

## АКУСТИЧЕСКИЕ АСПЕКТЫ ИНТОНАЦИОННОГО ПОВЕДЕНИЯ В ДИКТОРСКОЙ РЕЧИ

**А. Б. АНАНЬЕВ**

*канд. техн. наук, НТУУ «КПИ»  
e-mail: anatolak@yandex.ru*

**М. Э. АВРАМЕНКО**

*студент кафедры акустики и акустоэлектроники НТУУ «КПИ»*

Работа посвящена объективному анализу акустических параметров интонации в разговорной речи. Приведены примеры детального графического описания интонационного речевого поведения. Путем корреляционного анализа исследована связь между поведением функций частоты основного тона и энергии в произносимой фразе. Разработанные средства позволяют анализировать интонационное поведение дикторов в связи с логическим построением как отдельной фразы, так и сообщения, содержащего текстовый блок фраз, выражающих завершённую мысль.

This study is devoted to objective analysis of acoustic parameters inherent to intonation in human speech. Within the study we introduced graphic descriptions of speech intonation patterns. Using correlation analysis we examined relation between pitch and sound energy in spoken phrases. A technique developed in this project allows us to relate intonation in the spoken phrase to the logical stresses within not only this phrase, but also within a contextually linked sequence of phrases.

### ВВЕДЕНИЕ

Изучение интонационного поведения в разговорной речи – важный и актуальный аспект изучения человеческого поведения в целом. Акустические методы и средства дают объективную базу для такого изучения. Особый практический интерес представляет изучение интонации в дикторской речи, поскольку его результаты могут быть применены при подготовке и оценке профессиональных качеств дикторов.

Диктор – это лицо, речь которого целенаправлена. Речевой ресурс диктора, а именно тембр голоса, артикуляция, интонационное поведение – все призвано служить поставленной цели.

Под интонацией (от лат. *intonō* – громко произношу) понимают изменение во времени ряда просодических<sup>1</sup> характеристик, основными из которых являются высота основного тона (в англоязычной литературе *pitch*), интенсивность и длительность. Функции интонации разнообразны, в контексте речевого поведения диктора важнейшей из них представляется расстановка логических (фразовых) ударений в произносимом потоке текстовых фраз.

В рамках работы в рассмотрение взята речь дикторов новостных программ. При изложении новостей само сообщение дает почву для расстановки фразовых ударений, что помогает изучать характерные особенности интонационного поведения каждого конкретного диктора.

---

<sup>1</sup> Просодия (от гр. *προσῳδία, prosōidia*) – ритмико-мелодическая схема организации звучащей связанной речи. Просодические характеристики – это невербальные характеристики речи, передающие значение, ударение и эмоцию высказывания.

## 1. ОБЗОР ЛИТЕРАТУРНЫХ ИСТОЧНИКОВ

Значению интонации в человеческой речи уделяется большое внимание в литературе, где широко распространены рассуждения о роли различных интонационных компонентов и взаимосвязях между ними. Однако какие-либо объективные количественные данные о просодических характеристиках речи встречаются крайне редко.

В работе [1] рассмотрены результаты экспериментальных исследований речевой интонации, проведенных на материале английского и русского языков. Суть экспериментов сводится к выяснению того, каким образом диктор, произнося фразу, добивается различного вида интонации в ней (простое заверщенное повествование, незавершенное повествование, повествование-ответ на вопрос, общий вопрос, восклицание). Один из выводов состоит в том, что наиболее точное восприятие интонационных различий достигается при одновременном действии всех трех компонентов интонации, однако параметры частоты основного тона при этом являются доминирующими.

Вопросам, связанным со слуховым восприятием интонации, уделено внимание в [4]. В частности, в данной работе отмечено, что интенсивность как просодическая характеристика речи играет вспомогательную роль в реализации ударений, а ее небольшие (порядка 3..5 дБ) самостоятельные изменения вовсе практически не влияют на восприятие ударения в слове.

Существенное внимание интонационным аспектам разговорной речи уделяется в специализированной литературе, посвященной задачам синтеза и клонирования речи [2], распознавания говорящего по голосу [3] и пр. Так, например, в работе [2] описана методика построения так называемых мелодических портретов, которые отражают поведение частоты основного тона на протяжении разговорной фразы. Такие портреты используются при синтезе речи для создания моделей интонационного поведения.

В работе [3] интонационное поведение обсуждается как один из признаков для идентификации говорящего по голосу. Автор указывает на актуальность исследования интонационных особенностей речи, а также на его сложность, поскольку крайне трудно указать, какой интонационный вариант является предпочтительным в конкретной ситуации.

