

## ОБ ИНТОНАЦИОННОЙ ИНДИВИДУАЛЬНОСТИ УСТНОГО ЧТЕНИЯ СЛОЖНЫХ ПРЕДЛОЖЕНИЙ

*Б. М. Лобанов*

Данное исследование вызвано желанием достичь, в конечном счёте, максимально возможной естественности и персонализации интонационного оформления синтезированной речи при «чтении» художественных текстов, включающих, как правило, большое количество распространённых и сложных предложений. В настоящее время в Интернете в свободном доступе находятся ряд «движков» синтеза русской речи по тексту, таких как Nikolay, Olga, Alena и др. Они основаны на «Corpus Based» подходе и обеспечивают высокое качество звучания речи на сегментном уровне, однако, интонационное оформление речи, особенно при прослушивании созданных на их основе аудиокниг, в значительной степени неудовлетворительно.

В данной работе под общим понятием «интонация» исследуется лишь её мелодическая составляющая, не включая другие известные просодические характеристики. Кроме того, мелодический контур исследуется в жёстких, одинаковых условиях относительно таких важных характеристик фразы, как интонационный тип, положение границ составляющих её синтагм и позиции выделенных слов в каждой синтагме. Тем самым рассматривается вариативность, обуславливаемая лишь персональными дикторскими вариациями мелодического контура. Для реализации этих условий предложена следующая методика эксперимента.

В качестве «эталонной» выбрана звукозапись (аудиокнига) чтения профессиональным диктором отрывка повести А. П. Чехова «Драма на охоте». Каждое предложение в выбранном отрывке размечалось опытным аудитором на синтагмы, а в каждой синтагме знаком (+) отмечались ударные гласные просодически выделенных слов. Ниже приведен пример предложения, состоящего из 9-ти синтагм.

- i. *В оди+н из апре+льских полу+дней /*
- ii. *ты+сяча восемьсо+т восьмидесятого го+да //*
- iii. *в мо+й кабине+т /*
- iv. *воше+л сторож Андре+й /*
- v. *и таи+нственно доложи+л мне, /*

- vi. *что в реда+кцию /*
- vii. *яви+лся какой-то господи+н //*
- viii. *и убеди+тельно про+сит /*
- ix. *свида+ния с реда+ктором. //*

Здесь знаком (/) отмечены короткие межсинтагменные паузы (вплоть до нулевой длительности), а знаком (//) – относительно долгие паузы.

Затем нескольким контрольным дикторам (непрофессионалам) предлагалось прочесть каждое предложение с соблюдением эталонного синтагматического членения и просодической выделенности отмеченных слов. При этом дикторам ни разу не предоставлялась возможность прослушивания эталонных записей профессионального диктора. Кроме того, запись речи контрольных дикторов осуществлялась до 10-ти раз в разное время с тем, чтобы исключить возможность самоимитации.

Полученные записи анализировались с использованием модели представления интонации синтагм последовательностью мелодических Портретов Акцентных Единиц (ПАЕ). ПАЕ-модель была предложена автором более 20 лет назад и успешно используется до сих пор в различных моделях синтеза речи по тексту [1 – 3]. В соответствии с ПАЕ-моделью, минимальной интонационной единицей является Акцентная Единица (АЕ), состоящая из одного или более слов, и имеющая в своём составе только одну, главноударную гласную. Каждая АЕ, в свою очередь, состоит из ядра (главноударная гласная), предъядра (все фонемы, предшествующие главноударной гласной) и заядра (все фонемы за главноударной гласной). Наиболее важное предположение ПАЕ-модели состоит в том, что топологические свойства просодических параметров для определенного типа интонации синтагмы не изменяются (или изменяются незначительно) с изменениями высоты голоса, фонетического контекста и числа слогов в пред-и заядре АЕ.

