



Электрический аналог голосового аппарата человека Новиков А. А., Мешков А. Ю.

Разработан электрический аналог голосового аппарата человека, который позволяет генерировать электрические сигналы, эквивалентные голосовым сигналам человека. Определен характер зависимости параметров электрического аналога от конфигурации и физических параметров голосового аппарата человека. Установлено, что начальный аналог, описанный в источниках, не учитывает индивидуальные особенности человека.

В ходе работы установлено зависимость параметров электрического аналога от показателей физического состояния человека (температуры, влажности и т. п.). Также установлено функциональную зависимость параметров голосового аппарата человека от его антропометрических показателей (роста, возраста, пола). С их помощью эти параметры были учтены в эквивалентной электрической схеме.

С помощью усовершенствованного электрического аналога было смоделировано эквивалентные электрические сигналы для нескольких дикторов с учетом их индивидуальных особенностей. Полученные сигналы были сравнены с голосовыми сигналами дикторов. Результаты свидетельствуют о том, что эквивалентные электрические сигналы с достаточной точностью повторяют голосовые сигналы человека и могут использоваться в качестве их аналогов.

На сегодняшний день известны множество методов диагностики. Каждый из них имеет свои недостатки и преимущества; каждый вид диагностики с развитием техники и технологий видоизменяется и совершенствуется. Основной целью совершенствования диагностики в современных условиях есть уменьшение влияния диагностических методов на здоровье человека и улучшение точности диагностических данных.

Известно, что любые изменения в физическом или психоэмоциональном состоянии человека определенным образом отражаются в его голосе. Голос – это вид звуковых колебаний, что создается голосовым аппаратом человека. Таким образом, по изменению голоса, наверняка, можно определить изменения, которые произошли в организме человека. Поэтому возникла идея разработать электрический аналог голосового аппарата человека с целью его диагностики по изменениям голоса.

Объектом данного исследования есть изучение влияния на голос человека его физического состояния.

Предметом данного исследования есть голос человека.

Целью данного исследования есть разработка системы диагностики человека по изменениям в его голосе.

Задания данного исследования:

- установить характер изменения голоса человека в зависимости от его заболевания;
- разработать электрический аналог системы образования голоса человека;
- исследовать на разработанном аналоге изменение голоса человека как изменение выходных параметров модели;
- разработать программное обеспечение для диагностики человека по изменениям его голоса.

Для исследования голосового аппарата человека было создано эквивалентную электрическую схему, которая выступает аналогом голосового тракта человека (см. приложение 1). С помощью этой схемы есть возможность генерировать электрические сигналы, которые по своим характеристиками есть сходными с голосовыми сигналами человека, и могли бы быть сравнены с ними.

Эквивалентная электрическая схема состоит из трех основных элементов.

Источник электрических колебаний при моделировании голосового возбуждения голосового тракта человека выступает аналогом голосовых связок. Он состоит из последовательного соединения 30 источников переменного электрического напряжения, каждый из которых характеризует действие соответствующей гармоники спектра голоса человека.

Величина эквивалентного напряжения источника, что моделирует первую гармонику голоса (основной тон), может быть определена как

$$U_0 = \sqrt{kP},$$

где U_0 – напряжение первого источника;

k – масштабный коэффициент для перевода акустических величин в электрические;

P – давление воздушного потока в голосовой щели.

Частота колебаний электрического напряжения первого источника, которая равна частоте основного тону голосовых связок, которая может быть определена как для натянутой струны:

$$F_0 = \frac{1}{2} \sqrt{\frac{T}{ML_x}}$$

где F_0 – основная частота голоса;



Т – сила натяжения голосовых связок;
 М – масса голосовых связок;
 L_x – длина голосовых связок в натянутом состоянии.

На основную частоту голоса влияют параметры голосовых связок, которые в свою очередь связаны с антропометрическими параметрами тела человека.

Длина голосовых связок L (в ненапрянутом состоянии) зависит от роста человека по следующим формулам:

- для мужчин: $L = 0,18Ht - 15,8$;
- для женщин: $L = 0,24Ht - 28,1$.

Рост человека в формуле задается в сантиметрах, длина голосовых связок – в миллиметрах. [8]

При анализе литературных источников не было найдено прямых зависимостей параметров голосовых связок от массы тела человека. Было указано, что корреляционной связи между массой тела человека и массой его голосовых связок не существует. Эмпирические исследования также не дали положительных результатов в этом направлении. Соответственно, на данном этапе исследования учесть такую антропометрическую характеристику как масса тела человека не удалось. Однако исследования в этом направлении продолжаются.

В таком случае массу голосовых связок выразим через другие параметры:

$$M = \rho LDb$$

де ρ – плотность тканей голосовых связок;

D – глубина голосовой щели.

b – толщина тканей голосовых связок.

Сила натяжения голосовых связок определяется их растяжением при фонации, а также их упругостью:

$$T = k_T \Delta L$$

где k_T – коэффициент упругости голосовых связок;

$\Delta L = L_x - L$ – удлинение голосовых связок по отношению к их первоначальной длине.

В литературных источниках указано, что относительное удлинение голосовых связок при фонации достигает 25% [3]. Соответственно, $\Delta L = 0,25L$ или $L_x = 1,25L$.

