Inton Trainer

SPEECH INTONATION ANALIZER&TRAINER

https://intontrainer.by/

LAST NEWS:

(Version: 2019-12-02)

- *The ability to set the desired number of test files (pronunciations) and to obtain an average rating of their similarity with the reference phrase (see section 8.3) is added*
- *An auxiliary software module (Multi-Lingual Launcher) for launching any desired set of IntonTrainer modules is added (see section 8.4)*
- *The new IntonTrainer modules: Belarusian IntonTrainer demo and Singing Lessons IntonTrainer demo are added*

(Version: 2019-08-22)

- *Added displaying the position of the boundaries of voice sections and the boundaries of accent units of phrases.*

(Version 2019-05-04

- *The ability to analyze and compare the melodic portraits without the need for preliminary manual marking of it into pre-nuclear, nuclear and nuclear sections is added.*
- *The possibilities of parametric display of melodic intonation portraits are expanded.*
- *The ability to analyze and store in numerical form a set of prosodic signs of the analyzed phrases has been added.*
- *The accuracy of segmentation and labeling of the analyzed speech signals is increased.*
- *Updated user interface.*

ATTENTION, PLEASE!

- ✓ **"IntonTrainer" is an open system and allows modification of the settings used and audio data.**
- ✓ **Reference databases - PATTERNS - can be supplemented or modified by the user in accordance with the task, or re-formed to work with new language applications.**
- ✓ **Please send your questions, suggestions and comments to: intontrainer@gmail.com**

SPEECH INTONATION ANALIZER AND TRAINER

- *USER GUIDE* -
(On example of British English)

1. Goals

The ***"IntonTrainer"*** software module (hereinafter referred to as **SWM**) is designed to analyze, display and comparison 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 (MP), namely:: universal - **UMP**, or normalized - **NMP**. The present version of the **SWM** also provides for obtaining numerical values of various kinds of prosodic features of the analyzed phrases described below.

The software package is recommended for use in the following popular fields:

- **In linguistic education used as a means of visualizing intonation**.

Primary introduction and study of the basic tone patterns of oral speech and songs, their pairwise comparisons, peculiarities of their usage as well as their actualization in dialogues prose and poetry.

- **In individual intonation training for correct pronunciation used as a means of feedback**.

Individual training for correct pronunciation of tone patterns when studying a foreign language or improving intonation skills of one's native language in some professions: call center operators, radio and TV announcers, etc.

- ***In scientific and practical research used as a means of comparing intonation from different sources***.

Study of individual, emotional and stylistic features of intonation. Comparative evaluation of speech and singing intonation in norm and pathology. Estimation of the intonational quality of synthesized speech.

The prototype of the IntonTrainer system is implemented as a separate SWM under Windows OS (xp, 7, 8, 10) and Linux (Debian, Ubuntu). The implementation uses the Qt library as a debugging user interface and SWM kernel, and the MathGL library for plotting graphs. To implement the main user interface, an SWM written in HTML / JS is used. Algorithms for calculating and processing data are written in C using the GNU Scientific Library, Speech Signal Processing Toolkit (SPTK) and OpenAL libraries.

2. Getting Started

The initial SWM window (it opens after the SWM is started) is shown in Fig. 1.

Fig. 1 The initial window of the SWM

Before you start, you can preview the settings of the SWM and correct them. Clicking the button - **Settings** - opens the **Main settings window** (see fig. 2), which provides ample opportunities for changing and setting the desired personal settings.

