

# A subsystem of the “IntonTrainer” for learning and training Belarussian intonation

Evgeniya Zenovko

*The State Scientific Institution*

*«The United Institute of Informatics Problems  
of the National Academy of Sciences of Belarus»*

Minsk, Belarus

evgeniakacan@gmail.com

Boris Lobanov

*The State Scientific Institution*

*«The United Institute of Informatics Problems  
of the National Academy of Sciences of Belarus»*

Minsk, Belarus

lobanov@newman.bas-net.by

**Abstract**—The article presents Belarussian module of program complex "IntonTrainer" for training learners in producing a variety of intonation patterns of Belarussian speech. The system is designed for visual learning basic Belarussian tonal types, comparing melodic portraits of a reference phrase and a phrase spoken by a learner. It involves an active learner-system interaction, where the user can not only listen to an audio, but also try to repeat correctly this phrase and achieve the best result. The main principle of the program is to analyse and compare student's results with standard variants without teacher's involvement.

**Keywords**—program complex, intonation, tone, melodic contour, melodic portrait, linguaacoustic database, module, reference phrase

## I. INTRODUCTION

Today experimental studies of oral speech are of particular relevance, since they reveal the qualitative characteristics of the phonetic, and directly, prosodic means in the implementation of the speaker's communicative intention and the level of information perception. The prosodic level is the least studied and developed in comparison with phonetic, phonematic or other levels of the language. However, in recent decades it has acquired an increased interest in this area of the language due to advances of information society and the need to update computer technology. Modern software market offers a wide selection of computer-based training systems with an emphasis on vocabulary, grammar and syntax. These applications offer not only explanations of grammatical material, examples, texts, tests, but also exercises for listening and pronunciation. Some of them include tasks for listening to a speech and sound recording. Another use the interval repetition method, when at certain periods you are asked to repeat the words that have already been studied for their final fixation in memory. It is important that educational applications include various methods of computer monitoring and error correction, thereby individualizing the learning process, avoiding the subjectivity of assessment and expanding the possibilities of independent work for students. Unfortunately, we should notice, that all the variety of programs are produced for

different languages including Russian but not for the Belarussian language. We can only find bilingual thematic vocabularies, interactive phrasebooks and some special programs. This fact demonstrates the necessity to create new applications for our native language for the purpose of its promotion and possibility to learn it in more detail.

Rhythmic-intonational skills directed on intonationally and rhythmically correct design of speech (the correct statement of stress, the observance of the necessary melody, rhythm, etc.) call for special attention. For such skills we need special software complexes for studying prosody, in particular intonation, and its components. Computer programs and systems designed for effective teaching of intonation must meet the following requirements and be able to:

- 1) detect deviations in the pronunciation of the learner from the codified pronunciation variants (words and phrases);
- 2) classify these deviations;
- 3) show the learner the difference in pronunciation and correct pronunciation;
- 4) have the means (video, graphics, exercises) to bring the pronunciation implementation to the necessary degree of correct pronunciation;
- 5) act as a simulator of pronunciation skills;
- 6) to have a mechanism for evaluating the correctness and correctness of achieved results.

The IntonTrainer fills all the bill of the same. It is designed to analyze, display and compare intonation (pitch) contours of the reference and spoken phrases, as well as a numerical evaluation of their intonation similarity. Evaluation of intonation similarity is carried out on the basis of the representation of intonation in the form of Melodic Portraits. For now the software includes subsystem modules with the sets of reference phrases that represent the main intonation models of Russian, English (British and American versions), German and Chinese speech. In addition, modules for emotional intonation of the Russian and English languages have been developed. Also in the program you can find song modules of these

two languages. The main advantage of this program is an opportunity to create new modules for different languages, including Belarusian.