Выразим упругость голосовых связок через модуль упругости:

$$k_T = \frac{ES}{L}$$

где E – модуль упругости голосовых связок;

S – поперечная площадь голосовых связок;

В данном случае $S = Db$.

Тогда можем записать:

$$F_0 = \frac{1}{2} \sqrt{\frac{k_T \Delta L}{\rho LDbL_x}} = \frac{1}{2} \sqrt{\frac{0,25EDb}{\rho LDbL_x}} = \frac{1}{2} \sqrt{\frac{0,25E}{1,25\rho L^2}} = \frac{1}{2L} \sqrt{\frac{E}{5\rho}}$$

Видно что, частота колебаний голосовых связок, и соответственно, основная частота голоса не зависит от массы и толщины голосовых связок, а также глубины голосовой щели, а задается только плотностью тканей голосовых связок ($\rho = 1,04 \frac{г}{см^3}$, согласно данным [6]), их длиной, которая определяется по соответствующим формулам, а также их упругими свойствами.

В работе [9] приведена табличная зависимость основной частоты голоса человека от возраста. Однако, эта зависимость не учитывает влияния других антропометрических факторов. В исследовании [7] было изучено изменение модуля упругости голосовых связок с возрастом с учетом половых особенностей человеческого организма. На основе данных, приведенных в указанной публикации, было получено функциональную зависимость между параметрами голосовых связок и возрастными и половыми характеристиками человека (см. приложение 2). Полученные данные хорошо согласуются с данными [9], где указано, что повышение основной частоты голоса, которое наблюдается после 40-50 лет у представителей обоих полов, может быть вызвано увеличением упругости голосовых связок.

Однако, данные, полученные в [7], характеризуют голосовые связки в ненапрянутом состоянии. В [3] указано, что для натянутых связок эти значения должны быть увеличены в 5 – 10 раз, и при растяжении голосовых связок на каждые 3 мм их упругость увеличивается в 2 раза. Таким образом, необходимо ввести некоторый поправочный коэффициент q, который бы учитывал изменение упругости голосовой мышцы в напряженном состоянии с учетом ее растяжения при фонации:

$$q = 10 * 2^{\frac{0,25L}{0,003}}$$

Таким образом, основная частота голоса человека может быть определена с учетом его антропометрических параметров по следующим формулам:



$$F_0 = \frac{1}{2L} \sqrt{\frac{qE}{5\rho}}$$

$$L = 0,018Ht - 0,0158 \text{ (для мужчин); } L = 0,024Ht - 0,0281 \text{ (для женщин)}$$

$$q = 10 * 2^{\frac{0,25L}{0,003}}$$

$$E = -0,0621Ag^3 + 11,659Ag^2 - 686,61Ag + 16780 \text{ (для женщин)}$$

$$E = -0,0989Ag^3 + 16,806Ag^2 - 854,44Ag + 15647 \text{ (для мужчин)}$$

Каждый последующий источник синусоидального напряжения имеет частоту колебаний, кратную частоте основного тона. Амплитуды напряжения каждого источника зависят от амплитуды первого источника. В работе [4] указано, что наиболее точное совпадение спектра синтезированного сигнала со спектром реальной речи наблюдается при использовании спектральной характеристики источника колебаний (аналога голосовых связок) с наклоном огибающей спектра -12 дБ/октаву. При этом в работе [5] указано, что в области высоких частот наблюдается падение огибающей спектра со скоростью 6 дБ/октаву. В ходе исследования было установлено, что наиболее точное совпадение сигналов электрического аналога с голосовыми сигналами человека дает вторая частотная характеристика с наклоном огибающей спектра -6 дБ/октаву. Именно ее предлагается использовать при моделировании голосовых сигналов. График зависимости относительного уровня амплитуд гармоник с учетом спектральных характеристик показан в приложении 3. По этим данным было рассчитано относительный уровень каждого электрического источника (см. приложение 4). Соответственно, при моделировании голосового источника будут учитываться гармоники его спектральной характеристики (до 30-й) с учетом рассчитанных уровней.

Импеданс голосового источника состоит из параллельно подключенного активного сопротивления, величина которого определяется как:

$$R_{гол} = k \left(\frac{12\mu d}{lw^2} + 0,875 \frac{\rho v}{(lw)^2} \right),$$

где $R_{гол}$ – импеданс голосового источника;

μ – коэффициент динамической вязкости воздуха при данных условиях;

d – глубина голосовой щели;

l – длина голосовых связок;

w – ширина голосовой щели;

ρ – плотность воздуха при данных условиях;

v – объемная скорость воздуха на выходе из голосовой щели.

