

===== ИНФОРМАЦИОННЫЕ ТЕХНОЛОГИИ В ТЕХНИЧЕСКИХ =====  
===== И СОЦИАЛЬНО-ЭКОНОМИЧЕСКИХ СИСТЕМАХ =====

## МОДЕЛЬ АКУСТИЧЕСКОГО ВЗАИМОДЕЙСТВИЯ РЕЧЕВОГО ТРАКТА, ПОДСВЯЗОЧНОЙ ОБЛАСТИ И ГОЛОСОВОГО ИСТОЧНИКА

К. С. Горбунов, И. С. Макаров

Учреждение Российской Академии Наук Институт Проблем Передачи Информации им.

А. А. Харкевича РАН

101447, Москва, Б. Каретный пер. 19

e-mail: [ksggor@gmail.com](mailto:ksggor@gmail.com), [speechprod\\_mak@mail.ru](mailto:speechprod_mak@mail.ru)

Поступила в редколлегию 02.05.2009

**Аннотация** – Статья посвящена построению математической модели акустического взаимодействия объемной скорости, протекающей через голосовую щель, речевого тракта и подсвязочной области (трахея, бронхи, легкие). Модель основана на аппроксимации акустических импедансов авторегрессионными моделями с подвижным средним. Результаты экспериментов хорошо согласуются с данными других работ по изучению взаимодействия голосового источника и тракта.

### Введение

Акустическое взаимодействие речевого тракта, подсвязочной области (трахея, бронхи и легкие) и голосового источника является важным компонентом в артикуляторном синтезаторе речи. По мнению ряда исследователей [1], учет такого взаимодействия играет решающую роль в придании синтетическому голосу натуральности.

Совместное влияние речевого тракта и подсвязочной области на объемную скорость, протекающую через голосовую щель, приводит к возникновению разнообразных эффектов. Среди них в литературе в первую очередь отмечаются два эффекта: форма импульса объемной скорости становится менее симметричной (по сравнению с соответствующим импульсом, вычисленным без учета акустического взаимодействия), и на переднем фронте появляются так называемые «складки» (ripples), представляющие из себя акустические колебания на резонансных частотах речевого тракта и трахеи [1]. Форма этих складок меняется от периода к периоду (такое явление в [2] было названо явлением суперпозиции).

Математически взаимодействие между трактом, голосовым источником и подсвязочной областью может быть описано следующим образом. Рассмотрим уравнение, связывающее объемную скорость  $u_g(t)$ , протекающую через голосовые складки, и перепад акустического давления под складками (на входе в трахею)  $p_s$  и над ними (на входе в речевой тракт)  $p_{tr}$  [3]:

$$\frac{k\rho u_g^2}{2S_g^2} + \frac{12\mu d_g l_g u_g}{S_g^3} + \frac{d}{dt} \left( \frac{\rho d_g u_g}{S_g} \right) = p_s - p_{tr} \quad (1)$$

Здесь  $t$  – время,  $S_g$  – площадь щели между голосовыми складками,  $d_g$  – толщина голосовых складок,  $l_g$  – длина голосовых складок,  $\rho$  – плотность воздуха,  $\mu$  – коэффициент вязкости,  $k$  – коэффициент динамического сопротивления.

Акустическое взаимодействие между речевым трактом, подсвязочной областью и голосовым источником объясняется колебаниями акустического давления в речевом тракте и трахее. Эти колебания создают переменные компоненты давления на входе и выходе голосовой щели и, таким образом, влияют на объемную скорость  $u_g$ , протекающую через голосовые складки. Если  $Z_{tr}(t)$  – импульсная импедансная характеристика над голосовой щелью (вход в речевой тракт), а  $Z_s(t)$  – импульсная импедансная характеристика под голосовой щелью (вход в трахею), то переменные компоненты давления на входе в трахею и в речевой тракт определяются следующими соотношениями:

$$\begin{aligned} p_s &= Z_s * (-u_g) \\ p_{tr} &= Z_{tr} * u_g \end{aligned} \quad (2)$$

Здесь символ «\*» означает свертку.

Напомним, что импульсная импедансная характеристика в некотором сечении тракта определяется как отклик давления в данном сечении на единичный импульс объемной скорости, приложенный в том же сечении. Преобразование Фурье от такой характеристики определяет акустический импеданс.

Подставляя (2) в (1), получаем математическое описание акустического взаимодействия между трактом, подсвязочной областью и объемной скоростью через голосовую щель:

$$\frac{k\rho u_g^2}{2S_g^2} + \frac{12\mu d_g l_g u_g}{S_g^3} + \frac{d}{dt} \left( \frac{\rho d_g u_g}{S_g} \right) = -Z_s * u_g - Z_{tr} * u_g \quad (3)$$

Уравнение (3) может быть решено относительно  $u_g$  только численными методами. Кроме того, обычно учитывают ряд обстоятельств, существенно упрощающих запись (3). Во-первых, полагают  $k = 1$ . Во-вторых, пренебрегают вторым и третьим членами на том основании, что для реальных конфигураций голосовой щели эти члены оказываются очень малыми по сравнению с первым членом – даже для  $S_g$ , близких к нулю [2, 4]. Таким образом, уравнение (3) фактически вырождается в (4):

$$\frac{\rho u_g^2}{2S_g^2} = -Z_s * u_g - Z_{tr} * u_g \quad (4)$$

Учитывая определение свертки и переходя в (4) от непрерывного времени  $t$  к дискретным отсчетам  $j$ , можем записать (4) как:

$$\frac{\rho u_g^2(j)}{2S_g^2} = -\sum_{i=1}^{\infty} Z_s(j-i)u_g(i) - \sum_{i=1}^{\infty} Z_{tr}(j-i)u_g(i) \quad (5)$$

Практически, суммирование в (5) осуществляется вплоть до отсчета, на котором модуль амплитуды импедансной характеристики уменьшается (по сравнению с модулем амплитуды на первом отсчете) не менее, чем на 50-60 дБ.

Если необходимые импульсные импедансные характеристики или соответствующие им акустические импедансы известны, то, теоретически, объемная скорость  $u_g$  может быть вычислена непосредственно из уравнения (5) [5]. Практически, однако, этот метод

обладает одним существенным недостатком [6] – импедансные отклики затухают сравнительно медленно, что приводит к вычислительным трудностям при быстрых изменениях формы речевого аппарата. Другая трудность заключается в сложности учета начальных условий для этих откликов при переходе от одного импульса объемной скорости к другому [2].

