

Учреждение образования
«Белорусский государственный университет
информатики и радиоэлектроники»

Государственное учреждение
«Администрация Парка высоких технологий»

Открытые семантические технологии проектирования интеллектуальных систем

Open Semantic Technologies for Intelligent Systems

МАТЕРИАЛЫ
МЕЖДУНАРОДНОЙ
НАУЧНО-ТЕХНИЧЕСКОЙ КОНФЕРЕНЦИИ

Выпуск 1

(Минск, 16-18 февраля 2017)

Основан в 2011 году

Минск
БГУИР 2017

УДК 004.822+004.89-027.31
ББК 32.813-02+73
О-83

ISSN 2415-7740

Редакционная коллегия:

*В. В. Голенков (отв. ред.), Л. С. Глоба, В. А. Головко, Н. А. Гулякина, О. П. Кузнецов, Б. М. Лобанов,
Г. С. Осипов, С. В. Смирнов, П. И. Соснин, Д. Ш. Сулейманов, В. Ф. Хорошевский*

Организаторы конференции:

Министерство образования Республики Беларусь

Министерство связи и информатизации

Учреждение образования «Белорусский
государственный университет информатики и
радиоэлектроники»

Российская ассоциация искусственного интеллекта

Государственное учреждение «Администрация Парка
высоких технологий» (Республика Беларусь)

Объединённый институт проблем информатики
Национальной академии наук Беларусь

Брестский государственный технологический
университет

ОАО «Савушкин продукт»

ФГБНУ «Институт управления образованием
Российской академии образования»

Кафедра молодежной политики РИВШ

ООО «Интеллектуальные процессы»

ЗАО «Qulix Systems»

ЗАО «Итранзишен»

ЧТУП «АйДжи Софт Бел»

ОДО «ВирусБлокАда»

ООО «Октонион технолоджи»

ИООО «Софтек Девелопмент»

ООО «Альторос Девелопмент»

ООО «РСЛ»

ООО «Аквео»

ООО «Датарокетс»

ООО «Соларвиндз МСП Технолоджи»

Техническая и информационная поддержка:

Научный журнал «Информатика»

Научный журнал «Онтология проектирования»

Научно-технический журнал «Автоматизация процессов
управления»

Научно-практический журнал для специалистов
«Электроника ИНФО»

Международный журнал «Программные продукты и
системы»

Научно-производственный журнал «Веснік сувязі»

*Издание осуществлено по заказу государственного учреждения
«Администрация Парка высоких технологий»*

О-83 **Открытые семантические технологии проектирования интеллектуальных
систем = Open Semantic Technologies for Intelligent Systems** : материалы
междунар. науч.-техн. конф./ редкол. : В. В. Голенков (отв. ред.) [и др.], ISSN
2415-7740; Вып.1 (Минск, 16-18 февраля 2017г.). – Минск : БГУИР, 2017. --
466 с.

ISBN 978-985-543-034-7.

Сборник включает прошедшие рецензирование статьи VII международной научно-технической
конференции «Открытые семантические технологии проектирования интеллектуальных систем».

Сборник предназначен для преподавателей высших учебных заведений, научных сотрудников,
студентов, аспирантов, магистрантов, а также для специалистов предприятий в сфере
проектирования интеллектуальных систем.

Материалы сборника одобрены Программным комитетом OSTIS и печатаются в виде,
представленном авторами.