С учётом разметки синтагм на АЕ показанный выше пример перепишется в виде:

В оди+н  из апре+льских  полу+дней  /	(3 АЕ)
ты+сяча  восемьсо+т  восьмидесятого го+да  //	(3 АЕ)
в мо+й  кабине+т  /	(2 АЕ)
воше+л  сторож Андре+й  /	(2 АЕ)
и таи+нственно  доложи+л мне , /	(2 АЕ)
что в реда+кцию  /	(1 АЕ)
яви+лся  какой-то господи+н  //	(2 АЕ)
и убеди+тельно  про+сит  /	(2 АЕ)
свида+ния  с реда+ктором . //	(2 АЕ)

На рисунках 1, 2 представлены примеры различных мелодических ПАЕ, полученных с использованием разработанной системы «Инто-Клонатор» [3], где по вертикальной оси отложены нормированные значения ЧОТ (в диапазоне от 0 до 1) в соответствии с формулой  $F_{0\text{ норм}} = (F_0 - F_{\text{min}}) / (F_{0\text{ max}} - F_{\text{min}})$ , а по горизонтальной оси – нормированные значения времени в терминах позиций предъядра, ядра и заядра. На рисунке 1 представлены ПАЕ 3-х акцентной синтагмы «*В один из апрельских полудней*» для 3-х дикторов: профессионального (1) и 2-х непрофессиональных – мужчину (2) и женщину (3).

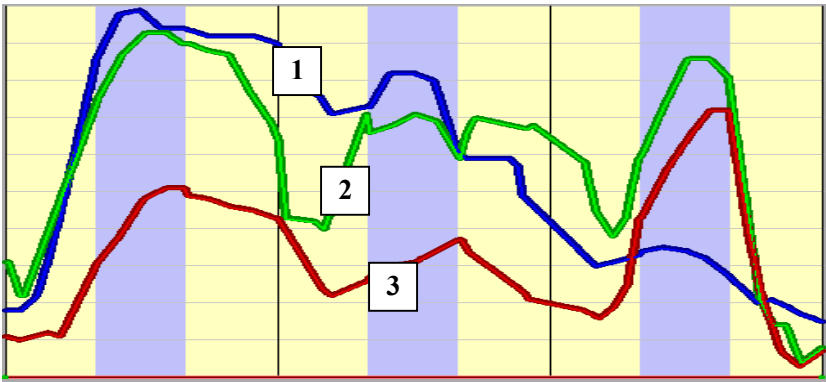


Рис. 1. ПАЕ синтагмы «*В один из апрельских полудней...*» для 3-х дикторов

Как видно из рисунка мелодические ПАЕ имеют существенные персональные различия.

На рисунке 2 представлены 6 реализаций ПАЕ одной и той же одноакцентной синтагмы «*что в редакцию*», принесённой в разное время диктором женщиной, а на рисунке (В) – диктором мужчиной. Жирные кривые отображают ПАЕ, усреднённые по 6-ти реализациям, а нижняя кривая – отклонение от средних ПАЕ.

Рисунки 2А и 2Б показывают, что индивидуальные отличия ПАЕ остаются достаточно стабильными у обоих дикторов.

Были проведены также исследования персональных особенностей реализации ПАЕ для четырёх различных интонационных типов: незавершённость, завершённость, восклицание и общий вопрос, в соответствии со следующей методикой. Два профессиональных диктора радио (Олег и Светлана) и три непрофессиональных (Борис, Елена, Лилия) зачитали один и тот же отрывок художественного текста. Затем в фонограммах записей каждого диктора были выделены одинаковые