В работе [6] было предложено вычислять некоторую функцию, затухающую во времени гораздо быстрее, чем импульсная импедансная характеристика. Такой подход в значительной степени решает проблему быстрых изменений формы речевого аппарата, однако не снимает проблемы учета начальных условий.

В ряде работ предложены различные схемы, упрощающие правую часть соотношения (5). В [2, 7] влияние речевого тракта и подсвязочной области моделировалось введением нескольких двухполюсных фильтров, каждый из которых был настроен на определенную форманту. Такой подход существенно облегчает решение (5), однако вносит новую проблему – необходимость по акустическим импедансам оценивать резонансные частоты речевого тракта и подсвязочной области. Для ряда конфигураций речевого тракта такая проблема представляет собой весьма сложную и нетривиальную задачу [8].

В работе [9] уравнение (5) решалось методом малого параметра при некоторых дополнительных упрощающих обстоятельствах. Существенным недостатком данного подхода является предположение о малых вариациях объемной скорости при учете акустического взаимодействия по сравнению с объемной скоростью, вычисленной без учета взаимодействия. По различным источникам [2, 4], в действительности такие вариации оказываются весьма значительными, так что метод малого параметра в общем случае неприменим.

Цель данной работы заключается в создании такого метода решения уравнения (5) (или (3)), который бы давал возможность наиболее полно (без существенных упрощающих предположений) учесть взаимодействие тракта, голосового источника и подсвязочных областей и, кроме того, был бы свободен от недостатков методов, построенных в [5, 6]. Идея построения заключается в аппроксимации необходимых акустических импедансов некоторой авторегрессионной моделью с подвижным средним. Использование такой модели вместо исходных импульсных импедансных характеристик приводит к тому, что правая часть в соотношении (5) может быть реализована посредством цифровой фильтрации (а не через свертку). Это обстоятельство, в свою очередь, решает проблему учета начальных условий.

### Теория

Предположим, что акустический импеданс  $\tilde{Z}(\omega)$ ,  $\omega$  - радиальная частота, можно аппроксимировать (в  $z$ -области) авторегрессионной моделью с подвижным средним:

$$\tilde{Z}(\omega) = \frac{\tilde{P}(\omega)}{\tilde{U}(\omega)} = \frac{\sum_{i=0}^N b_i z^{-i}}{1 + \sum_{i=1}^M a_i z^{-i}} \quad (6)$$

Здесь  $\tilde{P}(\omega), \tilde{U}(\omega)$  - Фурье-образы акустического давления и объемной скорости, соответственно,  $\{a_i\}, \{b_i\}$  - коэффициенты числителя и знаменателя модели,  $N, M$  – порядок соответствующих полиномов. В пространстве дискретных отсчетов соотношение (6) запишется как:

$$p(j) = \sum_{i=0}^N b_i u(j-i) - \sum_{i=1}^M a_i p(j-i) \quad (7)$$

Введем два подстрочных индекса – «s» и «tr» для обозначения параметров авторегрессионных моделей с подвижным средним, аппроксимирующих входной импеданс в трахею и в речевой тракт, соответственно. Тогда уравнение (5) с учетом (2) и (7) может быть переписано так:

$$\frac{\rho u_g^2(j)}{2S_g^2} + (b_{s,0} + b_{tr,0})u_g(j) = -\sum_{i=1}^{N_s} b_{s,i}u_g(j-i) - \sum_{i=1}^{N_{tr}} b_{tr,i}u_g(j-i) + \sum_{i=1}^{M_{tr}} a_{tr,i}p_{tr}(j-i) - \sum_{i=1}^{M_s} a_{s,i}p_s(j-i) \quad (8)$$

Уравнение (8) является квадратным относительно  $u_g(j)$ . Учитывая неотрицательность объемной скорости и приводя в (8) подобные члены, получаем следующее решение:

$$u_g(j) = \frac{S_g^2}{\rho} \left( -[b_{s,0} + b_{tr,0}] + \sqrt{[b_{s,0} + b_{tr,0}]^2 - \frac{2\rho}{S_g^2} C} \right) \quad (9)$$

Здесь  $C = \sum_{i=1}^{N_s} b_{s,i}u_g(j-i) + \sum_{i=1}^{N_{tr}} b_{tr,i}u_g(j-i) - \sum_{i=1}^{M_{tr}} a_{tr,i}p_{tr}(j-i) + \sum_{i=1}^{M_s} a_{s,i}p_s(j-i)$ .

Введем пару уравнений, определяющих, соответственно, акустическое давление под и над голосовыми складками:

$$\begin{aligned} p_s(j) &= -\sum_{i=0}^{N_s} b_{s,i}u(j-i) - \sum_{i=1}^{M_s} a_{s,i}p(j-i) \\ p_{tr}(j) &= \sum_{i=0}^{N_{tr}} b_{tr,i}u(j-i) - \sum_{i=1}^{M_{tr}} a_{tr,i}p(j-i) \end{aligned} \quad (10)$$

Соотношения (9) и (10) полностью определяют решение уравнения (5). В начальный момент времени  $j = 1$  начальные условия для объемной скорости и надсвязочного давления нулевые, а начальное условие для подсвязочного давления равно давлению в легких. Из (9) определяем значение объемной скорости  $u_g(1)$ , после чего находим из (10) значения надсвязочного и подсвязочного давлений в момент времени  $j = 1$ . Зная  $u_g(1)$ ,  $p_s(1)$  и  $p_{tr}(1)$ , из (9) определяем значение объемной скорости, протекающей через голосовую щель, в момент времени  $j = 2$ , после чего находим (из (10)) значения давлений в тот же момент времени, и т.д. В описанной схеме начальные условия учитываются автоматически, так что никаких проблем, свойственных алгоритмам из [5, 6], не возникает. Более того, поскольку учет акустического взаимодействия между речевым трактом, голосовым источником и подсвязочной областью в рамках построенного подхода реализован как цифровая фильтрация (а не как свертка), не возникает никаких трудностей, связанных с длинными импульсными импедансными характеристиками.

Практическая реализация построенной схемы требует решения двух задач – вычисления входных акустических импедансов в речевой тракт и подсвязочную область со стороны голосовой щели и аппроксимации этих импедансов моделями вида (6). Рассмотрим каждую из задач подробнее.

