

УДК 81'342
UDC 81'342

Евдокимова Вера Вячеславовна
Санкт-Петербургский государственный университет
г. Санкт-Петербург, Российская Федерация
Vera V. Evdokimova
St-Petersburg State University, St. Petersburg, Russian Federation
e-mail: postmaster@phonetics.pu.ru

Скрелин Павел Анатольевич
Санкт-Петербургский государственный университет
г. Санкт-Петербург, Российская Федерация
Pavel A. Skrelin
St-Petersburg State University, St. Petersburg, Russian Federation
e-mail: skrelin@phonetics.pu.ru

Евграфова Карина Владимировна
Санкт-Петербургский государственный университет
г. Санкт-Петербург, Российская Федерация
Karina V. Evgrafova
St-Petersburg State University, St. Petersburg, Russian Federation
e-mail: evgrafova@phonetics.pu.ru

Чукаева Татьяна Валерьевна
Санкт-Петербургский государственный университет
г. Санкт-Петербург, Российская Федерация
Tatiana V. Chukaeva
St-Petersburg State University, St. Petersburg, Russian Federation
e-mail: chukaeva68@mail.ru

Швалев Николай Вадимович
Мариинский театр
г. Санкт-Петербург, Российская Федерация
Nikolai V. Shvaley
Mariinsky Theater, St. Petersburg, Russian Federation
e-mail: dr-nix99@mail.ru

ИССЛЕДОВАНИЕ ПРОЦЕССА ФИЛЬТРАЦИИ
ГОЛОСОВОГО СИГНАЛА АРТИКУЛЯТОРНЫМ АППАРАТОМ
INVESTIGATING VOICE SOURCE SIGNAL FILTERING
BY ARTICULATION COMPONENT

Аннотация

Задача получения сигнала голосового источника отдельно от выходного речевого сигнала, а также исследование взаимодействия голосового источника и артикуляторного

аппарата остаются актуальными задачами, решение которых может быть полезно для различных направлений развития речевых технологий. В статье описан метод синхронной записи сигнала голосового источника в области голосовых связок, а также речевого сигнала микрофоном, расположенным вблизи губ. Сравнение двух полученных сигналов на гласных звуках русского языка позволило проанализировать частотные составляющие сигнала на разных стадиях речеобразования. Традиционная модель речевого тракта дополнена звеном обратной связи, что позволяет рассматривать речевой тракт как нелинейную систему. В статье описан метод записи сигнала, перцептивный эксперимент, представлены результаты акустического анализа и метода получения формантных характеристик гласных.

Abstract

The paper is concerned with the specification and improvement of the traditional source-filter model of the human vocal tract proposed by G. Fant and analyzed by many scientists. The new method of recording the glottal wave synchronously with an output speech signal was employed to obtain the experimental material. The comparison of the recorded signals enabled to analyze the structure of the speech signal at different stages of its generation. As a result, the classic vocal tract model was specified by distinguishing a feedback component which formalizes the processes in the vocal tract as a complex acoustic nonlinear system. One of the functions of the component is to transform the acoustic energy from the articulation system upstream. In the paper the recording method is described, the results of the perceptual and acoustic study are presented.

Ключевые слова: фонетика, фонация, голосовой источник, форманты.

Keywords: phonetics, phonation, voice source, formants.

1. Постановка задачи

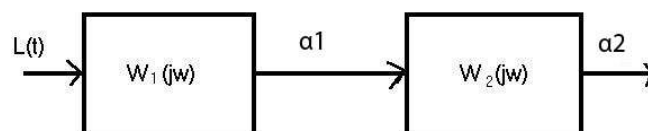
Традиционный подход к исследованию работы речевого тракта человека предлагает рассматривать несколько последовательных стадий его работы: инициализация, фонация, артикуляция и излучение речевого сигнала [Бондарко, 1998, с. 60; Fant, 1960; Кодзасов, 2001, с. 101]. Инициализация рассматривается как иницирующее внешнее воздействие со стороны костно-мышечной и легочной систем человека. На этом этапе под действием костно-мышечной системы поток воздуха через трахею поступает в гортань и энергетически обеспечивает движение голосовых связок в режиме близком к автоколебательному.

Импульс давления воздушного потока в гортани при раздвинутых связках чередуется с его отсутствием при закрытой голосовой щели. Таким образом, формируется во времени периодическая последовательность колоколообразных импульсов давления, получившая наименование глоттальной волны, которая порождает в речевом тракте акустические колебания [Fant, 1985, с. 1, 1997, с. 125]. Частота следования импульсов соответствует частоте основного тона (далее – Ω). Характер создаваемых в речевом тракте скачков давления (импульсов глоттальной волны) имеет сходную

форму у разных людей, но все же индивидуален и зависит от размеров, формы, эластичности, степени натяжения голосовых связок.