В работе [2] для расчета последней величины была использована следующая формула:

$$v = \frac{150}{1 + \left(\frac{F_0}{100}\right)^2} \text{ [см}^3\text{/сек.]}$$

Конфигурация голосового тракта человека аппроксимируется последовательным набором цилиндрических элементов, и моделируется в виде последовательного соединения фильтров, каждый из которых выступает эквивалентом участка тракта длиной l и площадью S , которая в пределах данного участка принимается постоянной. Значения электрических элементов, которые составляют фильтрующее звено, рассчитываются по формулам:

$$L = \frac{k\rho l}{S} \quad C = \frac{Sl}{k\rho c^2} \quad R = \frac{krl}{S^2} \sqrt{\frac{F_0\rho\mu}{2}} \quad G = \frac{Sl(\eta - 1)}{k\rho c^2} \sqrt{\frac{\lambda F_0}{2c_p\rho}}$$

где L – индуктивность фильтра;

l – длина участка голосового тракта;

S – площадь участка голосового тракта;

C – емкость фильтра;

c – скорость звука при данных условиях;

R – сопротивление фильтра;

r – радиус участка голосового тракта;

F_0 – основная частота голоса;

μ – коэффициент динамической вязкости воздуха при данных условиях;

η – показатель адиабаты воздуха при данных условиях;

λ – коэффициент теплопроводности воздуха при данных условиях;

c_p – удельная теплоемкость воздуха при постоянном давлении при данных условиях.[5]

Излучение звуковых колебаний через ротовое отверстие моделируется при помощи излучателя, который состоит из последовательного резистивного и индуктивного элементов. Значения элементов эквивалентных электрических цепей могут быть определены по приведенным формулам:



$$R_n = \frac{128k}{9\pi^2} \quad L_n = \frac{8ka}{3\pi c},$$

где R_n – сопротивление излучательной нагрузки;

L_n – индуктивность излучательной нагрузки;

a – радиус излучательного участка голосового тракта.[5]

Параметры воздушного потока, что находится в голосовом тракте, а именно, плотность, скорость звука, динамическая вязкость, показатель адиабаты, теплопроводность и теплоемкость, зависят от давления, температуры и влажности воздуха. В ходе работы было установлено функциональный вид этих зависимостей, таким образом, можно было учитывать такой показатель, как температура тела человека. Давление и влажность воздуха рассматривались как косвенные величины и их значения принимались постоянными.

В работе [4] были экспериментально определены параметры голосового тракта с помощью магнитной резонансной томографии (см. приложение 5). Результаты этого исследования использовались для синтеза речи и определения спектральных характеристик голосового сигнала. Эти данные использовались в данном исследовании при моделировании голосовых сигналов. Однако они не учитывают индивидуальности параметров голосового аппарата человека.

На длину голосового тракта человека основное влияние оказывают рост и пол человека. В исследовании [10] были установлены эмпирические зависимости длины голосового тракта от роста человека с учетом половых особенностей (см. приложение 6). В том же исследовании приведена обобщенная зависимость длины голосового тракта от массы человека, однако, на наш взгляд, эта зависимость носит непрямой характер и вес человека не оказывает существенного влияния на этот параметр системы. Возрастные характеристики человека в аспекте длины голосового тракта также не являются основными, так как могут быть учтены при помощи вышеуказанных факторов.

В ходе работы вышеуказанные зависимости были обработаны для получения функциональной зависимости между длиной голосового тракта и ростом человека с учетом половых особенностей человека (см. приложение 6). В результате обработки приведенных данных были получены следующие зависимости:

– для мужчин: $VTL = 0,73Ht + 24,55$

– для женщин: $VTL = 0,47Ht + 55,09$

Длина голосового тракта задается в миллиметрах, рост человека – в сантиметрах.

Процедуру определения параметров голосового тракта человека с учетом его антропометрических показателей предлагается проводить методом масштабного перерасчета. В работе [1] было указано, что пропорциональное изменение площади поперечного сечения для всего голосового тракта не имеет влияния на частоты формант. Поскольку данные, представленные в работе [4], использовались для расчета формант и спектральных характеристик, то, на наш взгляд, данный метод можно использовать в данном случае.

Методика перерасчета функции площади поперечного сечения голосового тракта, с учетом вариации длины голосового тракта человека и его половых особенностей, заключается в следующем. Согласно вышеуказанных зависимостей определяется длина голосового тракта человека. Длина голосового тракта в исследовании [4] составляла 17 см. Соответственно, масштабный коэффициент определяется как отношение полученной по формулам длины голосового тракта к 17 см. Далее табличные значения площади через каждые 0,5 см. пересчитываются при помощи этого масштабного коэффициента. В результате получаем новую табличную зависимость площади поперечного сечения уже для текущей длины голосового тракта. Длина элементарной секции голосового тракта также пересчитывается с учетом масштабного коэффициента: изначальная длина в 0,5 см. умножается на вышеуказанный масштабный коэффициент. По полученным в результате данным моделируются параметры фильтрующей системы голосового тракта.

За счет проведения всех вышеуказанных операций мы получили возможность моделировать при помощи эквивалентной электрической схемы (см. приложение 1) электрические сигналы, которые являются эквивалентами голосовых сигналов человека. Для упрощения расчетов параметров эквивалентной схемы было разработано программу в среде Borland Delphi. По введенным антропометрическим параметрам человека (возраст, рост, пол) программа рассчитывала все параметры схемы.

Результаты, полученные при помощи данной схемы, были сравнены с реальными акустическими записями шестерых дикторов, данные которых представлены в таблице.

Для каждого диктора было проведено акустическую запись звука «А». Результаты были сравнены с электрическим сигналом, сгенерированным для каждого диктора при помощи электрической схемы с учетом его данных (см. приложение 7).