Вычисление акустических импедансов по функциям площадей поперечных сечений речевого тракта и подсвязочной области выполнялось с помощью обобщенной схемы длинной линии, построенной в [10]. В качестве конфигураций речевого аппарата были

использованы 4 набора площадей поперечных сечений для английских гласных /a, æ, i, u/, заимствованных из работы [11]. Во всех расчетах учитывалось наличие потерь на вязкое трение и теплопроводность, а также податливость стенок речевого тракта. В качестве модели подвязочной области использовалась полость некоторого объема, аппроксимирующая легкие, которая сочленялась с однородной трубой с потерями и податливыми стенками, моделирующей трахею (подробнее см. [12]). Все необходимые акустические импедансы вычислялись на дискретной сетке частот, содержащей 512 отсчетов и покрывающей частотный диапазон от 0 Гц до 5 кГц.

В общем виде входной акустический импеданс в некоторую полость (речевой тракт или трахею) определяется так [10]:

$$\tilde{Z}(\omega) = \frac{Z_{rad}D - B}{A - CZ_{rad}} \quad (11)$$

Здесь  $Z_{rad}$  - импеданс нагрузки (для речевого тракта это импеданс излучения на губах, для трахеи это импеданс легких),  $A, B, C, D$  - некоторые функции радиальной частоты и распределения площадей поперечных сечений полости. Из (11) следует, что в общем случае импеданс имеет как полюсы (определяемые знаменателем (11)), так и нули (определяемые числителем (11)).

Аппроксимация (11) моделями вида (6) осуществлялась в два этапа. На первом этапе по входному акустическому импедансу вычислялась его импульсная характеристика с помощью обратного преобразования Фурье. На втором этапе импульсная характеристика аппроксимировалась авторегрессионной моделью с подвижным средним методом Стейглитца - МакБрайда [13]. При этом порядок полиномов в числителе и знаменателе выбирался равным 20. Ошибка аппроксимации импульсных импедансных характеристик и акустических импедансов моделями (6) составила доли процента.

### Эксперименты

Конфигурации речевого тракта представляли из себя набор площадей поперечных сечений для 4-х гласных американского английского языка /a, æ, i, u/, измеренных с помощью ядерно-магнитной томографии [11]. На Рис. 1 представлены функции площадей для гласных /a, i/, соответствующие передаточные функции и входные акустические импедансы для данных конфигураций. На Рис. 2 показаны те же графики для гласных /æ, u/.

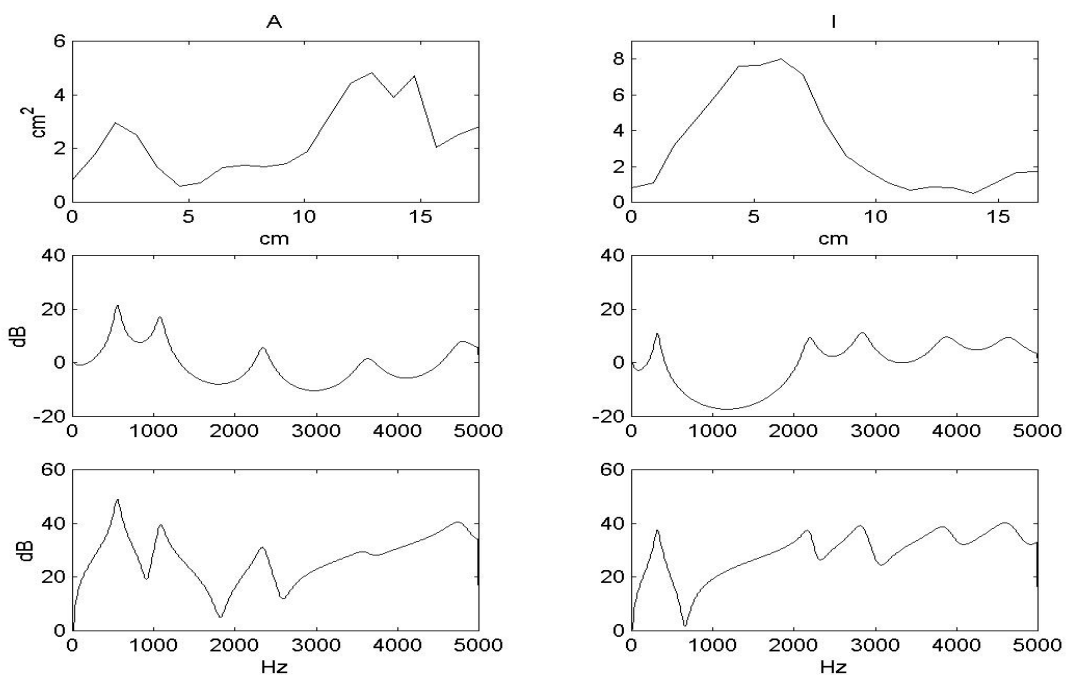


Рис. 1

Гласные /a/ (левый столбец) и /i/ (правый столбец).

Первая строка – площади поперечного сечения.

Вторая строка - передаточные функции.

Третья строка – входные акустические импедансы в тракт со стороны голосовой щели.

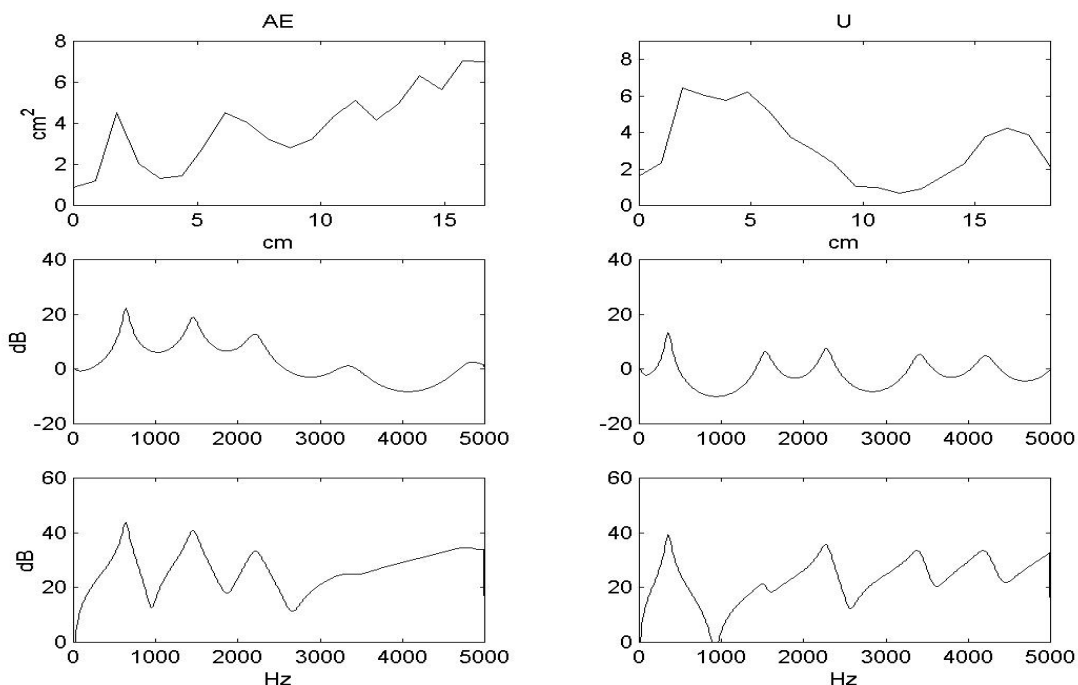


Рис. 2

Гласные /æ/ (левый столбец) и /u/ (правый столбец).