После фонации (первичной генерации акустического сигнала) эта последовательность импульсов давления фильтруется артикуляторным аппаратом, активные артикулирующие органы которого перестраиваются мозгом. Осознавая сложность нейрофизиологического управления процессом речеобразования, фонетики, не детализируя процесс нейрофизиологического управления ими, ограничиваются описанием трансформации акустического речевого потока [Бондарко, 1998, с. 60; Кодзасов, 2001, с. 102; Fant, 1960; Stevens, 1998; Flanagan, 1972].

Речевой тракт при таком подходе может рассматриваться как обобщенная динамическая система, в которой выделяются два последовательно соединенных звена, каждое из которых обладает своими динамическими свойствами (рис. 1).



Р и с у н о к 1. Динамическая система речевого тракта человека, состоящая из двух динамических звеньев

П р и м е ч а н и е к р и с у н к у 1. Введены обозначения:

$W(j\omega)$ – частотная передаточная функция голосового источника,

$W(j\omega)$ – частотная передаточная функция артикуляторного аппарата,

$L(t)$ – инициирующее внешнее воздействие,

$\alpha_1(t)$ – выходной сигнал голосового источника,

$\alpha_2(t)$ – выходной речевой сигнал.

Существуют различные модели голосового источника, которые применяются для исследования и использования в автоматических системах. Так, при решении задачи синтеза использовалась схема подачи на вход артикуляторного аппарата последовательности прямоугольных импульсов давления либо просто белого шума, что требовало длительного этапа подгонки речевого сигнала под эталон аппаратными методами [Hallahan, 1995].

Оптимальным вариантом структуры голосового источника стала предложенная на основе обобщения экспериментальных данных LF-модель выходного сигнала голосового источника [Fant, 1985, с. 5, 1997, с. 126]. Эта модель описывает глоттальную волну как последовательность импульсов заданной формы, описываемых набором стандартных параметров. Импульсы следуют с частотой OT . Форма импульсов в целом соответствует профилю экспериментально замеренной глоттальной волны.

В качестве эталона соответствия выбранной формы выходного сигнала голосового источника и, соответственно, входного сигнала артикуляторного аппарата выбрана спектральная плотность голосовой составляющей, выделенная из реального экспериментально полученного речевого сигнала методом обратной фильтрации, то есть путём подавления формантных частот артикуляторного аппарата. Сравнение с эталоном показало, что голосовая составляющая наилучшим образом моделируется по спектру функцией производной глоттальной волны [Fant, 1985, с. 5, 1997, с. 126]. График производной глоттальной волны, описывающий весь цикл работы голосовых связок (открытый и закрытый интервалы), по форме существенно отличен от синусоиды. Это говорит о значительной доле присутствия высших гармоник (тембровых частот). Весомость каждой из них может быть задана формой графиков глоттальной волны и её производной. В LF-модели для настройки предусмотрены четыре специальных безразмерных переменных параметра. Их выбор обеспечивает воспроизведение индивидуальных особенностей голоса [Fant, 1985, с. 5, 1997, с. 126].

Важнейшим отличием LF-модели от других рассмотренных выше моделей является то, что в структуру выходного сигнала голосового источника заложены не только частота Ω_T , но и её высшие гармоники, что обеспечивает базу для исследования взаимодействия голосового источника и артикуляторного аппарата.

Внедрение LF-модели стало основой разработки систем качественного голосового синтеза [Carlson, 1997, p. 402; Carre, 2004, p. 227]. Однако обратная задача определения параметров реального голоса по модели Г. Фанта оказывается слишком громоздкой и требует очень большого объёма вычислений.

Попытки формирования альтернативной LF-модели биомеханической модели речевого аппарата с описанием физического движения голосовых связок и аэродинамических процессов, с учётом строения мышц и тканей голосового аппарата [Леонов, 2005, с. 101; Макаров, 2004, с. 389], оказываются предельно громоздки и целесообразны только при решении специфических задач, например при изучении патологий голосового аппарата [Howe, 2010, p. 1; Mergell, 1997, p. 141; Zanartu, 2007, p. 1119]. Несмотря на сложность, они тоже не могут претендовать на полноту описания, так как не учитывают психофизиологическую сторону процесса.