**УДК 004.822+004.89-027.31
ББК 32.813-02+73**

ISBN 978-985-543-034-7

© УО «Белорусский государственный
университет информатики
и радиоэлектроники», 2017

О СПЕЦИАЛИЗАЦИИ МОДЕЛЬНО-УПРАВЛЯЕМОГО ПОДХОДА ДЛЯ СОЗДАНИЯ ИНТЕЛЛЕКТУАЛЬНЫХ СИСТЕМ ПОДДЕРЖКИ ПРИНЯТИЯ РЕШЕНИЙ ПРЕЦЕДЕНТНОГО ТИПА	151
Дородных Н.О., Юрин А.Ю.	
СЕМАНТИЧЕСКОЕ ПРОСТРАНСТВО В ФАКТОРИАЛЕ ТЕОРИИ АВТОМАТИЧЕСКОГО ПОРОЖДЕНИЯ АРХИТЕКТУРЫ ЗНАНИЙ (ТАПАЗ-2)	155
Гордей А. Н.	
СЕМАНТИЧЕСКОЕ КОДИРОВАНИЕ ДЛЯ СЕМАНТИЧЕСКОГО ВЫВОДА ЗНАНИЙ	158
Бойко И.М.	
ПРОТОТИП КОМПЬЮТЕРНОЙ СИСТЕМЫ ОБУЧЕНИЯ РЕЧЕВОЙ ИНТОНАЦИИ	163
Лобанов Б.М., Житко В.А.	
МУЛЬТИГОЛОСОВОЙ СИНТЕЗ РЕЧИ ПО ТЕКСТУ ДЛЯ ПОСТРОЕНИЯ ЕСТЕСТВЕННО-ЯЗЫКОВЫХ ИНТЕРФЕЙСОВ ИНТЕЛЛЕКТУАЛЬНЫХ СИСТЕМ	167
Захарьев В.А., Петровский А.А.	
ЛИНГВИСТИЧЕСКИЕ И АКУСТИЧЕСКИЕ РЕСУРСЫ КОМПЬЮТЕРНОЙ СИСТЕМЫ ОБУЧЕНИЯ ДИАЛОГОВОЙ ИНТОНАЦИИ АНГЛИЙСКОЙ РЕЧИ	171
Здоронок Ю.А.	
ГОЛОСОВАЯ ИДЕНТИФИКАЦИЯ ПОЛЬЗОВАТЕЛЯ В СИСТЕМАХ КОНТРОЛЯ ДОСТУПА	175
Меньшаков П.А., Мурашко И.А.	
ПОПОЛНЕНИЕ ЭЛЕКТРОННОЙ ЛЕКСИЧЕСКОЙ БАЗЫ ДАННЫХ ЧЕРЕЗ КОМПЬЮТЕРНУЮ СИСТЕМУ «ГЕНЕРАТОР ПАРАДИГМЫ СЛОВА»	179
Зеновко Е.С.	
РЕАЛИЗАЦИЯ ИНСТРУМЕНТОВ ОБУЧЕНИЯ С ПОДКРЕПЛЕНИЕМ ДЛЯ ИНТЕЛЛЕКТУАЛЬНЫХ СИСТЕМ ПОДДЕРЖКИ ПРИНЯТИЯ РЕШЕНИЙ РЕАЛЬНОГО ВРЕМЕНИ	183
Еремеев А.П., Кожухов А.А.	
О ФОКУСИРОВКЕ ВНИМАНИЯ НА РЕЛЕВАНТНОЙ ИНФОРМАЦИИ ПРИ РАССУЖДЕНИЯХ АГЕНТОВ В ИНТЕЛЛЕКТУАЛЬНЫХ СИСТЕМАХ ЖЁСТКОГО РЕАЛЬНОГО ВРЕМЕНИ	187
Фоминых И.Б., Романчук С.В.	
АЛГОРИТМ ИЗВЛЕЧЕНИЯ ТЕМПОРАЛЬНЫХ ПРЕЦЕДЕНТОВ ДЛЯ МЕХАНИЗМА РАССУЖДЕНИЙ НА ОСНОВЕ ТЕМПОРАЛЬНЫХ ПРЕЦЕДЕНТОВ	191
Куриленко И.Е., Гулякина Н.А.	
ОБНАРУЖЕНИЕ АНОМАЛИЙ В ДАННЫХ НА ОСНОВЕ АНАЛИЗА СЕТЕВОГО ТРАФИКА	195
Антипов С.Г., Вагин В.Н., Фомина М.В.	
ИНТЕЛЛЕКТУАЛЬНАЯ ОБУЧАЮЩЕ-ТЕСТИРУЮЩАЯ ПРОГНОЗИРУЮЩАЯ СИСТЕМА С КОГНИТИВНОЙ КОМПОНЕНТОЙ	199
Янковская А.Е., Шелупанов А.А., Шурыгин Ю.А., Дементьев Ю.Н., Ямшанов А.В., Ляпунов Д.Ю.	