Первая строка – площади поперечного сечения.

Вторая строка - передаточные функции.

Третья строка – входные акустические импедансы в тракт со стороны голосовой щели.

Из Рис. 1 и 2 видно, что передаточные функции и входные акустические импедансы обнаруживают локальные максимумы на одних и тех же частотах (соответствующих частотам формант). С другой стороны, импедансы имеют глубокие провалы (нули), отсутствующие у передаточных функций. Это свойство уже обсуждалось выше применительно к соотношению (11).

На Рис. 3 представлен входной акустический импеданс в подсвязочную область (трахея, бронхи и легкие) со стороны голосовой щели. При его вычислении длина и площадь поперечного сечения трахеи полагались равными 15 см и 3 см<sup>2</sup>, соответственно. Объем легких был равен 3000 см<sup>3</sup>, а сопротивление 40 акустическим Омам.

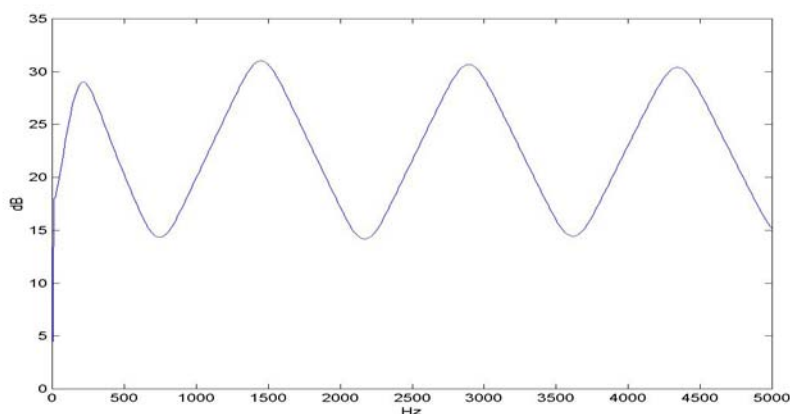


Рис. 3

Входной акустический импеданс в подсвязочную область со стороны голосовой щели.

В качестве начального условия для  $p_s$  использовалось значение, равное 1500 Па. В качестве модели площади просвета между голосовыми складками  $S_g$  (отметим, что уравнение (9) предполагает, что функция  $S_g(t)$  является заданной) использовалась модель, описанная в [14, 15]. Модель определяется следующим образом:

$$S_g(t) = \begin{cases} S_{\max} \left[ \sin\left(\frac{\pi t}{2t_1 T_0}\right) \right]^\alpha, & 0 \leq t \leq t_1 T_0 \\ S_{\max} \left[ \cos\left(\frac{\pi(t - t_1 T_0)}{2(t_2 - t_1) T_0}\right) \right]^\beta, & t_1 T_0 < t \leq t_2 T_0 \\ 0, & t_2 T_0 < t \leq T_0 \end{cases} \quad (12)$$

Здесь  $T_0$  – период основного тона,  $t_1$  – отношение интервала открытия голосовой щели к периоду основного тона,  $t_2$  – отношение интервала открытой голосовой щели к  $T_0$ ,  $S_{\max}$  – максимальное значение площади просвета между голосовыми складками,  $\alpha$ ,  $\beta$  – некоторые константы, определяющие форму голосовой щели. Все эксперименты были проведены для следующих значений параметров:  $t_1 = 0.4$ ,  $t_2 = 0.6$ ,  $S_{\max} = 0.2$  см<sup>2</sup>,  $\alpha = 2$ ,  $\beta = 1$ . Что касается

частоты основного тона, то эксперименты были проведены для трех ее значений – 80 Гц, 100 Гц и 200 Гц.

Рассмотрим сначала результаты для частоты основного тона 80 Гц. На Рис. 4 и 5 представлены данные для гласных /a, i/ и /æ, u/, соответственно. Первый столбец соответствует гласному /a/ (/æ/), а второй – гласному /i/ (/u/). Первая строка представляет собой форму площади голосовой щели  $S_g$ , вычисленной для трех периодов основного тона. Во второй строке показаны графики объемной скорости  $u_g$ , вычисленной без учета взаимодействия с речевым трактом и подвязочной областью (пунктир) и с его учетом (сплошная линия). В третьей строке представлены голосовые источники (первые производные от  $u_g$ ), вычисленные как для случая без учета взаимодействия (пунктир), так и с его учетом.

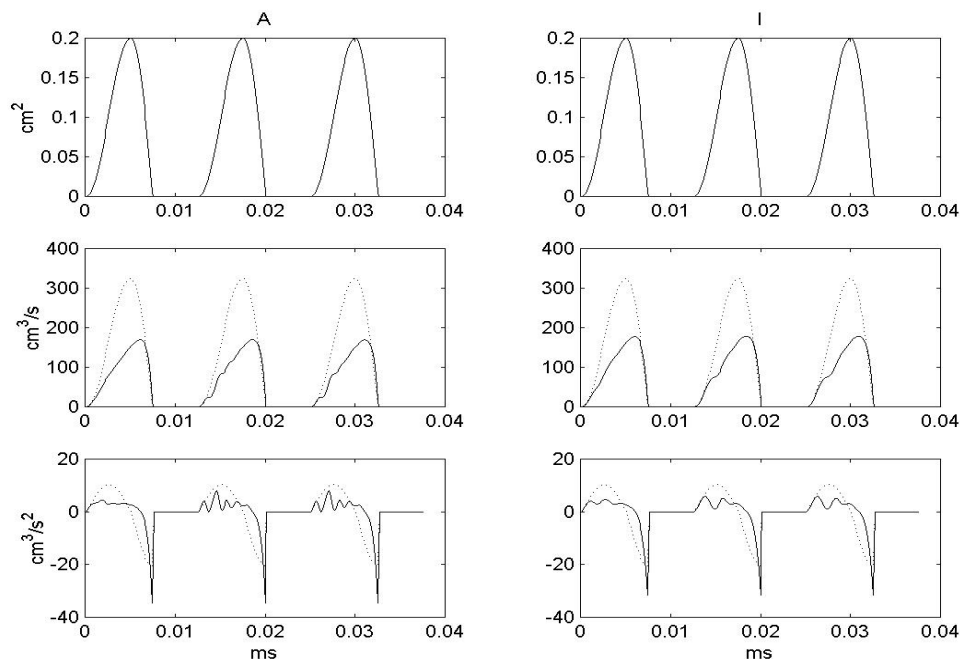


Рис. 4.