Безусловно, речевой тракт является нелинейной самонастраивающейся системой. Линеаризация его модели в рамках традиционной теории речеобразования с разделением на два основных отдела – голосовой источник и артикуляторный аппарат – является вынужденным шагом для изучения основных свойств его работы. Зафиксированные и исследованные факты взаимодействия двух частей речевого тракта дают возможность развития и дополнения классической теории речеобразования. В некоторых работах описывается это взаимодействие и то, как оно может проявляться [Hatzikirou, 2006, p. 468; Miller, 2005, p. 268; Titze, 2008, p. 2733]. Часто эти данные носят фрагментарный характер и требуют дальнейшего

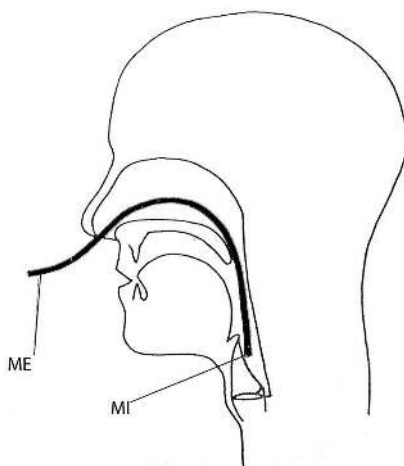
уточнения. Так, отмечено, что взаимодействие источника и фильтра происходит в речи мужчин и большей степени, чем в речи женщин и детей. Описаны сложные случаи взаимодействия для певческих голосов [Titze, 2008, p. 2733].

Таким образом, можно считать, что задача получения сигнала голосового источника отдельно от выходного речевого сигнала, а также исследование взаимодействия голосового источника и артикуляторного аппарата остаются актуальными задачами, решение которых может быть полезно для различных направлений развития речевых технологий.

2. Эксперимент

2.1. Методика записи

Запись проходила в студии звукозаписи с высокой шумозащитой Кафедры фонетики и методики преподавания иностранных языков Филологического факультета СПбГУ. После предварительного осмотра с помощью риноларингофиброскопа горла и носовых пазух микрофон QueAudio ($d=2.3$ мм, водозащита) был размещен в области голосовых связок диктора (MI). Процедуру проводил врач-фониатр Мариинского театра Н. В. Швалев. Сигнал с него и речевой сигнал с микрофона AKG HSC200, размещенного около губ диктора снаружи (ME), поступали на многоканальную систему записи звука MOTU Traveler (см. рис. 2).



Р и с у н о к 2 . Местоположение микрофонов при записи

Примечание к рисунку 2. MI – внутренний микрофон, расположенный в области голосовых связок (QueAudio, $d=2.3$ мм, водозащита). ME – внешний микрофон (AKG HSC200), размещенный около губ диктора снаружи.

Была проведена синхронная запись речевого материала длительностью 15 минут для двух дикторов (1 женский голос и 1 – мужской). Из собранного материала были сегментированы синхронно снятые участки гласных /i/, /e/, /a/, /o/, /u/, /ɨ/, произнесённые изолированно с нисходящей интонацией.

2.2. Перцептивный анализ

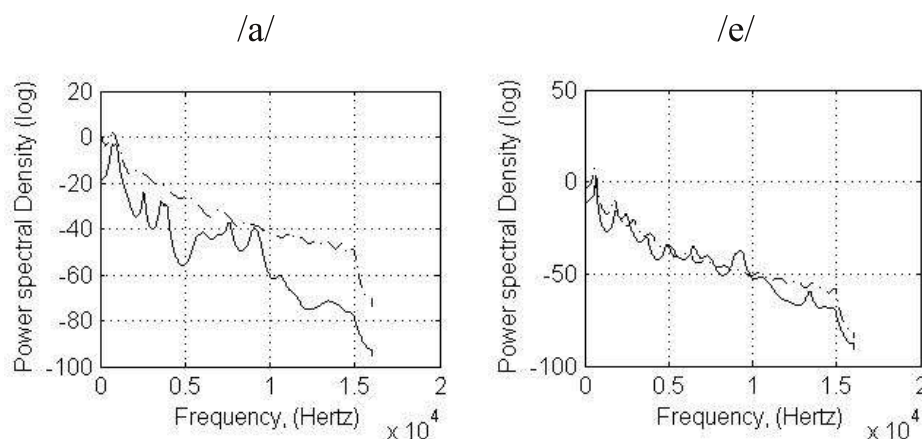
На основе полученного материала был проведён перцептивный эксперимент с той его частью, которая была записана на внутренний микрофон. Были отобраны две группы слушателей: первая группа – 5 человек были специалистами по фонетике, вторая – 20 человек – наивными носителями русского языка. Гласные звуки были предъявлены участникам в произвольном порядке. Вопрос звучал следующим образом: Можете ли Вы идентифицировать данные звуки как какой-либо гласный русского языка?

Результаты эксперимента не подтвердили предположение о том, что все звуки должны звучать сходным образом и включать в себя только сигнал голосового источника (это противоречит классической теории речеобразования, которая считает, что формантные характеристики образуются в надгортанных полостях речевого тракта и не влияют на голосовой источник). Группа фонетистов идентифицировала все предложенные звуки как принадлежащие какой-либо фонеме русского языка, однако, не все звуки были правильно отнесены к звучащим на выходе речевого тракта реализациям фонем. Группа наивных носителей также сделала много ошибок в идентификации сигнала. Кроме того, среди ответов были такие как: Я не могу отнести этот звук к какой-либо фонеме русского языка вообще.