A Prototype of the Computer System for Speech Intonation Training

Lobanov B.M., Zhitko V.A.

United Institute of Informatics Problems NAS Belarus
Minsk, Belarus

Email: lobanov@newman.bas-net.by, zhitko.vladimir@gmail.com

Abstract—In this paper we describe the importance of intonation in human comprehension of speech. We lay out fundamental principles of intonational theory of speech that is based on the concept of universal melodic portraits. In this paper we describe the main algorithms of analysis and interpretation of intonation in an utterance that underlie the developed computer-based system for teaching of speech intonation. We further show the system's output and discuss its usefulness in teaching foreign language intonation.

Keywords—Intonation of speech, speech analysis and synthesis, melodic portrait, intonation analysis and interpretation, computer system for teaching, intonation training.

I. INTRODUCTION

Intonation plays significant role during speech comprehension. Without intonation, it's impossible to understand the expressions and thoughts that go with words.

Speech intonation shows a communicative intention of an utterance, its logical meaning, a prominence of the most significant theme in relation to general themes (actual division of a sentence), a distinction between semantically associated segments of speech, and an integration of speech elements within these segments. We can therefore think of intonation as referring to the way we use the pitch of our voice to express particular meanings and attitudes. The exact relation between intonational patterns and informational structure (as part of semantics) is still to be investigated.

A common linguistic idea is that foreign accent appears more evident in intonation, and therefore, prosodic aspects of speech should be explicitly taught to students who wish to communicate intelligibly in a foreign language. Foreign accent arises due to contact between two different language systems, namely, at bilingualism, because of language interference. Intonation is the most important aspect of speech that provides both linguistic and sociocultural information. Considering that functions of intonation in speech are various, and that deviations in this area can lead to significant semantic differences, incorrect intonation can make a wrong impression during the speech of a non-native language speaker [1].

In this view, is it very important to emphasize intonational aspects of speech while teaching foreign language. In [2] an automatic intonation assessment system for computer aided language learning is described. The similarity between reference and test intonation templates is evaluated on a frame-by-frame basis by using the DTW alignment by estimation of the correlation between both F0 curves. Previously [3 –

5], we've published several scientific results in the area of speech intonation analysis and synthesis which can serve as a foundation for the creation of innovative intelligent systems of automatic intonation assessment for computer aided language learning.

II. GENERAL INFORMATION ABOUT SPEECH INTONATION

In phonetics and physiology of speech, a phrase is considered to be a comparatively independent unit in speech intonation [6]. Phrase independence manifests itself in articulatory integrity, semantic and syntactical association of linguistic units, and in presence of objective traits that allow singling it out from a speech stream. Any punctuation mark can set a boundary of intonation phrase in a written text. However, quite often the number of phrases can be greater than the number of punctuation marks. Generally, a phrase is a combination of one to five words, with a three-word phrase being the most common. A particular place of a phrase boundary is determined by the optimal satisfaction of the semantico-syntactic, phonetic and physiological requirements. The first of the named requirements prescribes a union inside a phrase between semantically connected words that cannot be split into two phrases. The second requirement denotes the tendency of language phonetic systems for definite rhythmical construction, for example, a group of two to three words combined in one phrase. Finally, the third requirement prescribes the formation of a phrase with the number of words that could be physically pronounced during the time required for one act of exhaling.

