

DOI: 10.34220/MAMSP_15-20

УДК 004.451.62, 004.451.644, 004.056.53

**ОДИН ИЗ ПОДХОДОВ ДЛЯ ОПРЕДЕЛЕНИЯ РЕЛЕВАНТНЫХ
ХАРАКТЕРИСТИК АУДИОСИГНАЛОВ НА ПРИМЕРЕ
ИССЛЕДОВАНИЯ МЕЖДОМЕТИЙ**

В.И. Анциферова¹, Т.В. Песецкая¹, И.И. Юлдошев¹, Лу Сянян¹, Ван Цин¹,
В.В. Лавлинский¹

¹ФГБОУ ВО «Воронежский государственный лесотехнический
университет имени Г.Ф. Морозова»

В данной статье описан один из подходов для определения релевантных характеристик аудиосигналов на примере исследований междометий. Сравнительная оценка проводится на основе технических характеристик аудиосигналов, записанных на диктофон междометий, являющихся, во-первых, одинаковыми по семантическому анализу, во-вторых, близкими по звучанию друг с другом, но произносимые с разной интонацией.

Ключевые слова: аудиосигналы, междометия, семантический анализ, системный подход, цифровая обработка сигнала.

**ONE OF THE TRIPS TO DETERMINE THE RELEVANT
CHARACTERISTICS OF AUDIO SIGNALS ON THE EXAMPLE
OF INTERJECTION RESEARCH**

V.I. Antsiferova¹, T.V. Pesetskaya¹, I.I. Yuldoshev¹, Lu Xianyang¹, Wang Qing¹,
V.V. Lavlinskiy¹

¹Voronezh State University of Forestry and Technologies named after G.F. Morozov

This article describes one of the approaches for determining the relevant characteristics of audio signals on the example of interjection studies. The comparative evaluation is based on the technical characteristics of the audio signals recorded on the dictaphone of interjections, which are, first, the same in semantic analysis, and second, close in sound to each other, but pronounced with different intonation.

Keywords: audio signals, interjections, semantic analysis, system approach, digital signal processing.

Исследование междометий в области цифровой обработки сигналов является одной из важной составляющей в изучении эмоций человека на уровне сигналов. Тем не менее, не очень много работ посвящено данным исследованиям с точки зрения изучения технических аспектов аудиосигналов междометий. Ввиду этого исследование характеристик аудиосигналов на примерах междометий для различных языков является своевременной и актуальной, а разработка методики сравнительной оценки релевантных характеристик аудиосигналов междометий на различных языках позволит лингвистам использовать средство для более глубокого изучения междометий и давать их сравнительную оценку, которая позволит относить сами стохастические аудиосигналы междометий к тому или иному закону распределения.

В работе [1] авторы рассматривают возможность формирования феноменологической математической модели психоакустики слуха человека. При формировании данной модели авторы уделяют внимание на параметры сигналов, которые воспринимает ухо человека. Эта особенность необходима, чтобы учитывать параметры аудиосигналов междометий. Так, например, авторы представляют в своей работе зависимости, которые оказывают гармонические колебания определённой частоты и с определёнными частотными характеристиками сигналов на мембрану уха человека (рисунки 1 и 2 соответственно [1]).

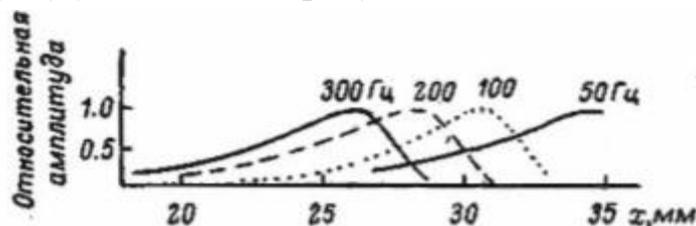


Рисунок 1 – Влияние гармонических колебаний определённой частоты на мембрану уха человека

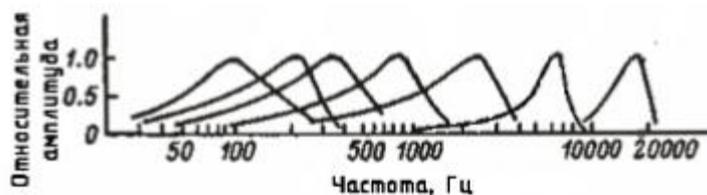


Рисунок 2 – Влияние частотных характеристик сигналов на мембрану уха человека

Также ещё одним параметром, который исследуется в работе [1] является высота тона. Данный параметр определяется на основе использования интеллектуальной системы человека. Высота тона измеряется в Мел.