Диктор	Пол	Возраст, лет	Рост, мм
Диктор 1	М	19	168
Диктор 2	М	19	165
Диктор 3	Ж	19	157
Диктор 4	Ж	19	160
Диктор 5	Ж	43	165
Диктор 6	М	49	172

Полученные результаты свидетельствуют о том, что по структуре эквивалентные электрические сигналы почти точно повторяют голосовые сигналы человека. Определенные отклонения во временных и амплитудных характеристиках могут объясняться некоторыми отклонениями, которые уже были в наличии в организме дикторов. Следовательно, можно также утверждать, что с помощью эквивалентных сигналов есть возможность выявлять эти отклонения и проводить диагностику общего состояния организма человека.

Таким образом, можно утверждать, что в ходе исследования было разработано электрическую систему, которая может генерировать электрические сигналы, которые являются эквивалентами голосовых сигналов человека. Система учитывает не только физическое состояние человека, но и его антропометрические особенности. Это позволяет создавать эквивалентные сигналы для каждого человека индивидуально. При сравнении эквивалентных сигналов с голосовыми сигналами человека наблюдаются отклонения, которые, на наш взгляд, свидетельствуют об определенных отклонениях в его организме, и могут быть определены по ним.

На текущем этапе исследования сравнение сигналов проводилось только в аналитическом виде. На данный момент проводится разработка программного обеспечения для цифрового сравнения полученных сигналов и определения критериев соответствия отклонений акустических записей от эталонных электрических сигналов.

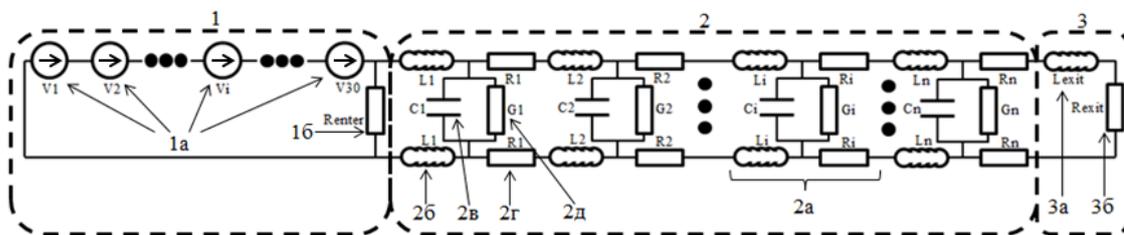
Список использованных источников

1. Квасов А. Н. Модель голосообразования и анализ речевого сигнала в норме и при патологии: автореф. дис. на соискание ученой степени кандидата технических наук: спец. 05.13.18 – Математическое моделирование, численные методы и комплексы программ / Квасов А. Н. – Томск, 2007. – 21 с.
2. Сапожков М. А. Речевой сигнал в кибернетике и связи. – М.: «Связь», 1963. – 451 с.
3. Сорокин В. Н. Теория речеобразования – М.: Радио и связь, 1985 г. – 312 с.
4. Фант Г. Акустическая теория речеобразования. Пер. с англ. под ред. В. С. Григорьева. М.: «Наука», 1964 г. – 284 с.
5. Фланеган Дж. Анализ, синтез и восприятие речи. Пер. с англ. под ред. А. А. Пирогова. М.: «Связь», 1968 г. – 397 с.
6. Brad H. Story, Ingo R. Tltze. Voice simulation with a body-cover model of the vocal folds // Journal of Acoustical Society of America. – 1995. – № 97 (2). – p. 1249–1260.
7. Eric Goodyer, Sandra Hemmerich, Frank Muller, James B. Kobler, Markus Hess. The shear modulus of the human vocal fold, preliminary results from 20 larynxes // Eur Arch Otorhinolaryngol. – 2007. – № 264 (1). – p. 45–50.
8. Filho J.A., de Melo E.C., Tsuji D.H., de Giacomo Carneiro C., Sennes L.U. Length of the human vocal folds: proposal of mathematical equations as a function of gender and body height // Ann Otol Rhinol Laryngol. – 2005. – № 114 (5). – p. 390–392.
9. Jody Kreiman, Diana Van Lancker Sidtis. Foundations of Voice Studies: An Interdisciplinary Approach to Voice Production and Perception. – 2011. – 512 p.
10. W. T. Fitch, J. Giedd. Morphology and development of the human vocal tract: A study using magnetic resonance imaging. // Journal of Acoustical Society of America. – 1999. – № 106 (3). – p. 1511–1522.

Приложения

Приложение 1

Электрическая схема, эквивалентная голосовому аппарату человека

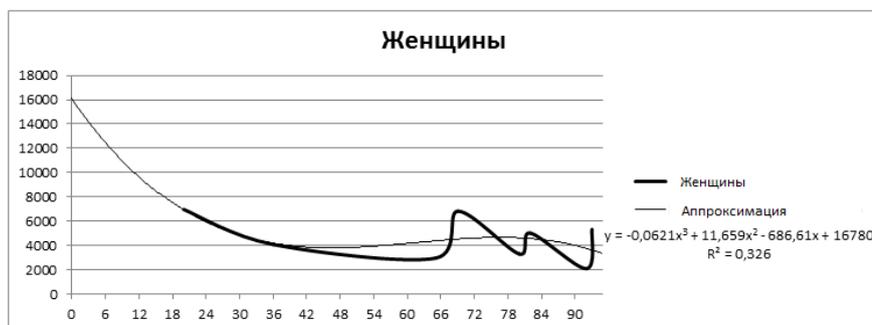


- 1 – эквивалентная схема голосового источника;
- 1а – источники переменного электрического напряжения;
- 1б – импеданс голосового источника;
- 2 – фильтрующая длинная линия;
- 2а – фильтрующий элемент длинной линии;
- 2б – индуктивность фильтрующего элемента;
- 2в – емкость фильтрующего элемента;
- 2г – сопротивление фильтрующего элемента;
- 2д – проводимость фильтрующего элемента;
- 3 – излучательная нагрузка;
- 3а – индуктивность излучательной нагрузки;
- 3б – активное сопротивление излучательной нагрузки.