Чаще всего правильно идентифицированы были гласные /a/, /e/, /i/. Однако часто /i/ идентифицировали как /u/, /i/ как /i/ и /u/ и /i/. Гласные /a/ и /i/ часто воспринимались как огубленные.

2.3. Акустический анализ

Собранная информация была обработана и представлена в форме спектрограмм сигналов (см. рис. 3–5).

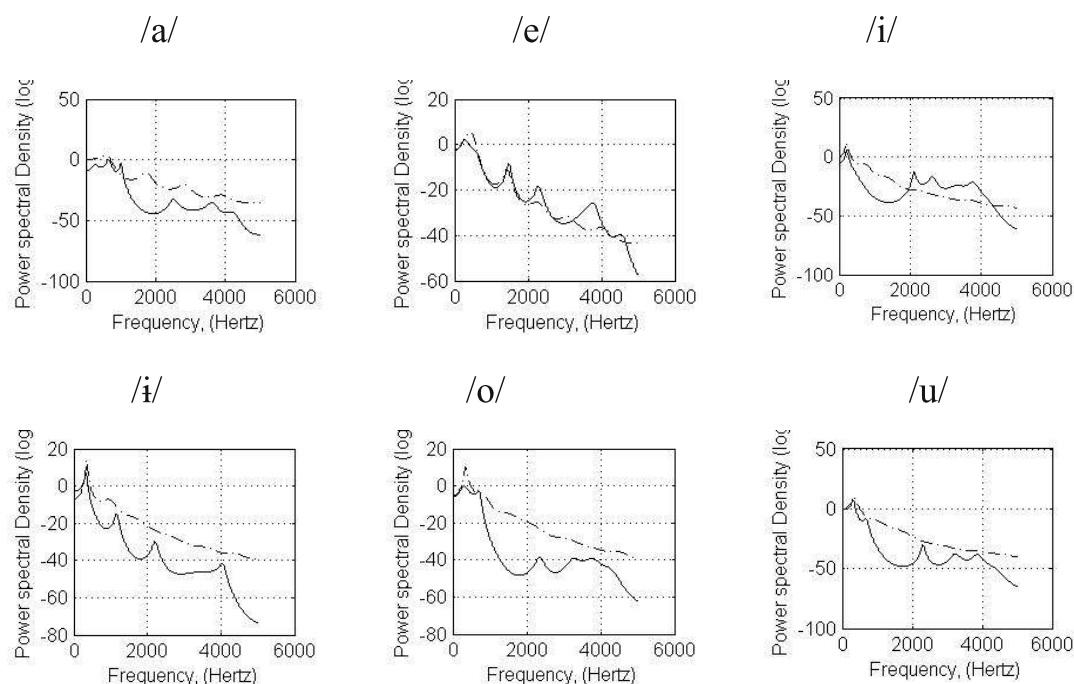


Р и с у н о к 3. Спектральные плотности сигналов записанных на внутреннем (пунктирная линия) и внешнем микрофоне (сплошная линия) для гласных /a/ и /e/

П р и м е ч а н и е к р и с у н к у 3. Спектр представлен до 16 кГц (ОХ – шкала частот в Герцах, ОУ – спектральная плотность (log)).

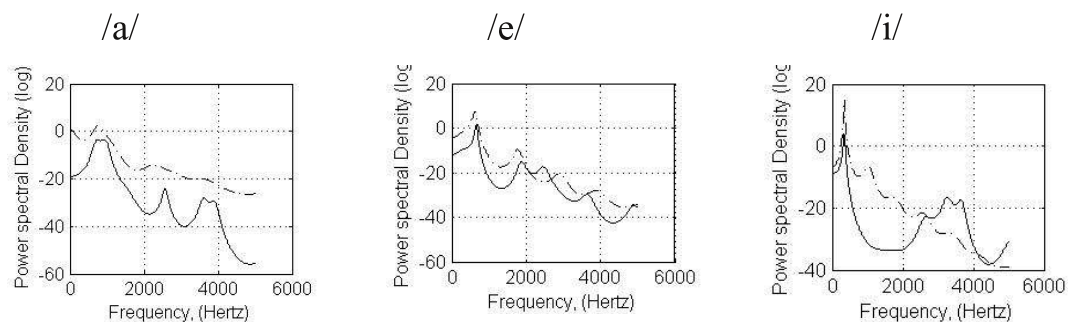
На рисунках видно, что разница в сигналах двух микрофонах видна не только в диапазоне формант, но и в более высокочастотном диапазоне. Однако в данном исследовании нас интересовал именно формантный диапазон гласных.