This work is a follow up study to the previously introduced model of universal melodic portraits (UMP) of accentual units (AU) for the representation for Belarusian intonation [1, 2]. The article depicts Belarusian module of IntonTrainer, its components and principles of its usage. The initial window of the Belarusian module is shown in Figure 1. For more detailed informa-

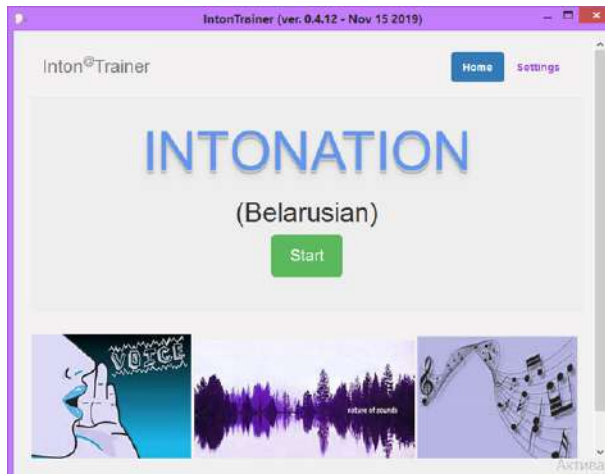


Figure 1. The initial window of the Belarusian module

tion you need to turn to official cite of the complex (<https://intontrainer.by> ).

## II. LINGUA-ACOUSTIC DATABASE FOR BELARUSIAN MODULE

Belarusian module is based on a lingua-acoustic database (LAD), which includes a large set of standard phrases. These reference phrases cover a greater number of Belarusian intonational patterns of speech according to the type of expression, stylistic features and their emotional coloring. The Belarusian module contains audio database of a female and a male voicies. They allow you to get acquainted with the four main melodic tones of the Belarusian literary language (falling, rising, falling-rising and rising-falling) on the examples of reference phrases, which reflect melodic contour of each phrase (Figure 2). In addition to information about the tone and the communicative type of sentence, the database depicts prosodic markup of sentences and the number of syntagmas in a phrase. The user can also listen to the audio file by clicking on the link.

Furthermore, LAD displays 26 groups of statements which provide a description of intonational features of Belarusian language in terms of the syntactic structure of a sentence, emotional coloring and a number of Syntagmas. Today Belarusian linguaacoustic database consists of 200 syntactic-accentual units. The module is

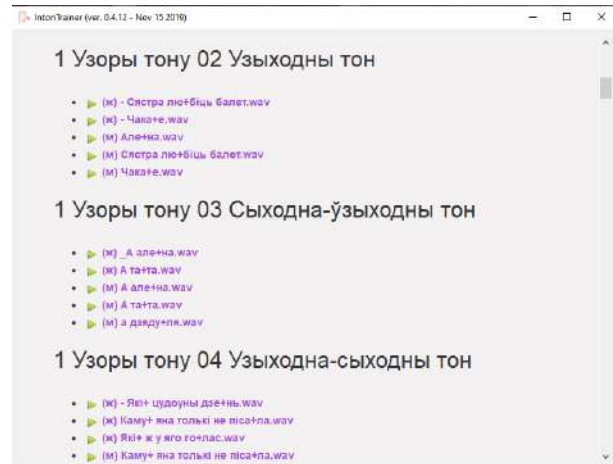


Figure 2. Database window of Belarusian module

in the process of refining, so in the future it will be added universal melodic portraits of phrases related to prose texts, poems and dialogues. The Belarusian module performs all the functions of IntonTrainer complex that allows you to learn the intonation of the language in its entirety.

## III. A STRUCTURE AND FUNCTIONS OF THE BELARUSIAN MODULE

The Belarusian module consists of 7 sections. First of all the program offers to get acquainted with basic tones (falling, rising, rising-falling and falling-rising) on the examples of sentences with different communicative types (figure 2). Then the user can learn intonation patterns of the phrases with different emotional colouring and types of intentional focus (for example sentences with neutral answer, alternative questions, appeals, forms of greeting and farewells, etc). Each example provides audio and visual representation of UMP, pairwise comparison of different Tone patterns (TP), explains peculiarities of TP usage.

When the student chooses the pattern for learning, he opens "Analysis results window" where the red column on the left shows the range of the melody change, i.e. frequencies of the pitch (F0), expressed in octaves (figure 4). On the right, a linear graph of the UMP is displayed in red, the nuclear syllable of which is marked with frequent vertical lines. Below the graphs, the minimum and maximum values of F0 for the selected phrase are listed, as well as the text of the phrase in which the nuclear vowel is indicated by the "+" sign. Listening to the selected phrase is carried out by pressing the «Play Record» button. An additional function of the program is the ability to select the markup of the studied phrase. At early version of the program it was necessary manually processed audio, highlighting pre-nuclear, nuclear and post-nuclear syllables. To date, it is also possible automatically to produces a tonal contour of

the phrase. Therefore, the user can choose the view mode of UMP: "Manually" or "Automatically" (see Figure 3, 4).