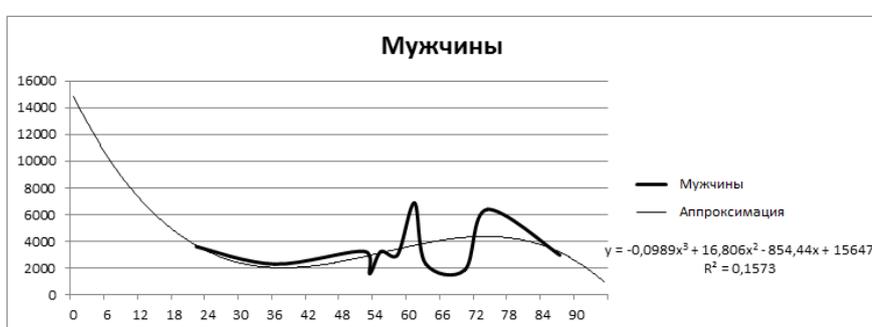
Приложение 2

Зависимость модуля упругости голосовых связок от возраста с учетом половых особенностей

Возраст, лет	Модуль упругости голосовых связок, Па
20	6977
36	4093
65	2966
69	6837
80	3317
82	5033
92	2108
93	5335



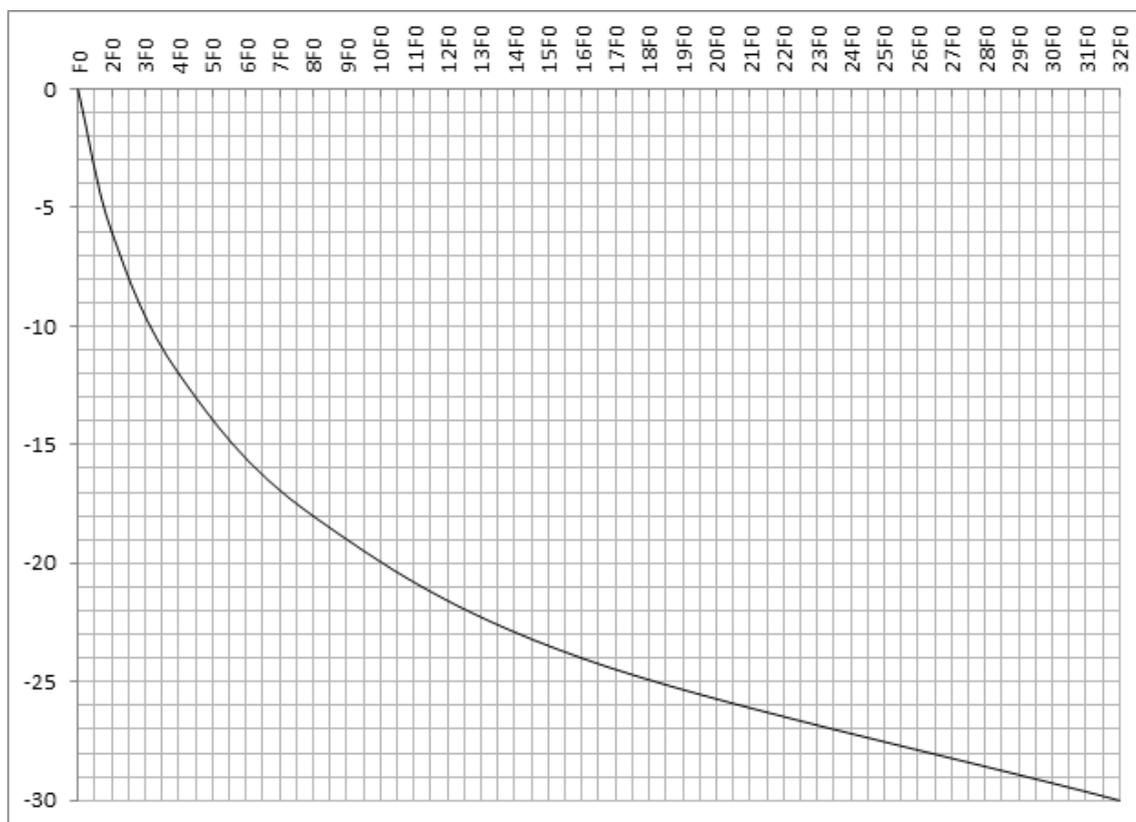
Возраст, лет	Модуль упругости голосовых связок, Па
22	3602
36	2299
52	3254
53	1601
55	3242
58	2999
61	6874
63	2326
70	1854
74	6387
87	2978





Приложение 3

Спектральная характеристика голосового источника (согласно данным [5])
Задаётся падением относительного уровня амплитуды гармоники – 6 дБ/октаву