Зависимости высоты тона и частоты, приведённые автором, представлены на рисунке 3 [1]. Данную особенность следует учесть при исследовании меж-

дометий, так как междометия отличаются друг от друга по эмоциональному окрасу из-за различного использования длительности звучания гласных и согласных букв для различных языков мира.

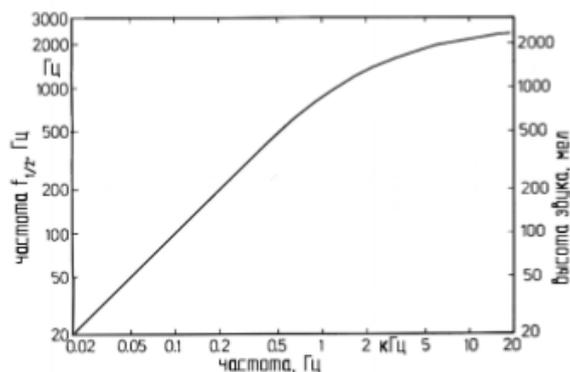


Рисунок 3 – Зависимость высоты тона, частоты от середины высоты звука

Ещё одним из параметров, который выделяется автором в работе [1] является громкость. Эта зависимость представлена на рисунке 4. Тем не менее, последняя характеристика сигнала определяет исключительно параметр, который присущ уху человека. Поэтому его использование не является обязательным при исследовании аудиосигналов междометий. В работе [5] авторами исследуются параметры сигналов для распознавания речи. Так основным методом обработки речевых сигналов является быстрое преобразование Фурье (БПФ). В этом случае авторы выполняли БПФ за интервальный промежуток 25 мс с временным окном в 100 п/с.

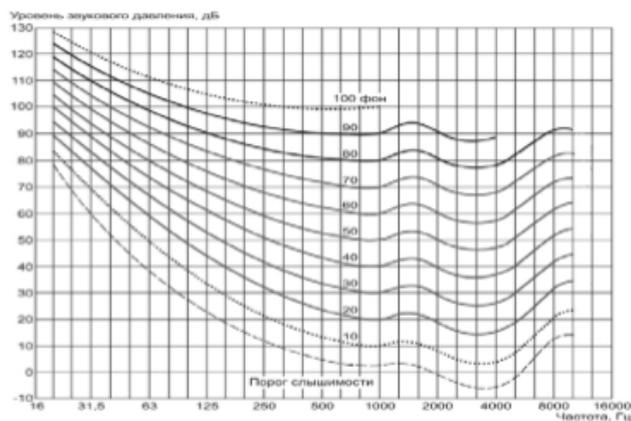


Рисунок 4 – Аппроксимация экспериментальных кривых равной громкости гармонических сигналов различной частоты

В этом случае будет сохранена энергия соседних частот. Причём следует отметить, что эта энергия, на основе мел-частотных фильтров, сохраняется в пределах каждого кадра. Используя логарифм, применительно к выходам фильтров, а также применительно к логарифмическим тоновым спектрам, можно осуществить декорреляцию с помощью дискретного косинусного преобразова-

ния. В этом случае формируется 13-мерный вектор кепстральных коэффициентов тоновой частоты (MFCC). Также коэффициенты MFCC могут быть заменены коэффициентами линейного перцепционного предсказания (PLP), что является более шумоустойчивым представлением сигнала. С помощью БПФ имеется возможность получения спектра аудиосигнала междометий и выполнения его спектрального анализа.