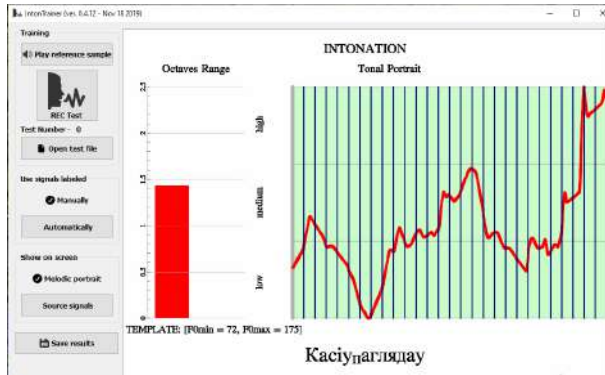


Figure 3. Melodic contour of the phrase "Kamu+ yana tolki ne pisa+la" marked manually

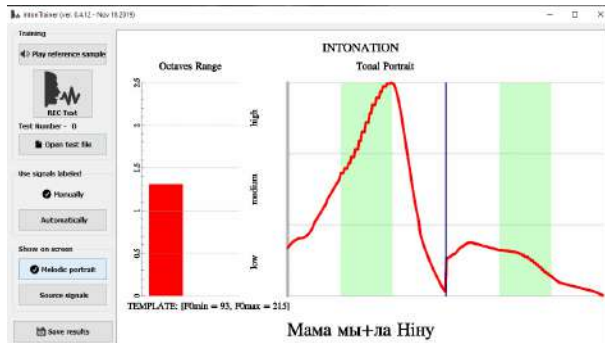


Figure 4. Melodic contour of the phrase "Kamu+ yana tolki ne pisa+la" marked automatically

If one wants to see graphical display of intonation contour with prenuclear, nuclear and post-nuclear marks and the oscillogram of the phrase, then needs to click "Show original" (figure 5).

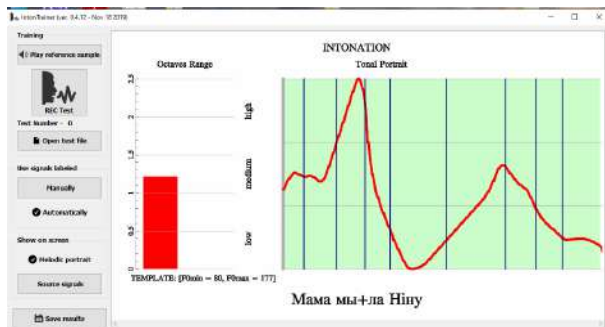


Figure 5. Additional window displaying F0 in real time

The next step is to get acquainted how to listen to reference phrase, repeat, learn it and receive the best result of learning. The system evaluates the student's pronunciation on a ten scale with percentage. The result

of the similarity of his pronunciation with the standard phrase is displayed on top of the screen (figure 6).

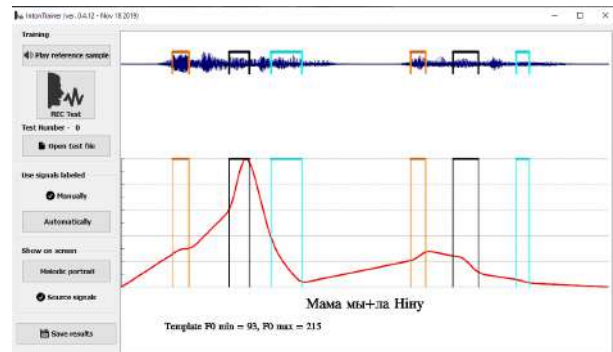


Figure 6. Comparison of reference phrase and pronounced phrase "Kamu+ yana tolki ne pisa+la"

The main advantage of this version is the ability to set the desired number of test files (pronunciations) and to obtain an average rating of their similarity with the reference phrase. The left part of the figure 7 (Octaves Range) graphically shows the results of calculating the similarities in the pitch frequency range for each pronunciation (thin columns, the last pronunciation corresponds to black) and the resulting averaged score (thick column). It shows the numerical value of the average score - the Range is 49 per cent. The right side (Tonal Portrait) shows the results of calculating the UMP (or NMP) of the spoken phrases (thin lines) and the resulting averaged curve (thick line). Above is the numerical value of the average score of similarity with the reference curve - Shape (58 per cent). Therefore, the user can see his progress from the very beginning of studying to the end.