Приложение 4

Относительные уровни электрических источников соответствующих гармоник

Номер гармоники	Относительный уровень						
1	1,000000	7	0,020683	13	0,006023	19	0,002827
2	0,251195	8	0,015850	14	0,005196	20	0,002552
3	0,111954	9	0,012534	15	0,004528	21	0,002316
4	0,063099	10	0,010160	16	0,003982	22	0,002111
5	0,040446	11	0,008402	17	0,003528	23	0,001932
6	0,028122	12	0,007064	18	0,003148	24	0,001775
						25	0,001636
						26	0,001513
						27	0,001403
						28	0,001305
						29	0,001217
						30	0,001137



Приложение 5

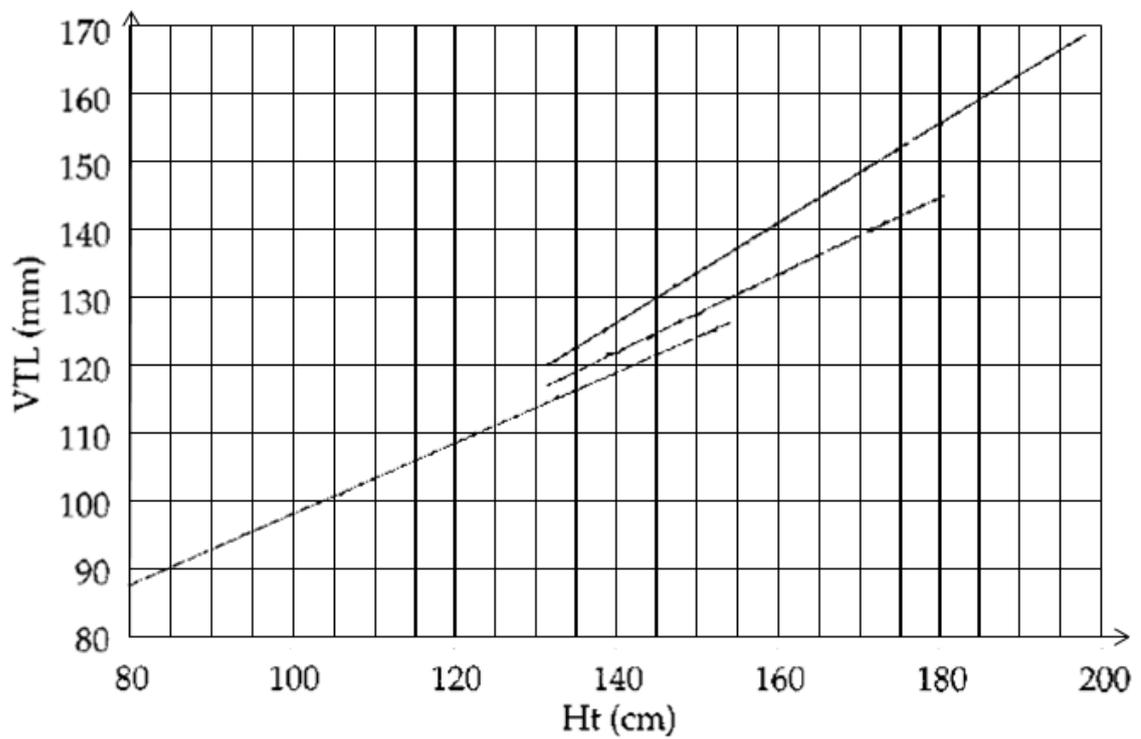
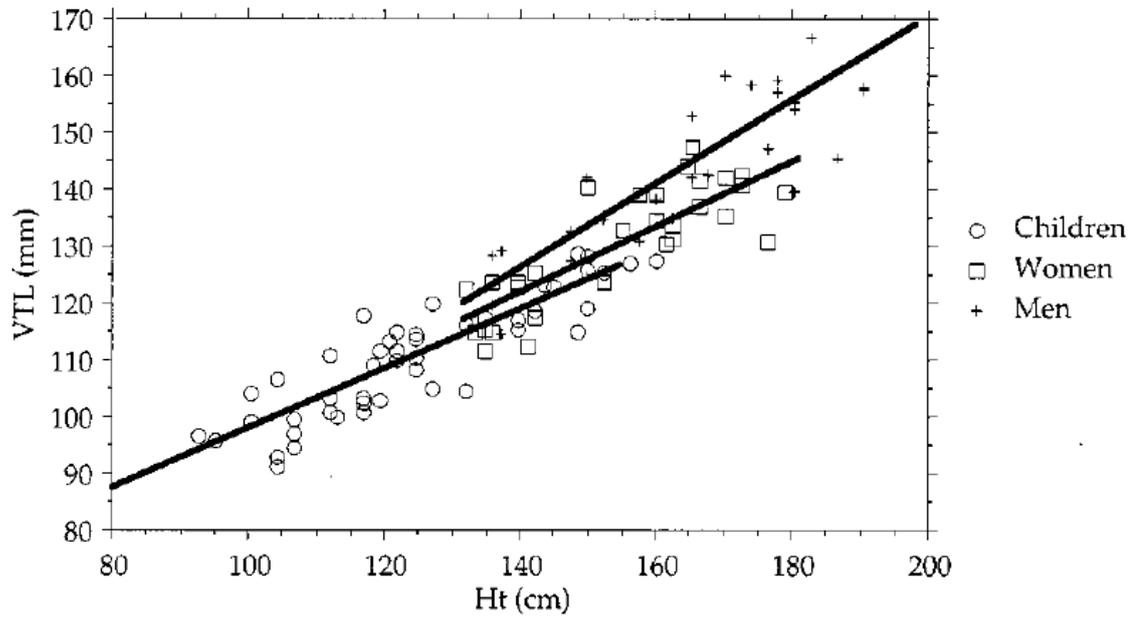
Конфигурация голосового тракта при фонации звука «А» (по данным [4])

