такого «недостатка» не может быть серьёзно принято в расчёт: что же, – тогда и телеграф нужно исключить из сферы телекоммуникаций?

Но если уж по серьёзному говорить об «узнаваемости», то тут дело обстоит, оказывается, **как раз наоборот**: использование **семантического кода** позволяет доказательно (юридически приемлемо) **распознать** говорящего по телефону (по любому обычному телефону), сопоставив произношение (фонетическую функцию – семантический код) опознаваемого субъекта со стандартным филологическим образцом кода для данного языка.

Итак, с «узнаваемостью» также открываются новые, существенно важные, возможности, уже давно понятые и применяемые на практике иностранцами!

6. Выводы

- 1/3 века тому назад всё изложенное выше было известно и доложено советским учёным; 1/4 века тому назад – опубликовано, 1/5 века тому назад подхвачено и использовано иностранцами, создан семантический код французского языка ...

- Начиная с 80-х годов в странах «семёрки» нет проблемы «речевой связи с машинами». Занимаются разработкой «интеллектуальных роботов»: – **искусственный интеллект (ИИ)**, его технология. Применения ИИ: разработка программного обеспечения, планирование производства, составление расписаний для школы и др. [15].

- В ФРГ, фонд SEL.: 1980г. – связь «человек-машина» в пунктах управления; 1981г. – естественный речевой диалог (искусственный интеллект); следующие годы – проблема речевой связи снята [16], после чего исследования были переключены на разработки ИИ-роботов (учрежденческая связь; автоматизация навигации для лоцманов и т.п.).

- В Японии ещё в 1983 г. было объявлено, что «... команды, отдаваемые голосом человека (капитана корабля «Кивокаве Маару»), являлись вполне достаточными для управления большим кораблём» [17]; и к 1990 году уже был «... установлен автомат для продажи железнодорожных билетов /который/ понимает человеческую речь. Он не имеет никаких кнопок, пассажир лишь сообщает ... куда едет; и спустя несколько секунд получает билет» [18].

- В сентябре 1992 года разработан интеллектуальный робот для перевода на три языка (японский, английский и немецкий); после успешной демонстрации этой системы, со словарём только в 400 слов, исследователи Университета Карнеги – Мелона работают над новой системой «переводчика» с расширенным словарём [19].

- Не удивительно, что сейчас уже представлено ПО (программное обеспечение) для распознавания речи со словарным запасом 2000 слов, расшифровывающее сообщение и выполняющее действия, указанные звонящими по телефону [20]. **Очень важно:** «ПО» позволяет создавать свой собственный словарь, **основываясь на фонетике, а не на так называемых «ключевых словах».**

- На той же выставке [20] представлена система *Telephony Recogniser*, распознающая речь **независимо** от говорящего: её не нужно «обучать» под данный голос (то есть в систему введён семантический код стандартного английского языка!); эта система очень ценна не только тем, что «понимает» любого, но **уникально** ценна ещё тем, что способна узнавать говорящего по специфическому «акценту» его речи.

- И совсем впечатляющая информация о том, что Ай-Би-Эм создаёт прототип карманного ПК с процессором Power PC, который имеет «...словарь объёмом 80 тысяч слов английского языка и обрабатывает 65 слов в минуту» ... с достоверностью 96% (!) [21].

Заметим, что разборчивость $W = 0,96$ соответствует почти полной фразовой разборчивости $J \approx 98,5\%$ и полной разборчивости двухзначных цифр $N = 100\%$ [22, с.98, 99, рис. 4.2]. Многие зарубежные ПК сейчас выпускаются со «встроенными аудио средствами».

А как обстоят дела у нас?

Да, – как уже отмечалось выше, наши учёные сделали большие вклады в развитие вокодерной телефонии для связи по дискретным каналам ёмкостью 1200...2400 бит/с.

Создали принципиально новый класс высокоэффективных **ортогональных** [3] вокодеров, выделенных Джеймсом Л. Фланаганом в отдельный раздел (8.6) во втором издании [23] его замечательной монографии. См. также [24,25].

Однако, даже после появления работ J.S. Lienard [8,9] в 1975-77 годах, «советские» учёные к основополагающим определениям фонетической функции – к **семантическому коду речи** продолжали относиться **предвзято отрицательно**.