Поскольку применение акустических методов и средств для объективного анализа интонационного речевого поведения в литературных источниках практически не освещено, проведение подобного анализа в настоящей статье представляется актуальным.

## 2. ФОРМАТ ПРЕДСТАВЛЕНИЯ ПРОСОДИЧЕСКОЙ ИНФОРМАЦИИ

Для анализа интонационного поведения в произносимой диктором фразе необходимо документировать поведение во времени частоты основного тона, «мгновенной» энергии сигнала и длительности гласных звуков на протяжении данной фразы. На рис. 1 показан разработанный нами формат представления просодических данных, который позволяет с высокой степенью подробности наблюдать поведение трех названных компонентов в интонационной картине произнесенной фразы<sup>2</sup>. В качестве

<sup>2</sup> Авторы выражают благодарность студентке НТУУ «КПИ» А. С. Солохе, активно участвовавшей в разработке этого формата.

примера представлено применение данного формата к речевой фразе, произнесенной диктором в телевизионном выпуске новостей.

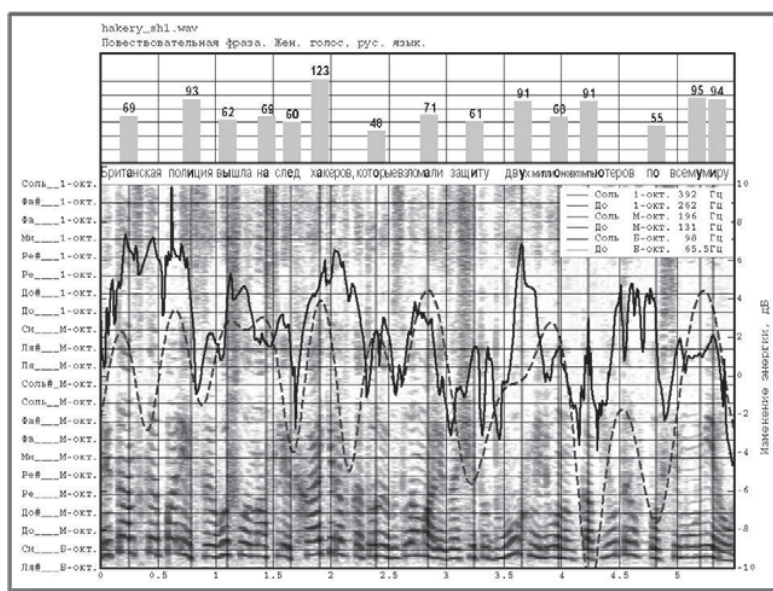


Рис. 1

На рисунке сплошная кривая отражает изменения питча (pitch), пунктирная – «мгновенной» энергии сигнала. Спектрограмма сигнала служит для наиболее точной привязки произносимого текста ко времени, поскольку позволяет судить о том, какой звук произносится в данный момент времени – гласный или согласный. Сам текст расположен выше, ударные гласные в нем выделены шрифтом. Выше текста указаны длительности ударных гласных звуков в миллисекундах. Моменты времени произношения ударных гласных выделены в нижней части графика вертикальными линиями. Выше поля рисунка приведена информация об обработанном файле: имя файла и краткое описание речевого материала.

Изменения питча показаны в нотной (музыкальной) шкале. Изменения энергии отображены в шкале децибел относительно среднего уровня энергии во фразе.

На спектрограмме отображены частоты от 0 до 8000 Гц. Этот диапазон позволяет отличить визуально произносимые гласные звуки от согласных.

Следует отметить тот факт, что длительность гласного звука является весьма неоднозначной характеристикой. Границы звука «размыты», определить их местонахождение на слух крайне затруднительно.

Поскольку, таким образом, исследование длительностей гласных как интонационного параметра представляет собой трудоемкую задачу, а их роль в интонационном оформлении речи не является ключевой, в дальнейшем мы рассматриваем поведение во времени только питча и энергии.

### 3. ВЗАИМНОЕ СООТВЕТСТВИЕ ЧАСТОТЫ ОСНОВНОГО ТОНА (ПИТЧА) И ЭНЕРГИИ

Сопоставление ряда графических результатов, подобных приведенному на рис. 1, со слуховым впечатлением от произносимых фраз подтверждает тот факт, что питч (по

сути, «мелодия», с которой диктор произносит каждое конкретное предложение) однозначно соответствует поведению интонации в произносимой фразе. Энергетическая же функция преимущественно отражает речевую артикуляцию.