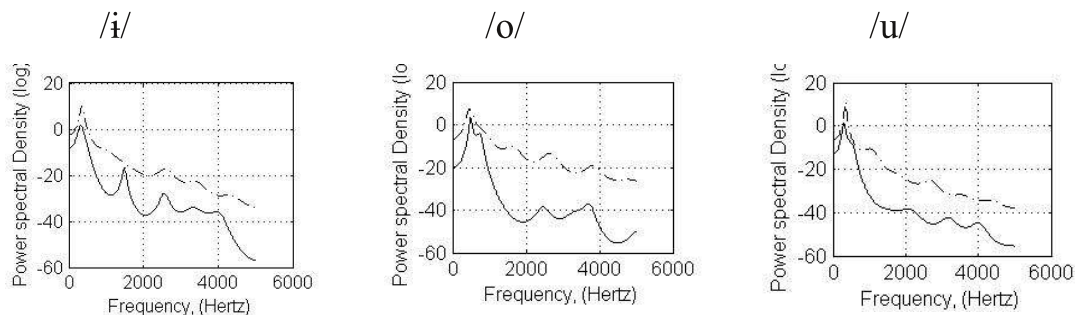
Рисунки 4 и 5 представляют примеры для сравнения сигналов, полученных с двух микрофонов для разных гласных для мужского и женского голосов. Представлено по одной реализации каждой фонемы в частотном диапазоне до 5 кГц.



Р и с у н о к 4. Спектральные плотности сигналов записанных на внутреннем (пунктирная линия) и внешнем микрофоне (сплошная линия) для разных гласных

Примечание к рисунку 4. Сигналы произведены женским голосом. Спектр представлен до 5 кГц (ОХ – шкала частот в Герцах, ОУ – спектральная плотность (log)).





Р и с у н о к 5. Спектральные плотности сигналов записанных на внутреннем (пунктирная линия) и внешнем микрофоне (сплошная линия) для разных гласных

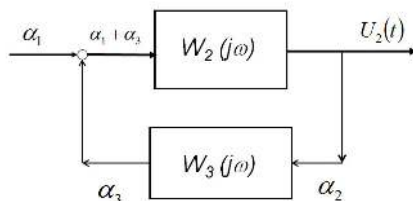
Примечание к рисунку 5. Сигналы произведены мужским голосом. Спектр представлен до 5 кГц (ОХ – шкала частот в Герцах, ОУ – спектральная плотность (log)).

Отметим присутствие формантных составляющих не только в речевом сигнале, но и в полости гортани. Однако иногда графики могут быть очень разными (например, для гласного /i/).

2.4. Моделирование

Собранная дополнительная информация дала возможность более детального изучения динамики речевого тракта.

Существует предложенная Фантом общепринятая динамическая модель структуры речевого тракта, в которой выделены последовательно действующие голосовой источник и артикуляторный аппарат. Размещение дополнительного микрофона в гортани позволяет детализировать эту модель и ввести звено обратной связи (см. рис. 6).



Р и с у н о к 6. Артикуляторный аппарат речевого тракта с введенным звеном обратной связи

Примечание к рисунку 6. Введены обозначения:

α_1 – глоттальная волна,

α_2 – речевой сигнал,

$U_2(t)$ – выходной речевой сигнал,

α_3 – отраженная обратно акустическая энергия,

$W_2(j\omega)$ – частотная передаточная функция артикуляции,

$W_3(j\omega)$ – частотная передаточная функция звена обратной связи.

В ней артикуляторный аппарат представляется как звено с обратной связью. Это звено формализует отражение части акустической энергии сформированного артикуляторным аппаратом речевого сигнала в гортань и объясняет присутствие формантных составляющих в гортани.

Кроме детализации процессов в речевом тракте собранный экспериментальный материал даёт возможность подтверждения предложенного ранее метода построения формантных картин гласных путём совместной обработки речевых сигналов различной длительности. Опишем этот метод.

В общем случае речевой сигнал может рассматриваться как смесь из акустических колебаний частоты ОТ и тембровых частот с одной стороны и колебаний со сформированными артикуляторным аппаратом, как фильтрующим звеном, частотами, характерными для конкретных фонем – с другой. Степень присутствия каждой из этих групп составляющих может быть изменена при статистической обработке исходного аудиоматериала путём выбора участка и длительности обрабатываемого сегмента. Если сегментировать участок действия конкретной фонемы, то присутствие в спектрограмме составляющих её собственных частот максимально.

С увеличением длительности обрабатываемого сегмента до фразы, нескольких фраз и далее до всего сообщения степень присутствия в спектрограмме частотных составляющих конкретной фонемы быстро падает из-за статистического усреднения. В то же время, составляющие частот голосового источника присутствуют в речевой посылке почти постоянно (кроме глухих согласных и пауз). Их интенсивность почти не ослабевает. В результате для достаточно длинной речевой посылки (в расчётах бралось несколько минут) можно прийти к полностью сглаженной на высокочастотном участке спектральной плотности выходного сигнала $U(t)$, сохранившей, в то же время, пики на частотах основного тона и её высших гармоник (если они были). Приблизительно такая спектральная плотность, в которой подавлено действие артикуляторного аппарата, может рассматриваться как спектральная плотность голосового сигнала $U(t)$.