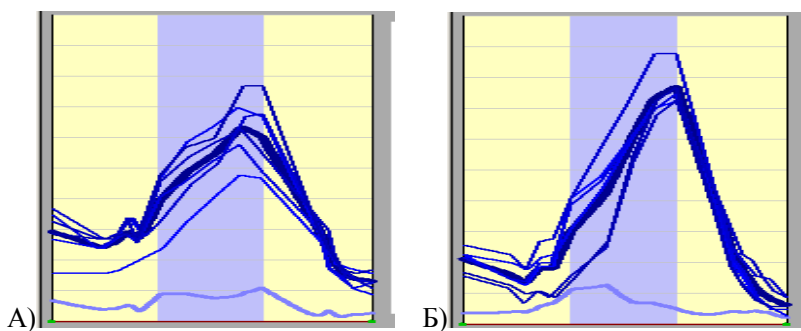


Рис. 2. ПАЕ синтагмы «что в редакцию...»  
для 6 произнесений 2 дикторами

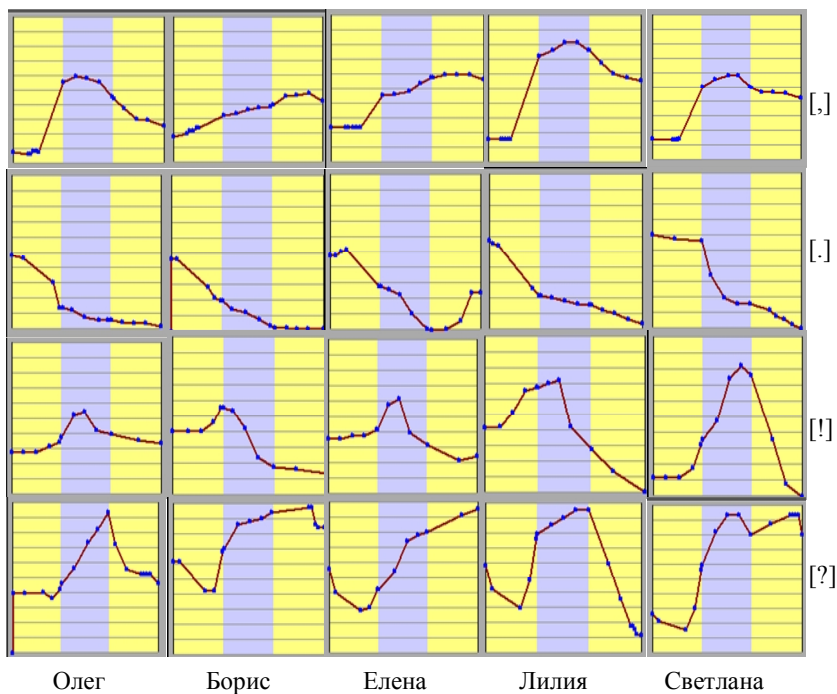


Рис. 3. ПАЕ для четырёх интонационных типов 5-ти дикторов

участки речи, на которых ими были реализованы указанные четыре интонационных типа, и на основании анализа контуров  $F_0$  построены ПАЕ в соответствии с разработанной методикой. На рис. 3 представлены полученные мелодические портреты конечных АЕ персональной реализации четырёх интонационных типов у 5 дикторов.

Как видно из рис. 3, полученные ПАЕ имеют достаточно ярко выраженные персональные особенности. Причём в наибольшей степени индивидуальные дикторские различия в ПАЕ проявляются на предъядре и заядре, в то время как на ядерных участках они менее значительны. В целом же, однако, сохраняется рисунок портретов, характерный для каждого исследуемого интонационного типа.

### Литература

Boris. Lobanov. The Phonemophon Text-to-Speech System // Proc. of the XIth International Congress of Phonetic Sciences – ICPhS'87, Tallin, 1987. С. 120–124.

Boris. Lobanov, Helena Karnevsaya, Lilia Tsirulnik. Language- and Speaker Specific Implementation of Intonation Contours in Multilingual TTS Synthesis // Proc. of the 3<sup>rd</sup> International conference “Speech Prosody”, Dresden, Germany, May 2–5, 2006. – P. 553–556.

Б. М. Лобанов, Л. І. Цірульнік. Компьютерный синтез и клонирование речи // Минск: Белорусская Наука, 2008. – 342 с. – Авт. также: Цирульнич Л. И.