Однако представляет интерес вопрос, в какой мере изменение энергии в речевом сигнале сопровождается изменением пика. В связи с этим было проведено исследование корреляционных связей между поведением функций частоты основного тона и «мгновенной» энергии сигнала в различных фразах из дикторской речи.

На рис. 2 представлены примеры графиков полученных взаимных корреляционных функций  $R(\tau)$ . Как видно из рисунка, такие функции, построенные для различных дикторов и образцов речи, либо не имеют явно выраженного максимума, либо имеют максимум при  $\tau = 0$ . Значения такого максимума достаточно близки между собой, и в большинстве случаев составляют величину порядка 0.3..0.4.

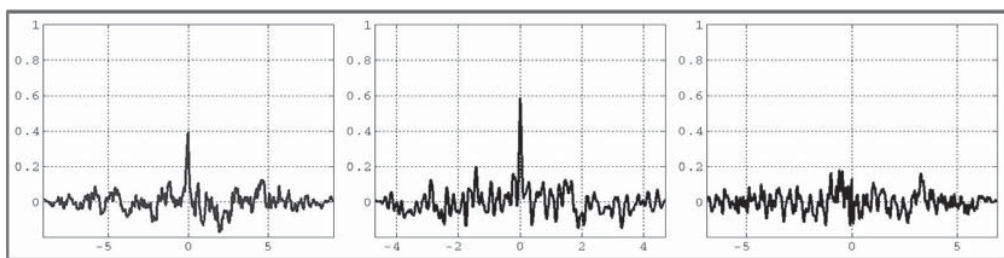


Рис. 2

Было отмечено, что фразы, для которых коэффициент корреляции частоты основного тона и энергии невелик (порядка 0.2..0.3), субъективно воспринимаются как эмоционально нейтральные, логические акценты в них, безусловно, присутствуют, однако выражены не ярко. В то же время образцы, для которых корреляционный максимум достигает значения порядка 0.6..0.7, воспринимаются как эмоционально насыщенные, вплоть до призыва слушателя к активным действиям.

Представляет интерес статистика распределения максимумов корреляционной функции пика и энергии в речевых фразах, полученных от различных информационных источников. На рис. 3 показана гистограмма распределения указанных максимумов, полученная для 255 образцов. Гистограмма построена для 15 каналов, пунктирными линиями указаны границы доверительных интервалов для доверительных вероятностей 0.6, 0.75 и 0.9. Как видно из рисунка, истинная кривая распределения с высокой вероятностью должна иметь сложный вид.

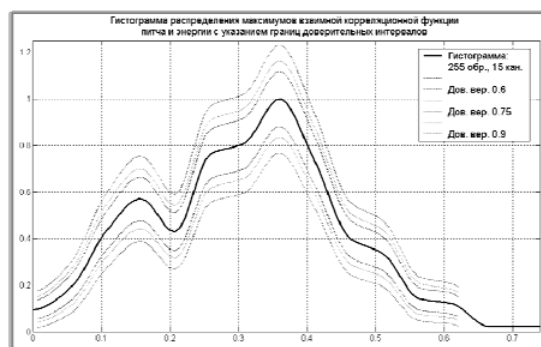


Рис. 3

Поскольку было замечено, что при одном и том же дикторе значение корреляционного максимума для разных фраз может быть различным, величину этого максимума можно рассматривать как меру эмоциональной выразительности речи.

Следующий эксперимент призван еще раз подтвердить решающую роль пика (по сравнению с энергией) в построении интонационного оформления произносимой фразы. В этом эксперименте в речевой фразе, величина максимума корреляционной функции для которой отражает высокую степень «сопровождения» основного тона энергией (и составляет значение 0.58), программными средствами были «подавлены» изменения энергии, что привело практически к полной декорреляции этих двух параметров (после обработки сигнала коэффициент корреляции составляет 0.07). После такой обработки субъективное восприятие интонационного оформления фразы осталось практически неизменным, при этом, однако, появилось ощущение некачественной записи звука.

Таким образом, для описания интонационного поведения в дикторской речи в дальнейшем можно ограничиться только одной кривой, отражающей поведение высоты основного тона.

#### 4. ИНТОНАЦИОННОЕ ОТРАЖЕНИЕ ИНФОРМАЦИОННОЙ ЛОГИКИ В ДИКТОРСКОЙ РЕЧИ

При исследовании интонации в речи диктора ключевым, на наш взгляд, является так называемый «крупномасштабный» анализ его интонационного поведения – исследование интонации в более продолжительные, чем одна фраза, промежутки времени. Такой анализ позволяет объективно выявлять, насколько речевое интонирование диктора соответствует информационной нагрузке, которую несет в себе сообщение.