After definition of a phrase, the next step in intonational speech analysis is determination of a phrase intonation type. The main intonation phrase types include: completed and uncompleted phrases, special and common questions, exclamation phrases, and some others types. The number of intonation subtypes of the main phrase intonation types could reach several dozens. The task of definition of an intonation subtype is achieved by the analysis of two factors: the position of a phrase in a text and by its semantic value. The first factor, a position in a text, is determined by the analysis of a phrase position in relation to the nearest punctuation marks. For example, its position at the beginning or the end of a text, its position before or after punctuation marks, and the type of punctuation mark in a sentence. The second factor, a semantic value of a phrase, is determined based on the meaning of a phrase and on its logically emphasized intonational center. In particular, it needs to be determined whether a phrase expresses intentions, explanations, follow up questions etc. The final

decision on an intonation type and subtype of a phrase is made after considering both factors.

Speech intonation is physically realized by the set of acoustic means, named the prosodical parameters:

- melodics - the movement of the frequency of the main tone (F_0);
- energetics - the current change of the force (amplitude) of the sound (A);
- rhythmics - the current change of the duration of the sounds and pauses (T).

Since melodics of speech is the most informative among these parameters, the main attention in this article is given to the melodics.

III. BLOCK DIAGRAM OF THE SUGGESTED COMPUTER SYSTEM FOR SPEECH INTONATION TRAINING

Figure 1 contains a block diagram illustrating sequence of algorithms for analysis of speech intonation within the developed computer training system. The main goal of the system is to provide a student with a compact and easily interpretable image for the results of analysis of melodic and energy contours of phrases with different intonation. The system would also provide a visual, auditory and numerical evaluation of the quality of learning of a foreign speech intonation by a student.

Block 1 contains the database of phrase samples with different intonation patterns which is compiled from multimedia textbooks (see, for example, [7] for Russian language, or [8] - for English). Every sample phrase has preliminary placed prosodic marks that include phrase boundaries and placement of its nucleus.

Based on a given goal of intonation learning, a student chooses the needed sample phrase, hears it and pronounces it. The pronounced phrase is recorded on the buffer (block 2).

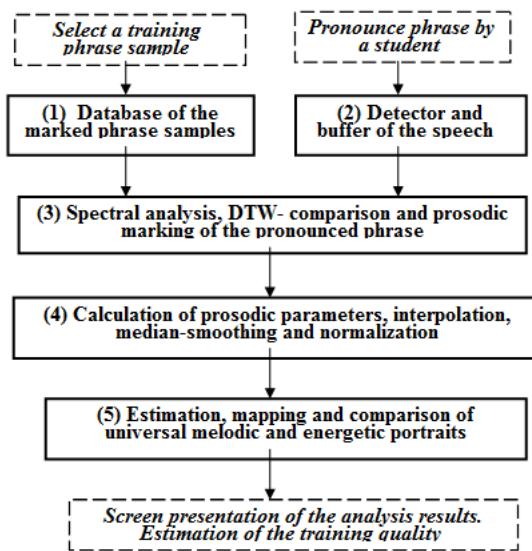


Figure 1. The block diagram of the computer training system of foreign speech intonation

In block 3, the signals from both sample and pronounced phrases are analyzed by calculating Mel-frequency cepstral coefficients (MFCCs) and then compared using the modified method of DTW – continuous dynamic time warping (CDTW) [9]. This is accompanied by determination the beginning and the end of the pronounced phrase, a transfer of prosodic marks (phrase boundaries and placement of its nucleus) and labeling of a pronounced phrase.

In block 4, prosodic phrase parameters of both sample and pronounced speech signals, such as F_0 – frequency of the basic tone (pitch), and A_0 – energy of the signal, are calculated. These parameters are further interpolated on the non-vocal areas of speech signal, and then they are median-smoothed and normalized on their minimum and maximum of phrase values.

In order to calculate the proper value of training quality, in block 5, an estimation and comparison of universal melodic and energetic portraits are produced.

Figure 4 presents illustration of system's output for the interrogative phrase: "Did Sasha eat the porridge?" The image shows successive processing $F_0(t)$ and $A_0(t)$ and a comparison of the sample phrase and the student-spoken phrase speech signals.