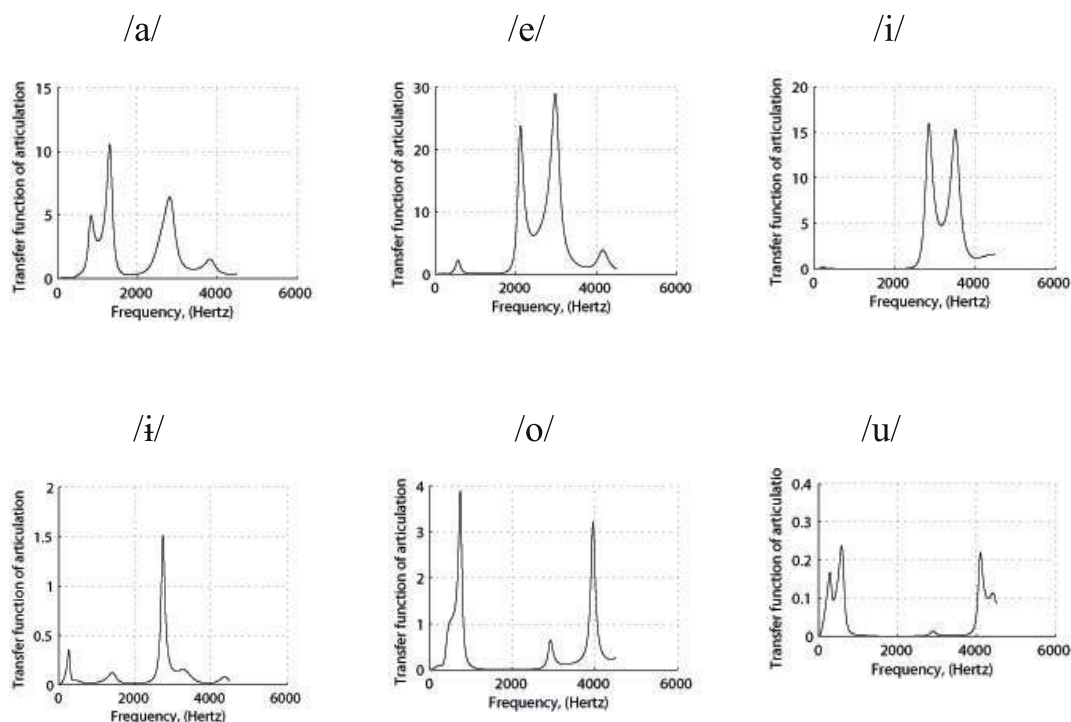
В отличие от частотной фильтрации описанный приём не вводит жёсткой частотной границы фильтрации и позволяет сохранить в сигнале $U(t)$ как составляющую частоты основного тона, так и составляющие тембровых частот.

Вторым сравниваемым участком выбираем сегмент конкретной гласной фонемы.

Остаётся, опираясь на описанную выше двухэлементную модель речевого тракта, определить соотношение полученных сигналов [Евдокимова, 2005, с. 74; Сергиенко, 2003, с. 249; Evdokimova, 2006, p. 210].

Результаты расчётов приведены на рисунке 7.

Сравнение кривых, полученных при обработке сигналов только из речевой посылки и при использовании длительного сигнала из гортани, подтверждает их почти полную идентичность и соответствие реальным формантным пикам на спектральной плотности сегмента соответствующей гласной.



Р и с у н о к 7. Передаточные функции артикуляторного аппарата при произнесении гласных /a/, /e/, /i/, /i/, /o/, /u/

П р и м е ч а н и е к р и с у н к у 7. Сигналы произведены женским голосом. Формантная структура хорошо видна и соответствует фонетическим данным по местоположению формант гласных русского языка. кГц (ОХ – шкала частот в Герцах, ОУ – спектральная плотность).

3. Выводы

1. Описанный эксперимент по синхронной записи речевого и голосового сигнала с двух микрофонов подтверждает явление отражения от артикулирующих элементов части акустической энергии речевой посылки обратно в гортань и наличие, таким образом, обратных связей в работе верхнего отдела речевого тракта. На основании этого в работе вводится дополнительное звено обратной связи в модель артикуляторного аппарата, что дополняет и развивает традиционную теорию речеобразования.

2. Результаты обработки экспериментальных данных подтверждают применимость метода совместной обработки речевых сигналов разной длительности для определения формантных картин гласных с высокой надежностью.

3. Дополненная модель артикуляторного аппарата полностью соответствует фонетическим принципам. Она позволяет уточнить существующие модели речеобразования и может быть использована для решения некоторых задач речевых технологий.