Мощный сдвиг вызвала замечательная теоретическая и экспериментальная работа В.Н. Соболева [26]. Прошло ещё 4 года ... и, наконец, появились работы Ю.К. Калининца [27,28] с деловым научным анализом фонетического представления речевых сигналов и, одновременно, [30] – другая замечательная работа В.Н. Соболева. Эта работа на опыте практических исследований незыблемо утвердила фундаментальное описание речевых сигналов и показала, что на основе оценки качества передачи по телефонному тракту **фонетической функции** речи «... при длительности речевого испытательного теста равной, примерно, 40 секунд ... величина ошибки соответствует погрешности в определении слоговой разборчивости, не превышающей 1%»! И не нужно никаких «артикуляционных бригад», часами сидящих и слушающих бессмысленные слоговые ГОСТ-овские таблицы.

Но, к сожалению, до машин, «понимающих» речь, а тем более – до «интеллектуальных роботов», нам ещё так далеко ... Хорошо ещё что J.S. Lienard 20 лет назад показал нам, что всё у нас правильно...

Да, – публикуются статьи, тематические сборники и монографии по филологии. Но при этом ключевое, основополагающее **первичное описание** речи – её **фонетическую функцию старательно замалчивают**. Её, также как и понятия семантического корректирующего кода речи, как бы не существует. И основополагающие отечественные фундаментальные работы [2,3,5,6,7,10,11,12,13,25,26,27,28,30 и др.] как бы не существуют!

Приведём в качестве примера одну из таких публикаций – монографию объёмом почти 400 страниц (!) [29]. Здесь есть и «синтез речи» с помощью ИКМ, и Δ -модуляция (запись, а не «синтез») – нет только «синтеза» с помощью грамофонных пластинок! Есть целые страницы математического кружева, посвящённого проблеме «...влияния голосового источника на натуральность [?] синтетической речи».

«Модель ... голосовой щели ... состоящую из 14 уравнений ...» предлагается, чтобы «... с её помощью достичь натуральности синтетической речи, имитировать дикторскую индивидуальность ...».

Итак, вместо $H_i - F_i$ («Хай Фиделити») динамиком будем конструировать голосок Аллы Борисовны! И т.д. и т.п. Посмотрите [29]!

НЕОБХОДИМО:

- покончить с театром абсурда – «незнания» **основы филологии – фонетической функции и семантического кода речи;**
- незамедлительно энергично составить **алфавит семантического кода русской речи, установить связь человек – ЭВМ, начать разработки отечественных интеллектуальных роботов;**
- на основе семантического кода речи широко внедрять чрезвычайно высокоэффективный ($B \approx 150$ бит/с) и надёжный «**телефонный телеграф**», прежде всего в системах **персональной космической связи** [14].

Литература

1. «Тысяча и одна ночь» – сборник арабских сказок, сложившийся в основном к 15 веку; повлиял на фольклор и литературу многих народов. Сов. Энциклопедический словарь, гл. ред. А.М. Прохоров, изд. 4-ое, М., 1989 г., 1632 с.
2. Пирогов А.А. Синтетическая телефония. – М., «Связьиздат», 1963 г., 120 с.
3. Вокодерная телефония. Методы и проблемы. Под редакцией А.А. Пирогова. – М., «Связь», 1974 г., 536 с.
4. Иофе В.К., Корольков В.Г., Сапожкова. – М., Связь, 1979 г., 312с.