В работе [2] авторы представляют речь на основе следующих характеристик, разбитых на группы:

- семантическая или смысловая. Эта группа характеризует смысл речи на основе понятий, передающихся при помощи речевого сообщения;
- фонетическая. Эта группа характеризует речь на основе её звукового состава;
- интонационная. Эта группа определяет стиль речи, ее эмоциональную окраску и даёт возможность различать незнакомые языки;
- физическая. Эта группа включает величины и зависимости, характеризующие речь как акустический сигнал.

Кроме того, авторы исследовали спектральный состав сигналов применительно к русской речи. В частности, ими исследовались статистические характеристики ЧОТ (F0T) и основные формантные частоты (F1–F4). Применительно к среднему мужскому голосу авторы выделили диапазоны формантных частот F1–F4, которые при их произнесении являются сонорными неносовыми звуками речи и приведены авторами в виде таблицы 1. В данной таблице частота основного тона F0T, представлены авторами для всего диапазона всех возможных мужских, женских и детских голосов.

Таблица 1 – Спектральный состав речевых сигналов для русского языка

F_{0T}	F_1	F_2	F_3	F_4
60–300 Гц	150–850 Гц	500–2500 Гц	1500–3500 Гц	2500–4500 Гц

Авторы рассматривают речевой сигнал (РС), как во времени, так и по частоте. Поэтому спектр сигнала ими получается на основе взаимного наложения друг на друга его гармонической (узкополосной) и формантной спектрально-временных структур (рисунок 5) [2]. Мощность сигнала, применительно для конкретной точки частотно-временного описания, представлена на рисунке 5. Градация по цвету определяется от чёрного (самого мощного) до белого (самого слабого). Кластеризация фонем представлена в работе [3]. Однако для данного исследования это потребует сбора большого количества характеристик и

параметров аудиосигналов междометий, применительно к русскому, узбекскому и французскому языкам в отдельности.

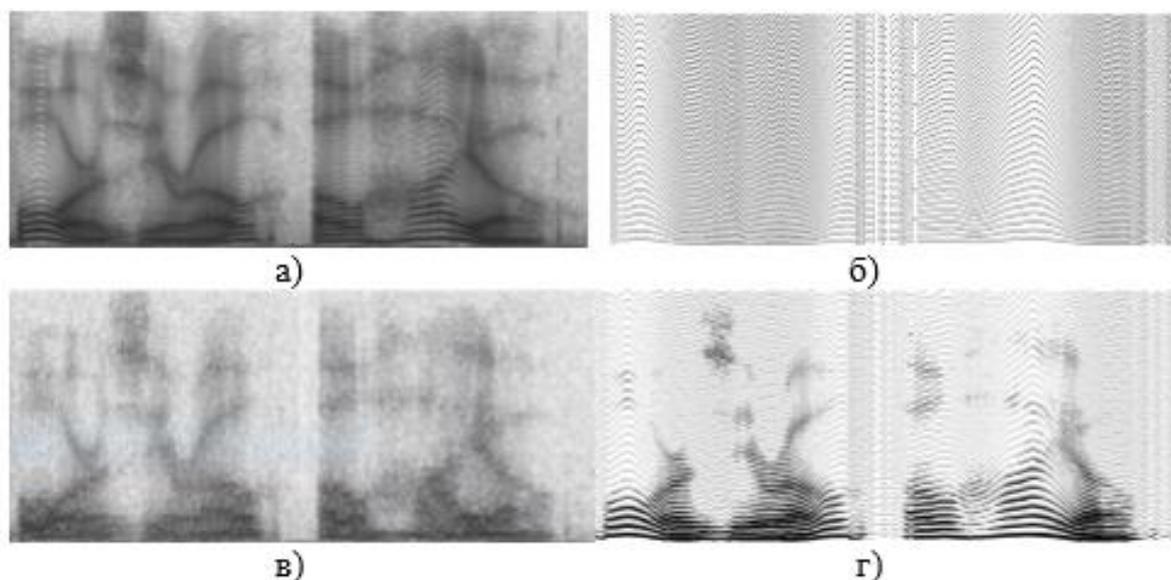


Рисунок 5 – Спектрограмма исходного РС (а),
гармоническая структура (б), формантная структура (в),
бинарная речевая подпись (г)

Таким образом, в ходе данных исследований были получены следующие результаты, которые представляют собой необходимость выполнения сравнительного анализа релевантных параметров аудиосигналов для междометий русского языка [1-13].