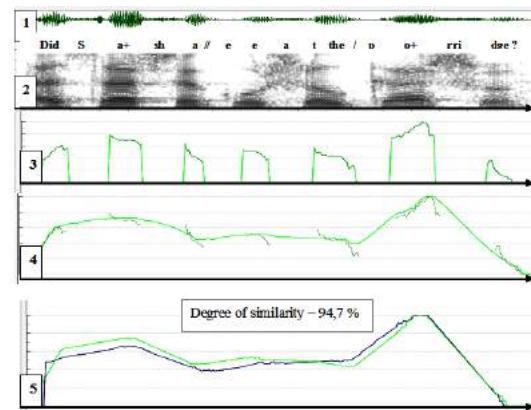


Figure 2. The illustration of speech signals processing: 1- oscillogram, 2 – spectrum, 3 – $F_0(t)$ (original), 4 – $F_0(t)$ (after interpolation and median smoothing), 5 – comparison of two melodic curves $F_0(t)$ - sample and spoken phrases

IV. SOFTWARE REALIZATION OF THE COMPUTER SYSTEM PROTOTYPE

Software realization of the prototype is written on C++ programming code by using Qt framework. It can be compiled under Windows platform (from xp to 10 versions), also under Linux platform.

Application used following libraries:

- OpenAL - used for multimedia interaction (audio input/output). OpenAL is a cross-platform audio application programming interface (API). It is designed for efficient rendering of multichannel three-dimensional positional audio;
- MathGL - used for draw graphics. MathGL is a library for making high-quality scientific graphics under Linux and Windows;

- GNU Scientific Library (GSL) - used for data calculation and processing. GSL is a numerical library for C and C++ programmers. It is free software under the GNU General Public License;
- SPTK - used for audio data analyzing. SPTK is a suite of speech signal processing tools for UNIX environments, e.g., LPC analysis, PARCOR analysis, LSP analysis, PARCOR synthesis filter, LSP synthesis filter, vector quantization techniques, and other extended versions of them.

Application core divided into several modules that implement standalone functions. Such modules can implement audio signal recording, voice detection, DP processing and so on. As this modules are independent from each other, we can easily build different applications by replacing this modules for other one or integrate them in external systems.

For build main user interface used built-in web engine. User interface built on html5, css3 and js (ReactJs js framework). "Developer mode" user interface build on standard Qt forms.

Main user interface is independent from application core and can be modified or even replaced by other one. Use html/css/js standard allow easy change application front-end for different purposes. For interaction with application core exists a number of special links formats that processed by application core. Such links can open different applications dialogs (like settings, developer mode and so on), process input audio signals and play audio files.

So we can easy build different training systems by replacing front-end and training data files.

The starting page of the User Interface is shown in figure 3. From this page, also possible to pass in "Developer mode".



Figure 3. Starting page of the User Interface

After clicking the "Start" button, a new window opens (see. Figure 4) in which a list of training phrases appears. When selecting one or another test phrase, there is an additional window where you can choose from 3 options: "Show pitch", "Show energy" or "Show spectrum".

When selecting the "Developer mode" one can see the full list of test phrases (see Figure 5). In this mode one can

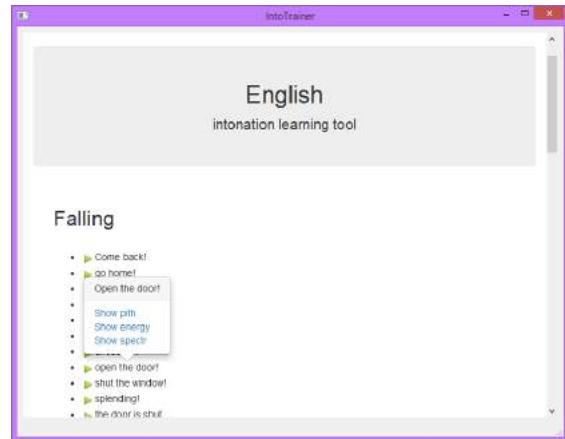


Figure 4. The page of the test phrases selection

choose additional options clicking the buttons: "Evaluation", "Record", "Play", "Remove", "Rename" and "Settings".