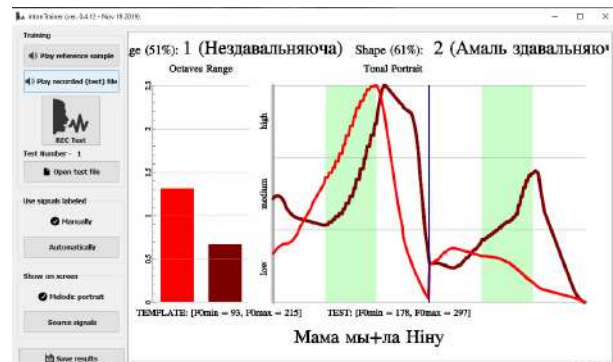


Figure 7. Comparison of reference phrase and multiple pronounced phrases "Mama my+la Ninu"

When you click the "Save Results" button, an additional icon appears and a page opens in EXCEL, on which a complete set of 10 prosodic features of the reference phrase is written (see Table 1). The obtained data is stored in the same folder where the reference phrase being studied is stored. In table 1, in addition to data on the parameters of the original NMP curve, data

on the value of its time derivative,  $d / dt$  (NMP), are also presented. Additional analysis of the parameters of the derivative of the NMP turns out to be useful for taking into account the dynamic characteristics of the CHO movement, which are characteristic of certain types of emotions.

Table 1 - Qualitative measures of the reference and pronounced phrase "Dama my-ia Nima"

| The data on the prosody                | Prosody     | Database | Reference data on intonation and in duration   | Reference data on Records                  |
|--|-------------|----------|--|--|
| Prosody this curve "Comparison"        | 70          | 36       | Mean [Hz] 1,575300285                          | Mean [Hz] 180                              |
| Prosody this curve "Integral"          | 71          | 29       | Mean [Hz] 1,7113366194                         | Mean [Hz] 182                              |
| Prosody this curve "Sum"               | 33          | 87       | Mean [Hz] 1,266566239                          | Mean [Hz] 1,6075002205                     |
| Average of the three above prosodies   | 54          | 87       | Mean [Hz] 1,575300285                          | Mean [Hz] 180,5                            |
| Prosody this "Range"                   | 87          | 31       | Mean Value of the curve NMP 1,1588018156       | Mean Value of the curve NMP 26,79450370    |
| Reference data on intonation           |             |          | Center of the curve NMP 0,38886705             | Center of the curve NMP 48,56527859        |
| Phrase [Hz]                            | 255         |          | Width of the curve NMP 0,344750157             | Width of the curve NMP 42,2288879          |
| Phrase [Hz]                            | 39          |          | Mean Value of the Derivative curve -0,25315731 | Mean Value of the Derivative 15,77518352   |
| Phrase [Hz]                            | 2,515017917 |          | Center of the Derivative curve NMP 1,288588819 | Center of the Derivative curve 25,32033359 |
| Register [Hz]                          | 154         |          | Width of the Derivative curve NMP 1,7019919    | Width of the Derivative curve 47,5879685   |
| Mean Value of the curve NMP            | 26,79450370 |          | Width Sounds Level 0,344750157                 | Width Sounds Level 42,2288879              |
| Center of the curve NMP                | 48,56527859 |          | Width Sounds Duration 0,75866761               | Width Sounds Duration 258                  |
| Width of the curve NMP                 | 42,2288879  |          |  |  |
| Mean Value of the Derivative curve NMP | 15,77518352 |          |  |  |
| Center of the Derivative curve NMP     | 25,32033359 |          |  |  |
| Width of the Derivative curve NMP      | 47,5879685  |          |  |  |
| Width Sounds Level                     | 42,2288879  |          |  |  |
| Width Sounds Duration                  | 258         |          |  |  |