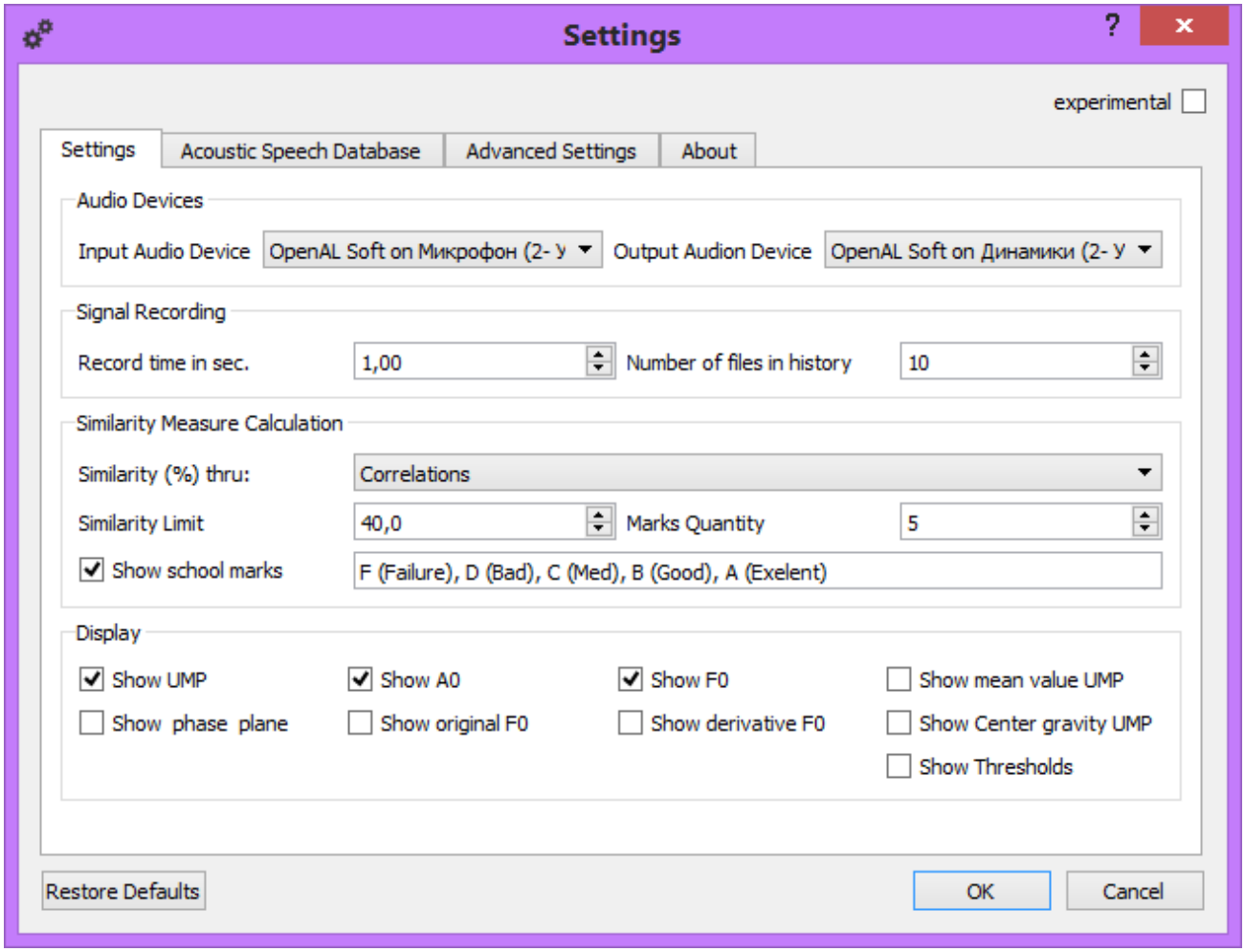

Fig. 2. Main settings window

In the **main settings window** the user can select:

- In the **Audio Devices** section, the user can select the type of audio devices used.
- In the section **Signal Recording**, you can select the length of the recording of the signal from the microphone - **Record time in sec**. In this case, recording takes place within N seconds + the duration of the selected phrase pattern. In addition, it is possible to save in the folder "RECORDS" the specified number of phrases recorded from the microphone - **Number of files in history**,
- In the **Similarity Measure Calculation** section, you can choose a method for assessing the intonation proximity of the spoken phrase to the reference one, using various similarity measures:- **Similarity (%) thru**:

1) cross-correlation - **Correlations**,
2) average value of the vectorl distance **– [1- Average Distance] (%)**,
3) maximum of the local distance **– [1 - Maximum Local Distance] (%)**,
4) average value of the three above-mentioned similarity measures – **Average (%)**.

The chosen method of calculating the intonational similarity is then used in calculating the school assessment of the intonation quality of the spoken phrase. To do

this, a checkmark is made in the **Show school marks** small window. For the selected method, a **Similarity limit** is defined, which corresponds to the worst scoring score. The total number of points used to assess the intonation quality is set by the desired number of them - **Marks Quantity**. By filling in the section - **Mark labels** - digital and verbal names of points are specified.

In the **Display** section, the user can select the desired mode of displaying the results of the intonation analysis of the studied signals. The default setting is marked with checkmarks. Features of other modes will be discussed below.

Two additional sections of settings (Fig. 2) - **Acoustic Speech Database** and - **Advanced settings** - are designed for experiments on the selection of the optimal mode, carried out by developers or "advanced" users of the program.

Button - **About** opens a window with information about the developers and a number of other information.

Returning to the original settings is performed by clicking the Restore defaults button (see the lower left corner in Figure 2).

3. Primary study of the basic tone patterns

After clicking the "**Start**" button, the **DB information window** opens (Fig. 3) containing a structured list of the phrases with different **Tone Patterns** examples of English speech, as well as other examples (see fig. 3), stored in the **"PATTERNS**" folder.