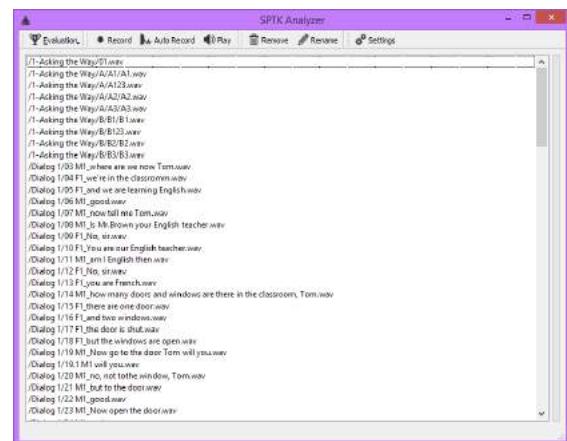


Figure 5. The "Developer mode" window

When selecting the "Evaluation" in the "Developer mode" window one can see the acoustic parameters of the test phrase (Figure 6), or the smoothed F0-parameter of the test phrase (Figure 7), or the F0-correlation between the test phrase and the same phrase pronounced by student (Figure 8).

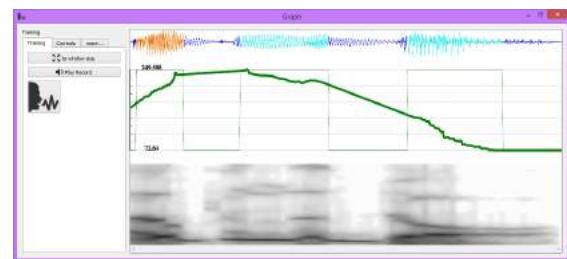


Figure 6. Showing of the acoustic parameters of the test phrase

When selecting the "Settings" in the "Developer mode" window one can change some speech analyzer parameters, such as: frame size, frame shift, type of frame windows, LPC order and others (see Figure 9).

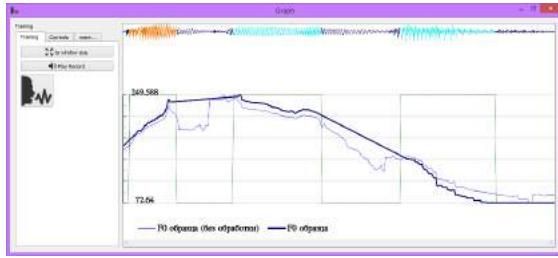


Figure 7. Showing of the smoothed F0-parameter of the test phrase



Figure 8. Showing of the F0-correlation between the test phrase and the same phrase pronounced by student

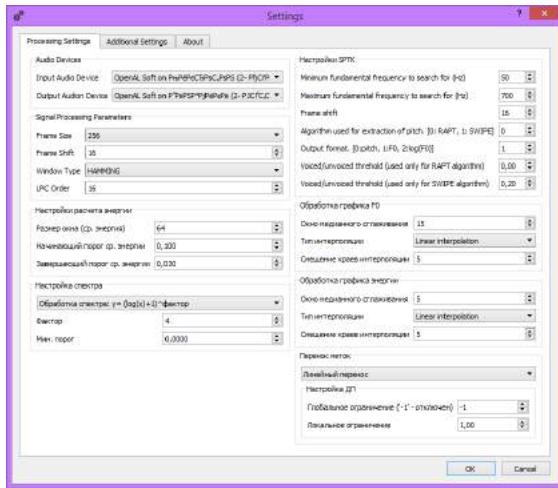


Figure 9. The “Settings” mode window

V. CONCLUSION

We believe, that there is a great potential in both domestic and international markets for a new and innovative product such as the proposed computer system for intonation training integrated into a foreign language educational courseware. To our knowledge, there is no satisfactory software available for such teaching system and, therefore, such system appears to be of great relevance. For example, there are considerable intonational differences between Russian and English languages. American English native speakers made the following interesting observation: “Ask an average American what they are thinking about the Russian accent, and the answer will be as follows: ”Russians don’t sound very friendly. I feel like they don’t like me at all. I am not sure whether it comes from their language or from their culture?“ One of the reasons why native Russian speakers sound unfriendly in English is the so called “flat” tone. Native Russian speakers often do not use language-