Благодарности

Работа проводилась при поддержке гранта СПбГУ «Фонетические аспекты синтеза речевого сигнала высокого уровня естественности», шифр 31.37.353.2015. Исследования были проведены с использованием вычислительных ресурсов Ресурсного Центра «Вычислительный центр СПбГУ» (<http://cc.spbu.ru>).

Список литературы

1. Бондарко, Л. В. Фонетика современного русского языка : учеб. Пособие [Текст] / Л. В. Бондарко. — СПб. : Изд-во СПбГУ, 1998. — 276 с.
2. Евдокимова, В. В. Выбор метода построения модели речевого тракта [Текст] / В. В. Евдокимова // Интегральное моделирование звуковой формы естественных языков : сб. статей / СПбГУ. Ин-т филологических исслед. — СПб. : Изд-во СПбГУ, 2005. — С. 74–78.
3. Кодзасов, С. В. Общая фонетика / С. В. Кодзасов, О. Ф. Кривнова. — М. : Изд-во РГГУ, 2001. — 592 с.
4. Кодовая книга для речевых обратных задач [Текст] / А. С. Леонов, И. С. Макаров, В. Н. Сорокин, А. И. Цыпликин // Информационные процессы. — 2005. — Т. 5. — № 2, — с. 101–109.
5. Макаров, И. С. Резонансы разветвленного речевого тракта с податливыми стенками [Текст] / И. С. Макаров, В. Н. Сорокин // Акустический журнал. — 2004. — Т. 50. — № 3. — С. 389–396.
6. Сергиенко, А. Б. Цифровая обработка сигналов [Текст] / А. Б. Сергиенко. — СПб. : Питер, 2003. — 603 с.
7. Carre, R. From acoustic tube to speech production [Text] / R. Carre // Speech Communication. — 2004. — N 42. — P. 227–240.
8. Carlson, R. Experiments with voice modeling in speech synthesis [Text] / R. Carlson, B. Granstrom, I. Karlsson // Speech Communication. — 1991. — N 10. — P. 481–489.
9. Evdokimova, V. V. The use of vocal tract model for constructing the vocal structure of the vowels [Text] / V. V. Evdokimova // SPECOM'2006, Saint-Petersburg, 25–29 June 2006. — P. 210–214.
10. Fant, G. Acoustic Theory of Speech Production [Text] / G. Fant. — Netherlands: Mouton, 1960. — 321 p.
11. Fant, G. A four-parameter model of glottal flow [Text] / G. Fant, J. Liljencrants, Q. Lin. — STL-QPSR. — 1985. — P. 2–3.
12. Fant, G. The voice source in the connected speech [Text] / G. Fant // Speech Communication. — 1997. — N 22. — P. 125–139.
13. Stevens, K. Acoustic Phonetics [Text] / K. Stevens. — Cambridge, MA 02141: The MIT Press, 1998. — 624 p.
14. Flanagan, J. L. Speech Analysis, Synthesis, and Perception [Text] / J. L. Flanagan. — Springer, New York, 1972. — 446 p.
15. Hallahan, W. I. DECTalk Software: Text-to-Speech Technology and Implementation [Text] / W. I. Hallahan // COMPAQ DIGITAL Technical Journal. — 1995. — V. 7. — N 5. — P. 5–19.

16. Hatzikirou, H. Voice instabilities due to source-tract interactions [Text] / H. Hatzikirou, W. T. S. Fitch, H. Herzel // *Acta. Acust. Acust.* 92, – 2006. – P. 468–475.
17. Howe, M. S. On the single-mass model of the vocal folds [Electronic resource] / M. S. Howe, R. S. McGowan // *Fluid Dyn. Res.* 42. – 2010. – 015001. doi: 10.1088/0169-5983/42/1/015001. – P. 1–16.
18. Miller, D. G. ‘Mixing’ the registers: Glottal source or vocal tract? [Text] / D. G. Miller, H. K. Schutte // *Folia Phoniatr Logop* 57. – 2005. – P. 278–291.
19. Mergell, P. Modeling biphonation –The role of the vocal tract [Text] / P. Mergell, H. Herzel // *Speech Commun.* 22. – 1997. – P. 141–154.
20. Titze, I. R. Non-linear source-filter coupling in phonation: Theory [Text] / I. R. Titze // *J. Acoust. Soc. Am.* – 2008. – Vol. 123. – P. 2733–2749.
21. Titze, I. R. Nonlinear source-filter coupling in phonation: Vocal exercises [Text] / I. R. Titze, T. Riede, P. Popolo // *J. Acoust. Soc. Am.* – 2008. – Vol. 123. – P. 1902–1915.
22. Zanartu, M. Influence of acoustic loading on an effective single mass model of the vocal folds [Text] / M. Zanartu, L. Mongeau, G. R. Wodicka // *J. Acoust. Soc. Am.* – 2007. – Vol. 121. – P. 1119–1129.