5. Пирогов А.А. К вопросу о фонетическом кодировании речи. Тезисы доклада на Пленарном заседании Конференции по теории кодирования и её приложениям, созванной Научным Советом при АН СССР по комплексной проблеме «Кибернетика». ОЭИС Мин. Связи СССР и НТОРЭС им. А.С. Попова. Одесса, ОЭИС, 29 мая 1963г., см. также – Пирогов А.А. К вопросу о фонетическом кодировании речи. НТБ НИИ Радио. Н-4198, П334 , 534.78 А.
6. Пирогов А.А. К вопросу о фонетическом кодировании речи. / Журнал «Электросвязь», 1967 г., № 5, с. 24-31.
7. Пирогов А.А. К фонетической теории речи. – Доклад 3 IV.9 на VI Всесоюзной Акустической Конференции. М., 1968 г.
8. Lienard J.S. Differential analysis of speech: some preliminary results. Eight International Congress of Phonetic Sciences Leeds, August 17-23, 1975.
9. Lienard J.S. at all. Diphone synthesis of French: vocal response unit and automatic prosody from the text. 1977. International Congress on Acoustic, Speech and Signal Processing, p. 560...563.
10. Куля В.И. О дискретном представлении вокодерных сигналов. / Журнал «Электросвязь», 1967 г., № 9.
11. Куля В.И., Пирогов А.А. О фонетическом коде речевого сигнала. / Журнал «Радиотехника», 1970 г., № 6.
12. Куля В.И., Пирогов А.А. Проблема фонетического кодирования речевых сигналов и методы её решения. VIII Всесоюзная Акустическая Конференция. – М., Академия Наук СССР, 23...30 июня 1973г., Доклад В IVв - 3.
13. Ditl Augustin, Dr. Adaptive Erkennung von Vokalen. / «Elektronik», 1974, Heft 3, s. 103.
14. Пирогов А.А. Не только смотреть, но и разговаривать через космос. / Журнал «Вестник связи», 1993, № 11, с. 31, 34, 35.
- 15: Масаширо Ямамото. Направления исследований в компьютерной связи. – Доклад на симпозиуме в МТУСИ, 1 февраля 1994 г.
16. К. Тилеке. Фонд SEL. / Журнал «Электросвязь», 1994 г., № 2, с. 40, 41.
17. Human voice moves ship. / «Moscow News», 26-28. 03. 1983 г.
18. Касса – автомат без кнопок. / Газета «Советская Россия», 10.03.1990.
19. Воспоминания о будущем. / «Computer World». – Moscow, 1994, № 10, с. 56.
20. Давид Родэ (Rohde). Некоторые экспонаты выставки компьютерной технологии. / «Computer Week», №16 (170). 27.04.1995 г., с. 6.
21. IBM создаёт ... карманный ПК. / «Computer World», М., № 43 (151), 10-16 ноября 1994 г., с. 3.
22. Покровский Б.Н. Расчёт и измерения разборчивости речи. – М., «Связьиздат», 1962 г., с. 391.
23. Flanagan James L. Speech and Physis, Synthesis and Perception. – N.-Y., Springer-Verlag, 1972, 444 p., 258 Figures.
24. Сапожков М.А., Михайлов В.Г. Вокодерная связь. – М., «Радио и связь», 1983г., с. 248.
25. Куля В.И. Ортогональные фильтры. – Киев. «Техника», 1967 г., с. 240.
26. Sobolev V.N. Studies of Phonetic Function... to automatic cuality testing of speech... Proc. XI-th International Congress of Phonetic Sciences, vol. 3 (of volumes), August, 1-7, 1987, p.p. 274...277.
27. Калинин Ю.К. Разборчивость речи в цифровых вокодерах. – М., «Радио и связь», 1991 г., 220 с.
28. Калинин Ю.К. Оценка информативности речевых сигналов на основе современных результатов анализа речи. / Журнал «Электросвязь», 1991г., № 8, с. 16...19.
29. Сорокин В.Н. Синтез речи. – М., «Наука», 1992 г., 392 с.
30. Соболев В.Н. Оценивание качества передачи речевого сигнала по изменению его спектральной динамики. / Журнал «Электросвязь», 1991, № 2, с. 19-21.

Москва, 26.06.95

Примечание. В 2001 году в Журнале ЖРФМ, № 1-12, опубликована уникальная статья В.Н. Соболева «Алгоритмы формирования фонетической функции речевого сигнала», в которой впервые приведёно аппаратное обеспечение создания каталога фонемных переходов (алфавита семантического кода **русской** речи).

Пирогов Андрей Андреевич, – доктор технических наук, профессор МТУСИ, Почётный член НТОРЭС им. А.С. Попова, действительный член Русского Физического Общества, Лауреат Премии Русского Физического Общества (1992), Лауреат Международного Культурного Диплома Почёта и титула «Человек года – 1994»: Постановление Американского Биографического Института, США, шт. Сев. Каролина (1995).



Опубликовано: ЖРФМ, 2001, № 1-12, с. 15-28; журнал «Русская Мысль», 2010, № 1-12, с. 99-109.