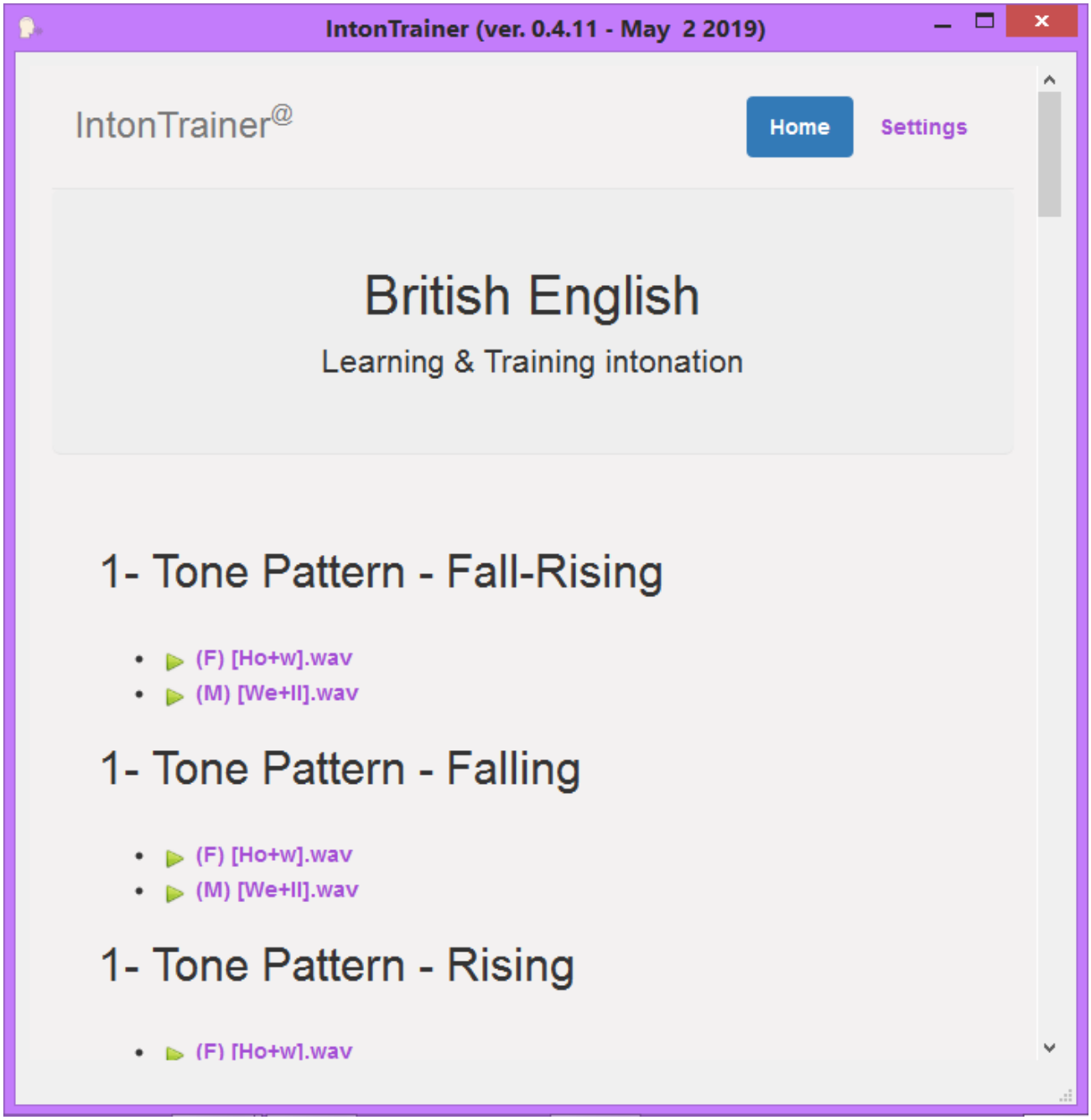

Fig. 3. DB Information window

By scrolling the page of the **DB information window** from top to button, the user is given the opportunity to see a lot of examples for the main tone patterns of English speech. The set of examples provides audio and visual representation of the main tone patterns

(TP1-TP3), pairwise comparison of different TPs, shows peculiarities of TP usage, as well as TP actualization in dialogues, prose and poetry.

On the page of the information window (Fig. 3), the user gets the opportunity of hearing and visual acquaintance with the basic intonation structures of English speech by selecting the required directory using the cursor, for example:

3- USSAGE - Falling-Rising (Exclamation) • She has ni+ce][ey+es!.wav

As a result of this choice, by default, the multifunctional main SWM window opens (see Figure 4).