Гласные /a/ (левый столбец) и /i/ (правый столбец). Частота основного тона = 80 Гц.

Первая строка – площади голосовой щели.

Вторая строка – объемные скорости без учета (пунктир) и с учетом взаимодействия.

Третья строка – голосовые источники без учета (пунктир) и с учетом взаимодействия.



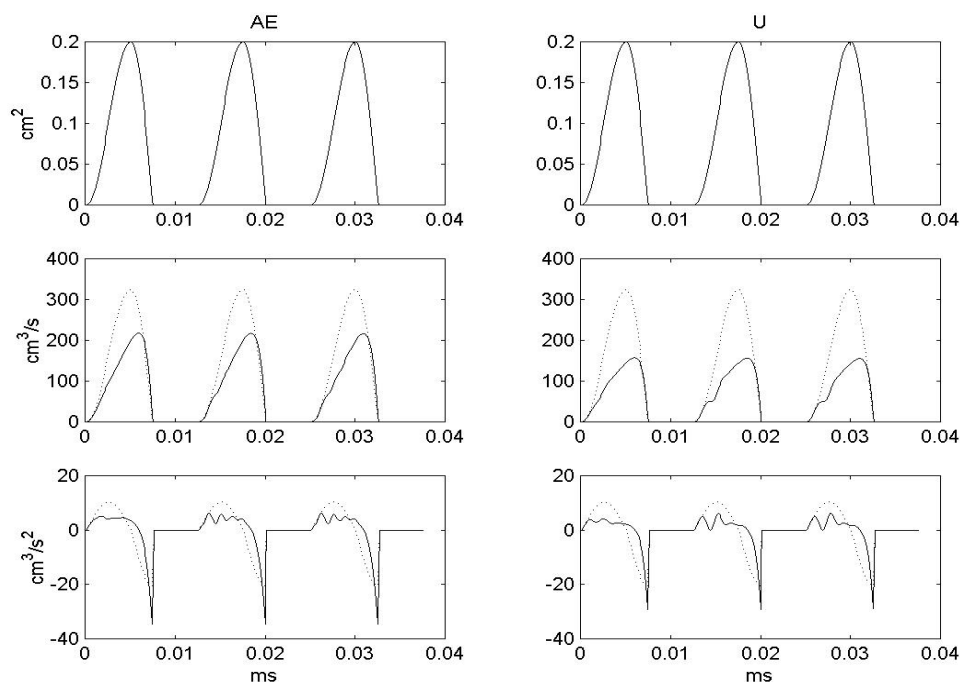


Рис. 5.

Гласные /æ/ (левый столбец) и /u/ (правый столбец). Частота основного тона = 80 Гц.

Первая строка – площади голосовой щели.

Вторая строка – объемные скорости без учета (пунктир) и с учетом взаимодействия.

Третья строка – голосовые источники без учета (пунктир) и с учетом взаимодействия.

Обращает на себя внимание несколько вещей. Во-первых, в случае, если никакое акустическое взаимодействие между объемной скоростью, протекающей через голосовую щель, речевым трактом и подсвязочной областью не присутствует, то форма импульса  $u_g$  в точности повторяет форму площади голосовой щели  $S_g$ . В самом деле, из (1) следует, что в этом случае объемная скорость через голосовую щель оказывается просто пропорциональной площади  $S_g$ . Напротив, при учете акустического взаимодействия формы импульса голосовой щели и объемной скорости, протекающей через голосовые складки, оказываются существенно различными.

Из сравнения импульсов  $u_g$  (вторая строка) видно, что объемная скорость, вычисленная с учетом акустического взаимодействия с речевым трактом и подсвязочной областью, по амплитуде оказывается существенно ниже, чем объемная скорость без взаимодействия. Сама форма импульса становится существенно несимметричной (с более пологим передним и более крутым задним фронтом), при этом на переднем фронте появляются дополнительные осцилляции, называемые складками (особенно они очевидны для гласных /a, i, u/, второй и третий импульсы). Как указывалось во введении, эти явления очень хорошо известны в литературе.

Сравнивая импульсы объемной скорости для одного и того же гласного во времени, мы видим, что при отсутствии взаимодействия форма импульсов не меняется от периода к периоду. Напротив, при наличии взаимодействия форма импульса может претерпевать изменения. Например, для гласных /a, i, u/ на первом периоде никаких дополнительных осцилляций на переднем фронте не наблюдается. Однако уже на втором периоде основного тона они становятся очевидными. Это явление (попериодное изменение формы импульса объемной скорости) также известно в литературе и носит название суперпозиции.

Сравнивая импульсы объемной скорости для разных гласных, мы можем наблюдать, что они отличаются друг от друга. Это связано с тем, что для разных гласных соответствующие акустические импедансы оказываются различными (см. Рис. 1 и 2, третья строка), что и отражается на  $u_g$ . Особенно четко различаются между собой голосовые источники (третья строка) между двумя группами гласных /a, æ/ и /i, u/. Внутри групп различие между голосовыми источниками незначительное. Это объясняется тем, что акустически гласные /a, æ/ четко противопоставлены гласным /i, u/ - первые имеют весьма высокую первую формантную частоту, в то время как для гласных /i, u/ значение первой формантной частоты оказывается весьма низким (см. Рис. 1 и 2, графики передаточных функций).

Рассмотрим теперь результаты экспериментов для тех же гласных звуков и частоты основного тона = 100 Гц (Рис. 6 и 7).