specific phonological representation of intonation during their conversation in English. Moreover, native Russian speakers tend to avoid using rising and falling intonations in English and, as a result, Americans may find their speech unfriendly and unpleasant.

Application of an intonation mapping analyzer as a part of a speech recognition system is expected to increase reliability of recognition through the prominence of accented words and intonational segmentation of a speech flow. Intonation analysis will also be helpful for subsystems of identification of individual and emotional factors of speaker’s speech. The use of intonation system in speech synthesis systems will give an opportunity to improve intonational prominence of synthesized speech so that it will positively affect listener’s comprehension.

REFERENCES

- [1] Chun, D. M. (1988). The neglected role of intonation in communicative competence and proficiency. *Modern Language Journal*, 72, 295-303.
- [2] Juan Arias, Nestor Yoma, Hiram Vivanco. (2010). Automatic intonation assessment for computer aided language learning. *Speech Communication* 52 254–267
- [3] Lobanov, B. (2006). Language- and speaker specific implementation of intonation contours in multilingual TTS synthesis / B. Lobanov, L. Tsirulnik, D. Zhadinets, E. Karnevskaya // Speech Prosody: proceedings of the 3-rd International conference, Dresden, Germany, May 2–5, 2006. – Dresden, – V. 2. – pp. 553-556.
- [4] Lobanov, B. (2014). Universal Melodic Portraits of Intonation Patterns of Russian Speech / Lobanov, B., Okut, T. // Computational Linguistics and Intellectual Technologies: Papers from the Annual International Conference “Dialogue”. Issue 13 (20). — Moscow. : RSHU, — pp. 330-339.
- [5] Lobanov, B. Comparison of Melodic Portraits of English and Russian Dialogic Phrases / Lobanov, B. // Computational Linguistics and Intellectual Technologies: Papers from the Annual International Conference “Dialogue”. Issue 15 (22). – Moscow.: RSHU, 2016. – pp. 382-392.
- [6] R. Ogden, at al. (2000) “Prosynth: an integrated prosodic approach to device-independent, natural-sounding speech synthesis,” / R. Ogden // Computer Language and Science, pp. 177–210.
- [7] Odintsova, I. (2011). Sounds. Rhythmic. Intonation. / I. Odintsova // Moscow.: Flinta-Science, pp. 253.
- [8] Ockenden M. (2005). Situational Dialogues / M. Ockenden // The English Centre, Eastbourne / Revised Edition. -Longman – pp. 98.
- [9] Boris Lobanov, Tatiana Levkovskaya (1997). “Continuous Speech Recognizer for Aircraft Application “Proc. of International Conference «Speech and Computer» SPECOM’97, Napoca, Romania, pp. 817-820.

ПРОТОТИП КОМПЬЮТЕРНОЙ СИСТЕМЫ ОБУЧЕНИЯ РЕЧЕВОЙ ИНТОНАЦИИ

Лобанов Б.М., Житко В.А.

В статье обосновывается важность интонации для восприятия и понимания речи человеком. Описаны фундаментальные принципы интонационной теории речи, которая основывается на концепции универсальных мелодических портретов. В статье описываются основные алгоритмы анализа и интерпретации фразовой интонации, которые лежат в основе разработанной компьютерной системы, предназначеннной для обучения речевой интонации. Показаны выходные характеристики системы и обсуждается её полезность в обучении интонации иностранного языка.