Расстояние от голосовых связок, см	Площадь поперечного сечения голосового тракта, см ²
0,0	2,6
0,5	1,6
1,0	1,3
1,5	1,0
2,0	4,0
2,5	2,6
3,0	1,6
3,5	1,0
4,0	0,65
4,5	0,65
5,0	0,65
5,5	1,0
6,0	1,3
6,5	1,6
7,0	2,0
7,5	2,6
8,0	2,6
8,5	1,6
9,0	3,2
9,5	4,0
10,0	5,0
10,5	6,5
11,0	8,0
11,5	8,0
12,0	8,0
12,5	8,0
13,0	8,0
13,5	8,0
14,0	8,0
14,5	8,0
15,0	6,5
15,5	5,0
16,0	5,0
16,5	5,0
17,0	5,0



Приложение 6

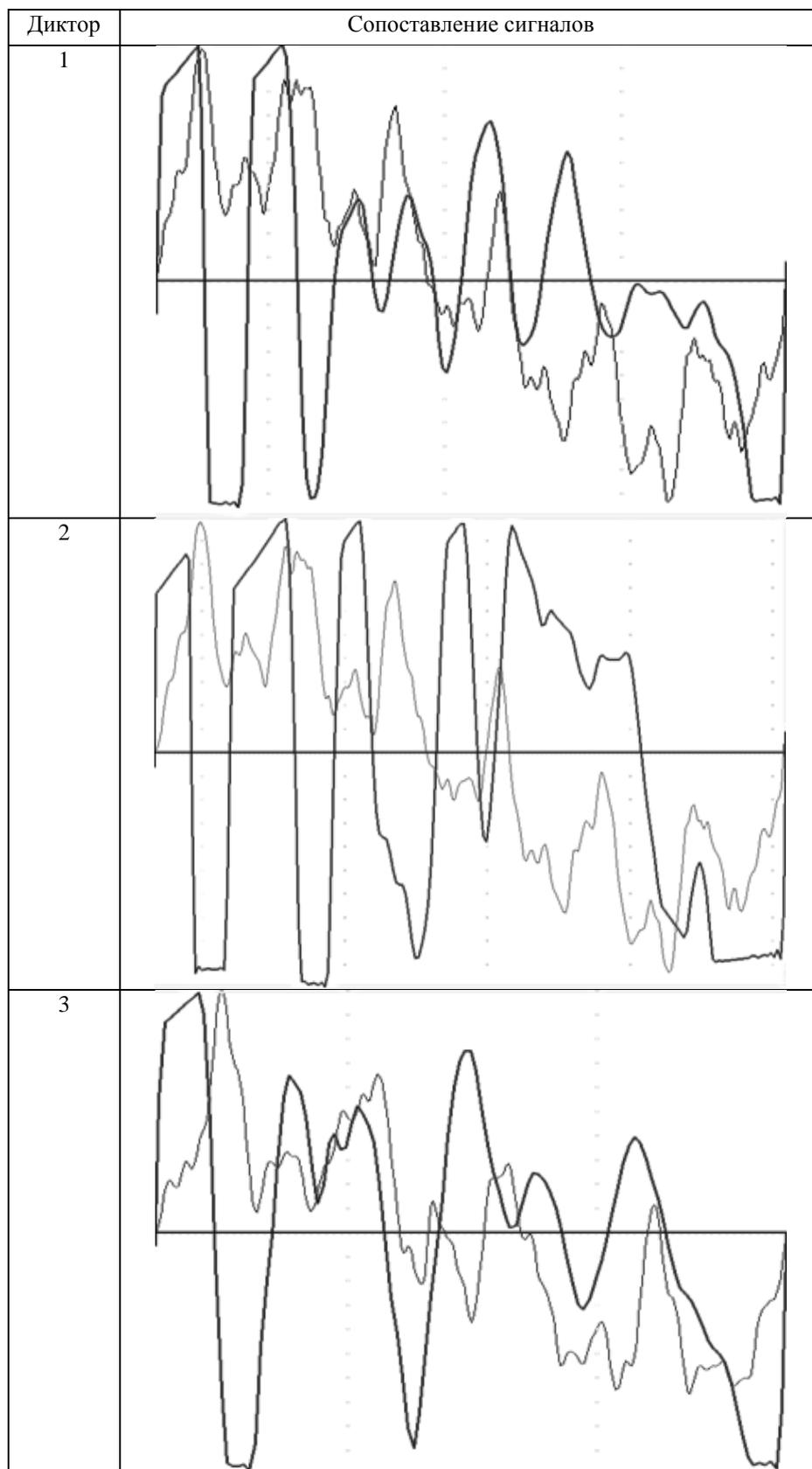
Зависимость длины голосового тракта человека от возраста с учетом половых особенностей