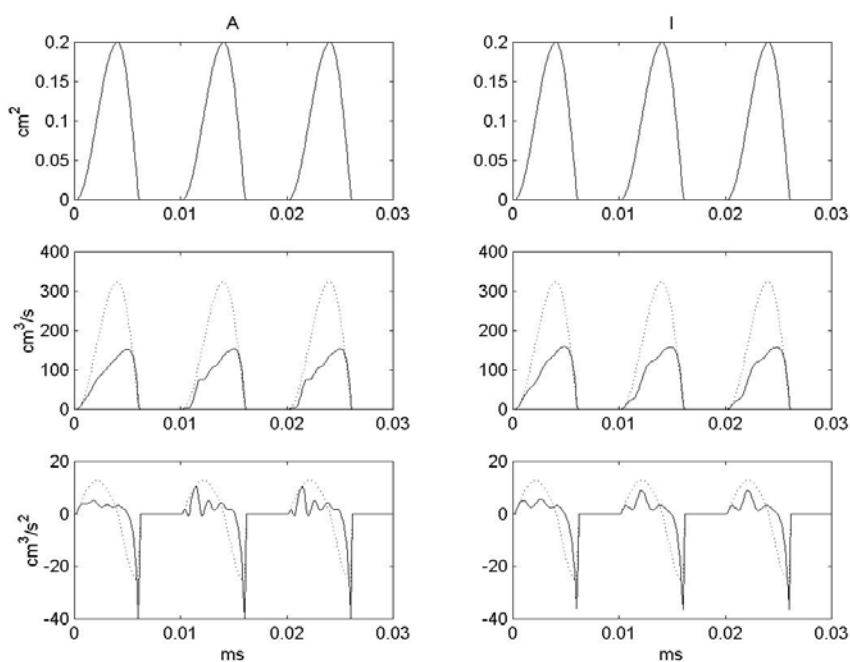


Рис. 6.

Гласные /a/ (левый столбец) и /i/ (правый столбец). Частота основного тона = 100 Гц.

Первая строка – площади голосовой щели.

Вторая строка – объемные скорости без учета (пунктир) и с учетом взаимодействия.

Третья строка – голосовые источники без учета (пунктир) и с учетом взаимодействия.

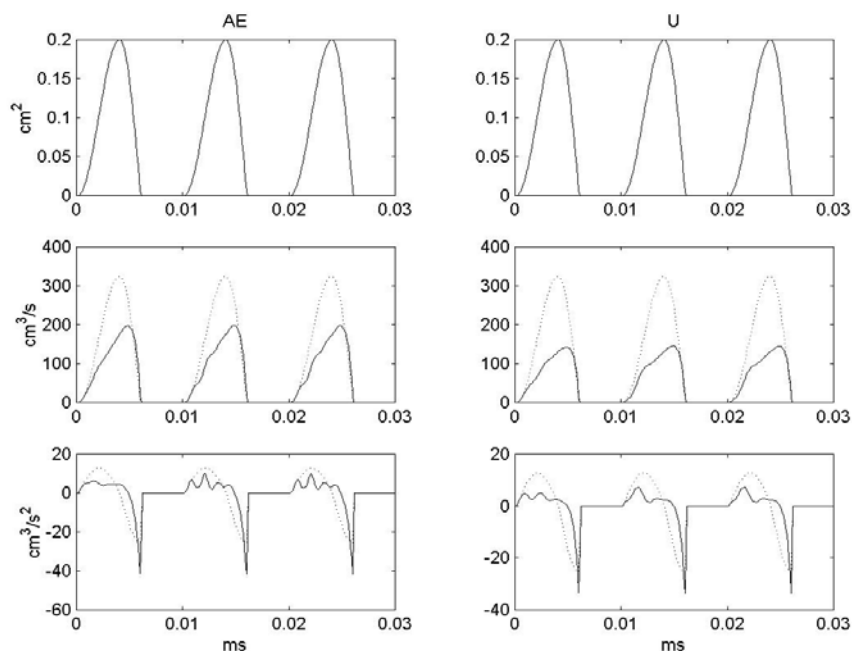


Рис. 7.

Гласные /æ/ (левый столбец) и /u/ (правый столбец). Частота основного тона = 100 Гц.

Первая строка – площади голосовой щели.

Вторая строка – объемные скорости без учета (пунктир) и с учетом взаимодействия.

Третья строка – голосовые источники без учета (пунктир) и с учетом взаимодействия.

На Рис. 6 и 7 наблюдаются те же явления, которые были отмечены выше: несимметричность импульсов объемной скорости, появление дополнительных осцилляций на переднем фронте и поперiodное изменение формы импульса объемной скорости, протекающей через голосовую щель (суперпозиция). Обращает на себя внимание тот факт, что для частоты основного тона = 100 Гц дополнительные осцилляции выражены значительно лучше (/a, i, u/, второй и третий импульсы), чем для частоты основного тона = 80 Гц. Кроме того, для разных гласных форма этих осцилляций (складок) оказывается различной (это особенно очевидно при сравнении импульсов объемной скорости для гласных /a/ и /i/).

Рассмотрим теперь результаты экспериментов для частоты основного тона = 200 Гц (Рис. 8 и 9).