References

1. Bondarko, L. V. *Fonetika sovremennogo russkogo jazyka : ucheb. Posobie* [Tekst] / L. V. Bondarko. — SPb. : Izd-vo SPbGU, 1998. – 276 s.
2. Evdokimova, V. V. *Vybor metoda postroenija modeli rechevogo trakta* [Tekst] / V. V. Evdokimova // *Integral'noe modelirovanie zvukovoj formy estestvennyh jazykov : sb. statej / SPbGU. In-t filologicheskikh issled.* – SPb. : Izd-vo SPbGU, 2005. – S. 74–78.
3. Kodzasov, S. V. *Obshhaja fonetika* / S. V. Kodzasov, O. F. Krivnova. – M. : Izd-vo RGGU, 2001. – 592 s.
4. *Kodovaja kniga dlja rechevyh obratnyh zadach* [Tekst] / A. S. Leonov, I. S. Makarov, V. N. Sorokin, A. I. Cyplikin // *Informacionnye processy.* – 2005. – T. 5. – № 2, – s. 101–109.
5. Makarov, I. S. *Rezonansy razvetvlenogo rechevogo trakta s podatlivymi stenkami* [Tekst] / I. S. Makarov, V. N. Sorokin // *Akusticheskij zhurnal.* – 2004. – T. 50. – № 3. – S. 389–396.
6. Sergienko, A. B. *Cifrovaja obrabotka signalov* [Tekst] / A. B. Sergienko. – SPb. : Piter, 2003. – 603 s.
7. Carre, R. From acoustic tube to speech production [Text] / R. Carre // *Speech Communication.* – 2004. – N 42. – P. 227–240.
8. Carlson, R. Experiments with voice modeling in speech synthesis [Text] / R. Carlson, B. Granstrom, I. Karlsson // *Speech Communication.* – 1991. – N 10. – P. 481–489.
9. Evdokimova, V. V. The use of vocal tract model for constructing the vocal structure of the vowels [Text] / V. V. Evdokimova // *SPECOM'2006, Saint-Petersburg, 25–29 June 2006.* – P. 210–214.
10. Fant, G. *Acoustic Theory of Speech Production* [Text] / G. Fant. – Netherlands: Mouton, 1960. – 321 p.

11. Fant, G. A four-parameter model of glottal flow [Text] / G. Fant, J. Liljencrants, Q. Lin. – STL-QPSR. – 1985. – P. 2–3.
12. Fant, G. The voice source in the connected speech [Text] / G. Fant // *Speech Communication*. – 1997. – N 22. – P. 125–139.
13. Stevens, K. *Acoustic Phonetics* [Text] / K. Stevens. – Cambridge, MA 02141: The MIT Press, 1998. – 624 p.
14. Flanagan, J. L. *Speech Analysis, Synthesis, and Perception* [Text] / J. L. Flanagan. – Springer, New York, 1972. – 446 p.
15. Hallahan, W. I. *DECTalk Software: Text-to-Speech Technology and Implementation* [Text] / W. I. Hallahan // *COMPAQ DIGITAL Technical Journal*. – 1995. – V. 7. – N 5. – P. 5–19.
16. Hatzikirou, H. Voice instabilities due to source-tract interactions [Text] / H. Hatzikirou, W. T. S. Fitch, H. Herzel // *Acta. Acust. Acust.* 92, – 2006. – P. 468–475.
17. Howe, M. S. On the single-mass model of the vocal folds [Electronic resource] / M. S. Howe, R. S. McGowan // *Fluid Dyn. Res.* 42. – 2010. – 015001. doi: 10.1088/0169-5983/42/1/015001. – P. 1–16.
18. Miller, D. G. ‘Mixing’ the registers: Glottal source or vocal tract? [Text] / D. G. Miller, H. K. Schutte // *Folia Phoniatr Logop* 57. – 2005. – P. 278–291.
19. Mergell, P. Modeling biphonation –The role of the vocal tract [Text] / P. Mergell, H. Herzel // *Speech Commun.* 22. – 1997. – P. 141–154.
20. Titze, I. R. Non-linear source-filter coupling in phonation: Theory [Text] / I. R. Titze // *J. Acoust. Soc. Am.* – 2008. – Vol. 123. – P. 2733–2749.
21. Titze, I. R. Nonlinear source-filter coupling in phonation: Vocal exercises [Text] / I. R. Titze, T. Riede, P. Popolo // *J. Acoust. Soc. Am.* – 2008. – Vol. 123. – P. 1902–1915.
22. Zanartu, M. Influence of acoustic loading on an effective single mass model of the vocal folds [Text] / M. Zanartu, L. Mongeau, G. R. Wodicka // *J. Acoust. Soc. Am.* – 2007. – Vol. 121. – P. 1119–1129.