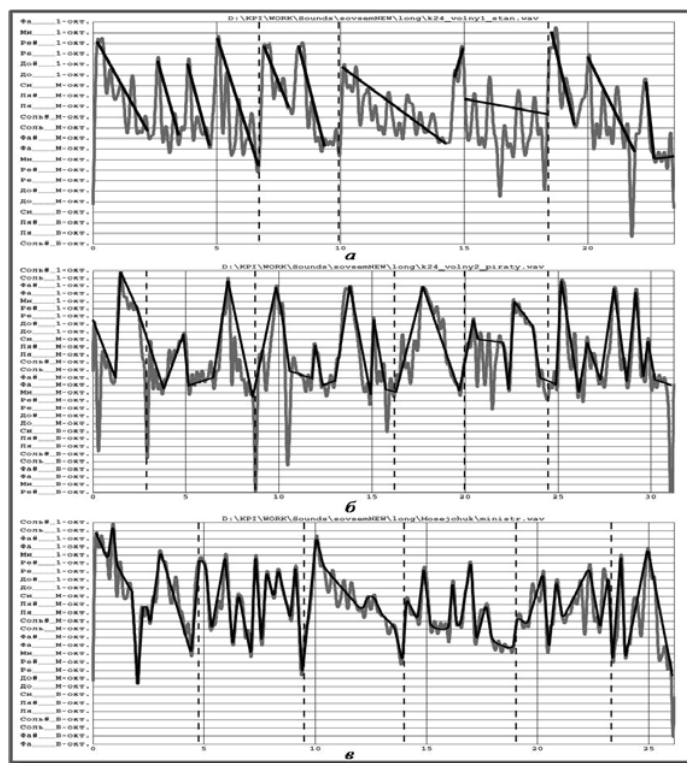


Рис. 4



На рис. 4 показаны примеры графических результатов подобного анализа, примененного к трем образцам из сообщений дикторов телевизионных новостных программ. Кривые отражают изменение питча во времени (в нотной шкале), а более темные линии, проведенные поверх кривых, иллюстрируют общую тенденцию в изменении «мелодии» при произнесении текста. Вертикальные пунктирные прямые разделяют все сообщение на логические составные части (отдельные завершённые фразы).

Образцы *a* и *б* воспринимаются на слух, как монотонные, «шаблонные», интонация в них связана, скорее, не с расстановкой логических ударений, а с собственным характерным интонационным поведением диктора. На рисунках это отражается общей тенденцией повторять повышения и понижения питча в одних и тех же пределах.

В то же время образец речи *в* воспринимается как эмоционально насыщенный, интонационное поведение в нём сопровождает логические акценты. Из рисунка видно, что кривая питча более разнообразна, чем у образцов *a* и *б*.

Таким образом, документирование с применением акустических средств интонационного поведения в разговорной речи позволяет выявлять характерные особенности, присущие конкретным дикторам. Применение такого документирования может быть использовано в процессе обучения дикторской речи и при контроле её качества. Отметим при этом, что критерии качества интонационного речевого поведения могут быть разработаны исключительно в сотрудничестве со специалистами в области постановки речи.

## ВЫВОДЫ

В данной работе представлены средства, позволяющие выполнять объективный анализ акустических интонационных параметров в связи с логическим построением как отдельной фразы, так и сообщения, содержащего набор фраз, выражающих завершённую мысль. Приведены примеры детального графического описания интонационного речевого поведения, которые вскрыли особенности, присущие конкретным дикторам телевидения в процессе их работы. Также исследована корреляционная связь между поведением функций частоты основного тона и энергии в произносимой фразе.

Полученные результаты могут быть предложены журналистским организациям для дальнейшего совместного всестороннего исследования интонационного речевого поведения в различных профессиональных и бытовых ситуациях, и разработки рекомендаций по применению объективного анализа интонационной картины сообщения при обучении дикторской речи.

## ЛИТЕРАТУРА

1. Кантер Л. Р. Системный анализ речевой интонации. – М.: Высш. шк., 1988. – 128 с.
2. Лобанов Б. М. Компьютерный синтез и клонирование речи / Б. М. Лобанов, Л. И. Цирульник. – Минск: Белорус. наука, 2008. – 343 с.
3. Рамишвили Г. С. Речевой сигнал и индивидуальность голоса. – Тбилиси: Мецниереба, 1976. – 183 с.
4. Физиология речи. Восприятие речи человеком / Л. А. Чистович и др. – Л.: Наука, 1976. – 388 с.