Список литературы

1. Жиляков, Е.Г. Феноменологическая математическая модель психоакустики слуха человека / Е.Г. Жиляков, С.Л. Бабаринов // Научные ведомости Белгородского государственного университета. Серия: Экономика. Информатика. – 2017. – № 16 (265). – С. 122-130.
2. Системное моделирование речеподобных сигналов и его применение в сфере безопасности, связи и управления / С.В. Дворянкин, С.В. Уленгов, Р.А. Устинов, Н.С. Дворянкин, А.О. Антипенко // Безопасность информационных технологий. – 2019. – Т. 26, № 4. – С. 101-119.
3. Савченко, В.В. Экспериментальное исследование фонетических свойств речевого сигнала на основе его теоретико-информационной модели / В.В. Савченко, Т.А. Соловьева // Научные ведомости Белгородского государственного университета. Серия: Экономика. Информатика. – 2018. – Т. 45, № 1. – С. 168-175.

4. Dobrushkin, G.O. Basic approaches to speech recognition (part II) / G.O. Dobrushkin, V.Ya. Danylov // Вісник Вінницького політехнічного інституту. – 2010. – № 2 (89). – С. 61-73.
5. Зацепин, П.М. Применение технологии NVIDIA CUDA для обучения и декодирования скрытых марковских моделей / П.М. Зацепин, Д.А. Гефке // Известия Томского политехнического университета. – 2014. – Т. 324, № 5. – С. 115-121.
6. Вейвлет-преобразование Добеши для низкочастотных сигналов, снятых с коры головного мозга человека / Д.В. Бибииков, Р.Б. Буров, В.В. Лавлинский, Ю.Г. Табаков // Моделирование систем и процессов. – 2013. – № 2. – С. 8-11
7. Лавлинский, В.В. Теоретические предпосылки решения проблем формирования моделей и методов взаимодействия информационных процессов / В.В. Лавлинский // Моделирование систем и процессов. – 2013. – № 2. – С. 30-36.
8. Лавлинский, В.В. Теоретические основы математического моделирования для описания целенаправленных систем / В.В. Лавлинский, Е.А. Рогозин, С.Н. Яньшин // Вестник Воронежского института МВД России. – 2017. – № 2. – С. 143-153.
9. Лавлинский, В.В. Формирование моделей и методов взаимодействия информационных процессов / В.В. Лавлинский, О.Г. Иванова // Приборы и системы. Управление, контроль, диагностика. – 2014. – № 5. – С. 39-50.
10. Лавлинский, В.В. Формирование моделей и методов взаимодействия информационных процессов / В.В. Лавлинский, О.Г. Иванова // Прикладная физика и математика. – 2015. – № 4. – С. 49-61.
11. Лавлинский, В.В. Метод обработки звуковых сигналов на основе wavelet-преобразований / В.В. Лавлинский, Д.М. Ковалевский // Моделирование систем и процессов. – 2016. – Т. 9, № 3. – С. 30-34.
12. Лавлинский, В.В. Критерий выявления закономерностей в информационных потоках речевых сигналов / В.В. Лавлинский, Т.В. Песецкая // Промышленные АСУ и контроллеры. – 2019. – № 3. – С. 36-42.
13. Лавлинский, В.В. Метод сравнения смешанных и несмешанных звуковых сигналов на основе wavelet-преобразований / В.В. Лавлинский, Д.М. Ковалевский, Д.С. Григорьев // Моделирование систем и процессов. – 2016. – Т. 9, № 4. – С. 40-43.