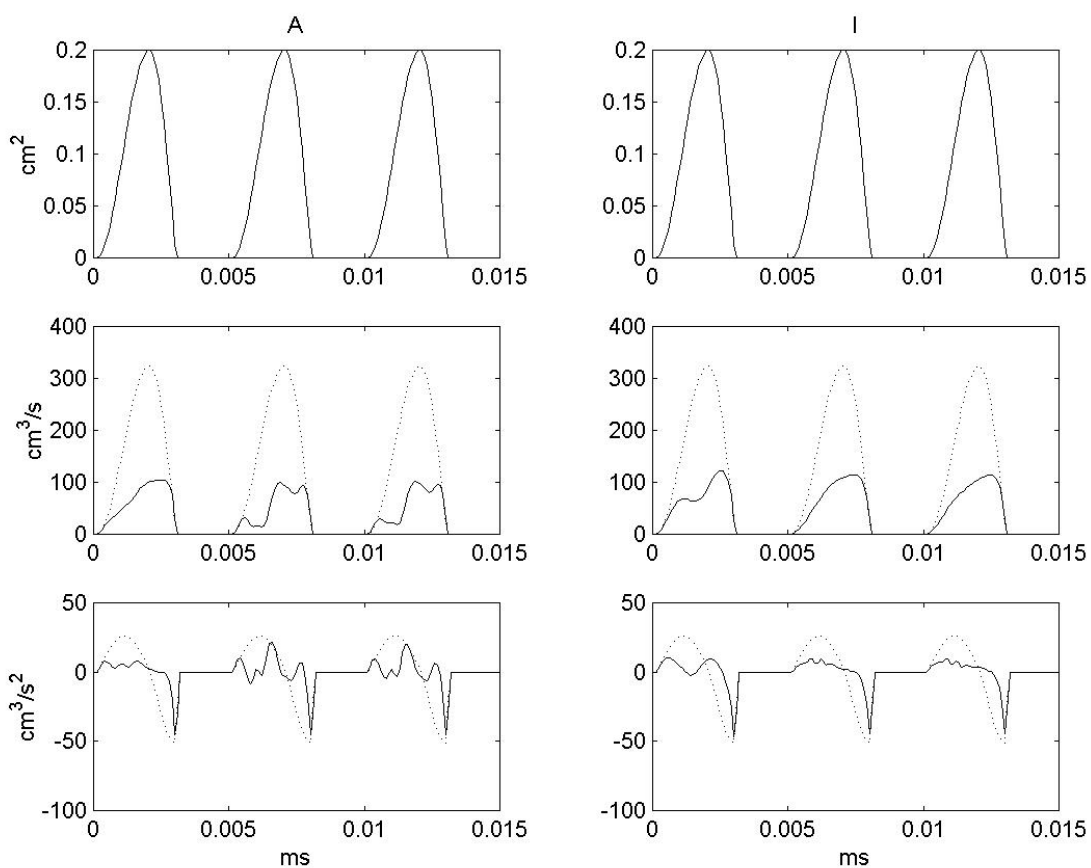


Рис. 8.

Гласные /a/ (левый столбец) и /i/ (правый столбец). Частота основного тона = 200 Гц.

Первая строка – площади голосовой щели.

Вторая строка – объемные скорости без учета (пунктир) и с учетом взаимодействия.

Третья строка – голосовые источники без учета (пунктир) и с учетом взаимодействия.

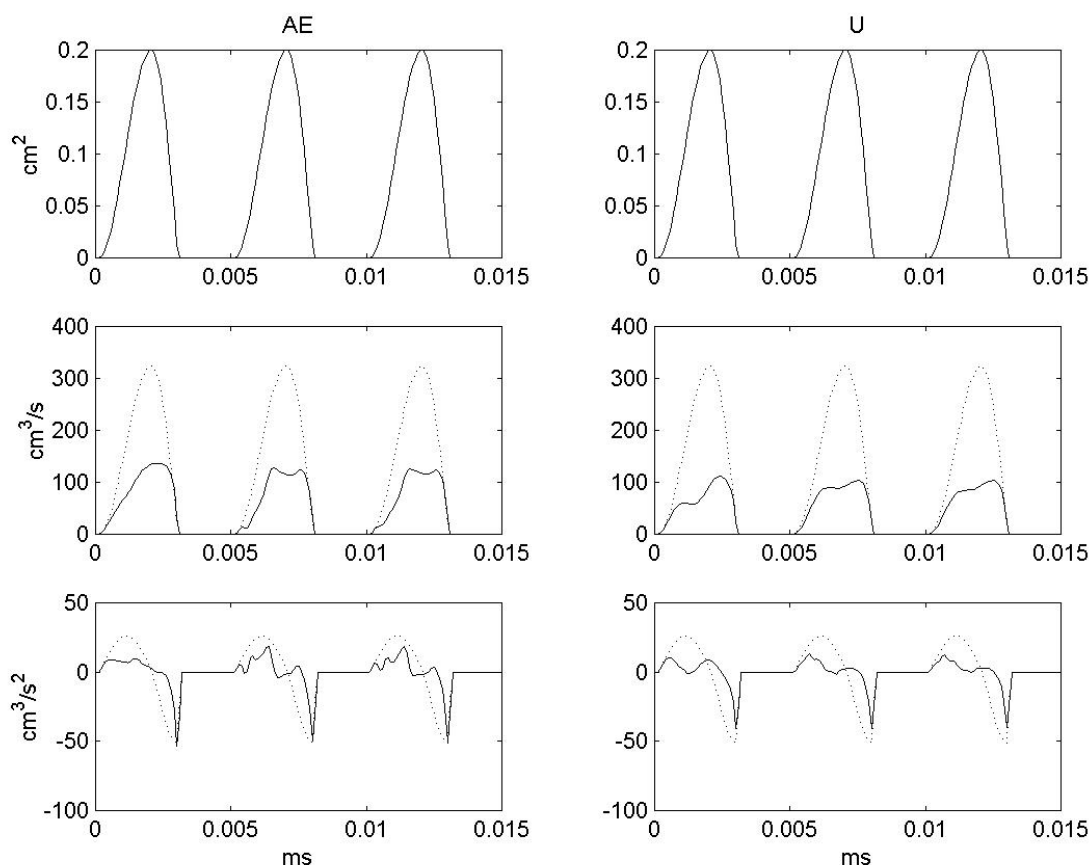


Рис. 9.

Гласные / æ / (левый столбец) и / u / (правый столбец). Частота основного тона = 200 Гц.

Первая строка – площади голосовой щели.

Вторая строка – объемные скорости без учета (пунктир) и с учетом взаимодействия.

Третья строка – голосовые источники без учета (пунктир) и с учетом взаимодействия.

Как и для частот основного тона, равных 80 Гц и 100 Гц, в случае 200 Гц наблюдаются несимметричность импульсов объемной скорости, наличие складок на переднем фронте и суперпозиция (Рис. 8 и 9). Вместе с тем, для данной частоты основного тона эти явления выражены наиболее сильно, причем даже для первого импульса (особенно – для гласных /a, æ/).

В проведенных экспериментах варьировалась только частота основного тона. Предварительные эксперименты показали, что форма объемной скорости, вычисленная при учете акустического взаимодействия, существенно зависит от параметров подсвязочной области. На Рис. 10 (первая строка) показаны импульсы  $u_g$  для разных длин трахеи: 15 см (пунктир) и 20 см (сплошная линия). На том же рисунке (вторая строка) форма объемной скорости вычислена для двух различных радиальных резонансных частот трахеи – 203 Гц (пунктир) и 101 Гц (сплошная линия). Более подробно зависимость объемной скорости, протекающей через голосовую щель, от параметров подсвязочной области исследуется в [16].