#### IV. CONCLUSION

The software package IntonTrainer is recommended for the initial acquaintance and study of the main intonational constructions of the Belarusian literary language, comparison of its different patterns, as well as their implementation in dialogue, prose and poetry. As a means of training intonation, the IntonTrainer gives an opportunity to improve intonation skills in various fields of activity. Today it is the only free program that offers to study several languages at once. The complex also allows to create new language modules based on its platform. Thus, the authors of this article have developed the first Belarusian language module, which makes it possible to learn how to use the Belarusian intonation correctly. It can be used as a means for teaching Belarusian as a foreign language, get practical results without involving a teacher and compare its melodic structure with another languages. The significance of such a development for the Belarusian language lies in the universality of its use by both students and linguists in scientific and practical research.

#### REFERENCES

- [1] IntonTrainer. Available at: <https://intontrainer.by/> (accessed 12.03.2019).
- [2] Anne Bonneau, Vincent Colotte.: Automatic Feedback for L2 Prosody Learning. Speech and Language Technologies, (2011), pp.55-70.
- [3] Juan Arias, Nestor Yoma, Hiram Vivanco.: Automatic intonation assessment for computer aided language learning. Speech Communication 52 (2010), pp. 254–267.
- [4] Lobanov, B.M. The program model of analysis and visualization of emotional intonation in oral speech / B.M. Lobanov, V.A. Zhitko // Open semantic technologies for designing intelligent systems: a collection of scientific papers. Issue 3. / BSUIR; Editorial: V.V. Golenkov [et al.]. - Minsk: BSUIR, 2019. - C. 297-300.
- [5] Lobanov B. M. Komp'yuternyy sintez i klonirovanie rechi [Computer synthesis and speech cloning]. Minsk, Belorusskaya nauka Publ., 2008. 357 p.
- [6] Lobanov, B. A Prototype of the Software System for Study, Training and Analysis of Speech Intonation / B. Lobanov, V. Zhitko, V. Zahariev // Speech and Computer: 20th International Conference, SPECOM 2018, Leipzig, Germany, September 18–22, 2018, Proceedings / Eds. Alexey Karpov, Oliver Jokisch, Rodmonga Potapova. – Springer, 2018. – P. 337-346.

- [7] Lobanov B. Software Subsystem Analysis of Prosodic Signs of Emotional Intonation / B. Lobanov, V. Zhitko // Speech and Computer: 21st International Conference, SPECOM 2019, Istanbul, Turkey, August 20–25, 2019, Proceedings / Eds. Albert Ali Salah, Alexey Karpov, Rodmonga Potapova. – Springer, 2019. – P. 280-288.
- [8] Nikolaeva T.M. Frazovaya intonaciya slavyanskikh yazykov [Phrasal intonation of Slavic languages]. Moskva, Nauka Publ., 1977. 278 p.
- [9] Vigonnaya L.T. Intanacyja. Nacisk. Arfaepija [Intonation. Stress. Orthoepy]. Minsk, Nauka i tehnika Publ., 1991/ 215 p.
- [10] Zianouka Ja.S. Linguaacousticresources for Belarusian speech synthesizers. Open semantic technologies for designing intelligent systems: a collection of scientific papers. Issue 3. / BSUIR; Editorial: V.V. Golenkov [et al.]. - Minsk: BSUIR, 2019. - C. 301-304.

### Подсистема «IntonTrainer» для обучения и тренировки белорусской интонации

Зеновко Е.С., Лобанов Б.М.

В статье представлен белорусский модуль программного комплекса «IntonTrainer» для обучения учащихся выработке разнообразных интонационных паттернов белорусской речи. Система предназначена для визуального изучения основных белорусских тональных типов, сравнения мелодических портретов референтной фразы и фразы, произносимой учеником. Он включает в себя активное взаимодействие ученика и системы, где пользователь может не только слушать аудио, но и пытаться правильно повторить эту фразу и достичь наилучшего результата. Основным принципом программы является анализ и сравнение результатов студента со стандартными вариантами без участия учителя.

Received 14.01.2020