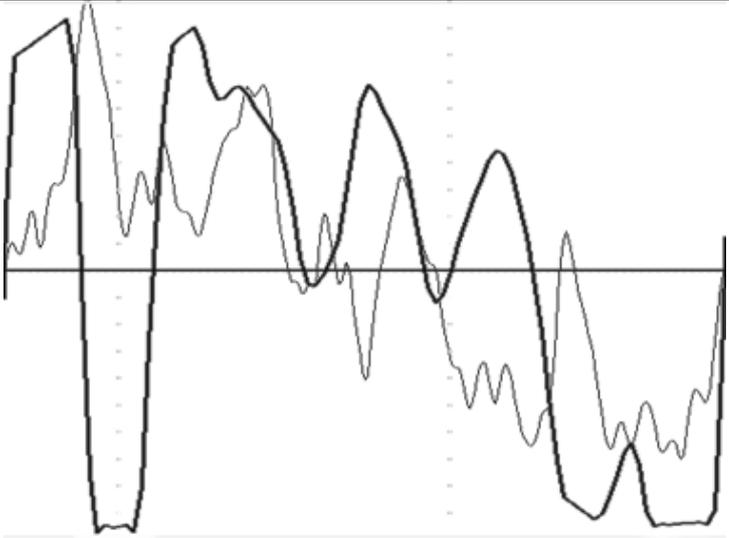
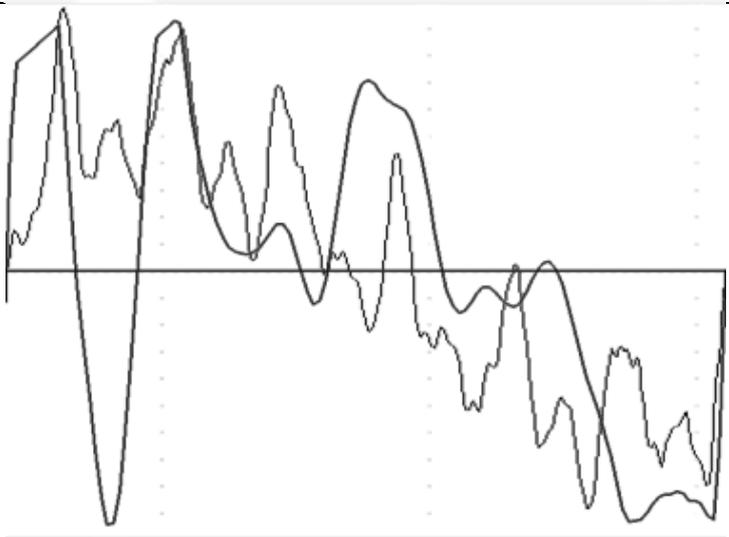
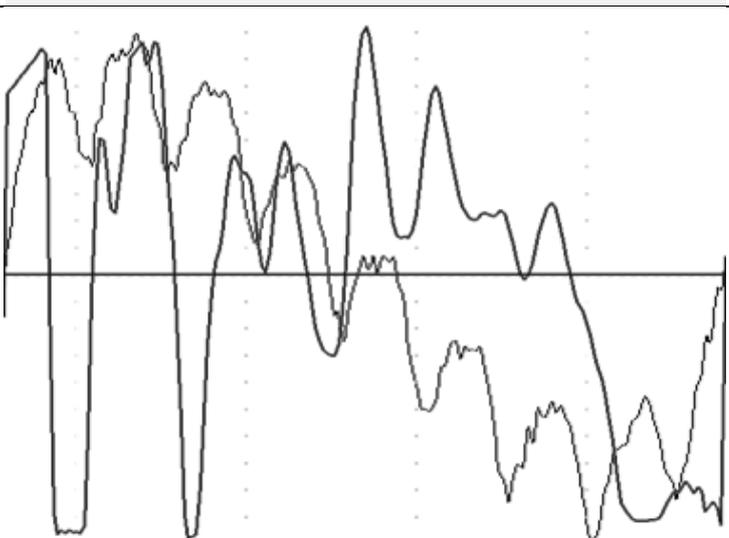


Приложение 7

Сопоставление акустических и электрических сигналов
Толстая линия – акустическая запись, тонкая – электрический сигнал





Диктор	Сопоставление сигналов
4	 The plot shows two overlapping waveforms. The top waveform is a thick black line, and the bottom is a thin grey line. Both show a similar sequence of peaks and troughs. Vertical dotted lines are placed at approximately 15%, 35%, 55%, and 75% of the horizontal axis. A solid horizontal line is drawn across the middle of the plot.
5	 The plot shows two overlapping waveforms. The top waveform is a thick black line, and the bottom is a thin grey line. The waveforms are highly similar in their oscillatory pattern. Vertical dotted lines are placed at approximately 15%, 35%, 55%, and 75% of the horizontal axis. A solid horizontal line is drawn across the middle of the plot.
6	 The plot shows two overlapping waveforms. The top waveform is a thick black line, and the bottom is a thin grey line. The waveforms are highly similar in their oscillatory pattern. Vertical dotted lines are placed at approximately 15%, 35%, 55%, and 75% of the horizontal axis. A solid horizontal line is drawn across the middle of the plot.