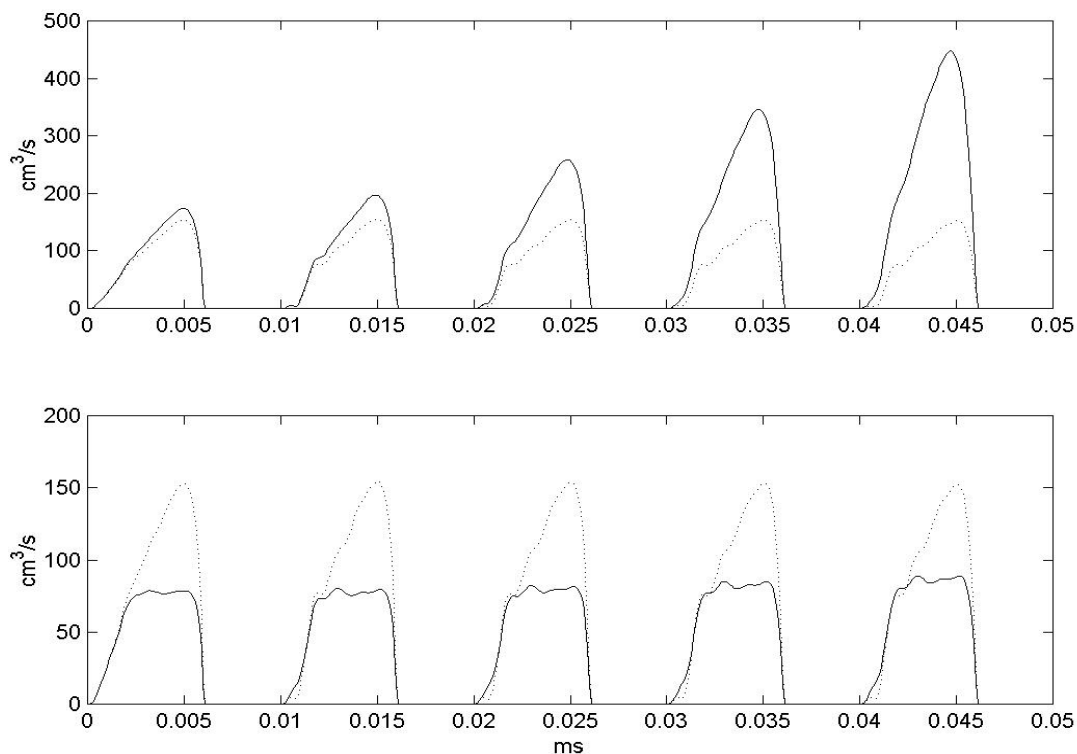


Рис. 10

Гласный /а/. Частота основного тона = 100 Гц.

Первая строка –  $u_g$  для длины трахеи = 15 см (пунктир) и 20 см (сплошная линия).

Вторая строка -  $u_g$  для радиальной частоты стенок трахеи = 203 Гц (пунктир) и 101 Гц (сплошная линия)

### Заключение

Статья посвящена математическому моделированию акустического взаимодействия между речевым трактом, подсвязочной областью (трахея, бронхи и легкие) и объемной скоростью, протекающей через голосовые складки. Построена модель данного взаимодействия, основанная на аппроксимации импульсных импедансных характеристик на входе в речевой тракт и трахею авторегрессионными моделями с подвижным средним. Модель протестирована на функциях площадей поперечных сечений, измеренных с помощью магнитно-резонансной томографии тракта, для трех значений частоты основного тона: 80 Гц, 100 Гц и 200 Гц. Результаты экспериментов хорошо согласуются с данными других работ по моделированию акустического взаимодействия тракта и голосового источника. В частности, импульсы объемной скорости при учете такого взаимодействия демонстрируют осцилляции на переднем фронте, а также поперiodное изменение своей формы (так называемое явление суперпозиции). Помимо этого, было обнаружено, что объемная скорость, протекающая через голосовую щель, существенно зависит от параметров подсвязочной области (геометрические размеры и поверхностные импедансы трахеи, бронхов и легких). Подробный анализ данной зависимости является темой отдельной статьи.

**Библиография**

1. В. Н. Сорокин, Синтез речи. – М.: Наука. 1992. – 392 с.
2. T. Ananthapadmanabha, G. Fant, Calculation of True Glottal Flow and Its Components // STL-QPSR. 1982. P. 1-30.
3. В. Н. Сорокин, Теория речеобразования. – М.: Радио и Связь. 1985. – 312 с.
4. K. Stevens, Acoustic Phonetics. – The MIT Press. 1998. P. 614.
5. J. Allen, W. Strong, A Model for Synthesis of Natural Sounding Vowels // J. Acoust. Soc. Amer. 1985. V. 78. P. 58-69.
6. M. M. Sondhi, J. Schroeter, A Hybrid Time-Frequency Domain Articulatory Speech Synthesizer // IEEE Trans. Acoust., Speech, Signal Process. Vol. ASSP-35. 1987. P. 955-967.
7. Q. Lin, Nonlinear Interaction in Voice Production // STL-QPSR. 1987. P. 1-12.
8. Q. Lin, A Fast Algorithm for Computing the Vocal Tract Impulse Response from the Transfer Function // IEEE Trans. Speech, Audio Process. 1995. Vol. 3. No. 6. P. 449-457.
9. T. Koizumi, S. Taniguchi, and S. Hiromitsu, Glottal Source – Vocal Tract Interaction // J. Acoust. Soc. Amer. 1985. Vol. 78. P. 1541-1547.
10. И. С. Макаров, Аппроксимация речевого тракта коническими рупорами // Акустический Журнал. 2009. Т. 55. № 2. С. 256-265.
11. T. Baer, J. Gore, V. Gracco, and P. Nye, Analysis of Vocal Tract Shape and Dimensions Using Magnetic-Resonance Imaging: Vowels // J. Acoust. Soc. Amer. 1991. Vol. 90. P. 799-828.
12. К.С. Горбунов, И.С. Макаров, Подсвязочная область в артикуляторных синтезаторах // Информационные процессы. 2009. Том 9, № 1. С. 34-43.
13. K. Steiglitz, L. McBride, A Technique for Identification of Linear Systems // IEEE Trans. Automatic Control. Vol. AC-10. 1965. P. 461-464.
14. В. Н. Сорокин, И. С. Макаров, Обратная задача для голосового источника // Информационные процессы. 2006. Т. 6. No. 4. С. 375-395.
15. В. Н. Сорокин, И. С. Макаров, Определение пола диктора по голосу // Акустический Журнал. 2008. Т. 54. No. 4. С. 659-668.
16. К. С. Горбунов, И. С. Макаров, Влияние подсвязочной области на характеристики голосового источника // Речевые Технологии. 2009. В печати.