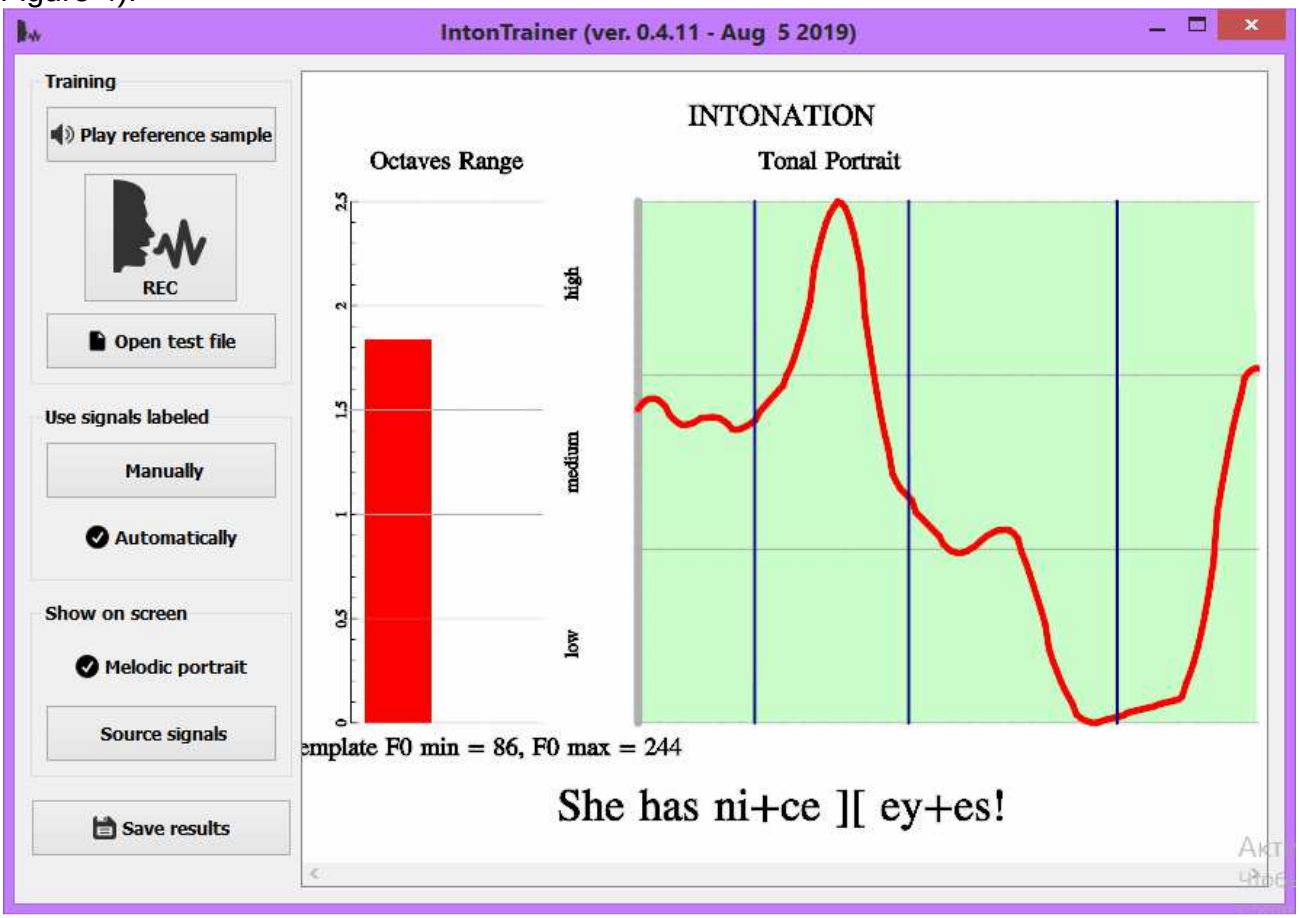

Fig. 4. Main window. Mode: "Automatically" - "Melodic portrait"

The left side of the window contains controls that allow listening to a speech signal, as well as setting various modes for graphical display of intonation analysis results.
The default graphical display modes are marked with checkmarks, namely:

- automatic signal marking method: (***Use signal labeled - Automatically***),
- view of the screen display of the result of the analysis: (***Show on screen - Melodic portrait***).

The right side of the window graphically shows the results of the analysis of the reference phrase. By default, the **Normalized Melodic Portrait (NMP)** of the analyzed phrase is displayed in the right part of the window (Fig. 5). Under NMP means the display of the melodic curve, which is normalized with respect to the minimum and maximum values of the F0(t) and the total duration of the voice sections of the phrase. The NMP curve displays the dynamics of the change in the frequency of the main tone (F0) only in the voiced sections of the phrase. The boundaries of the voiced sections are represented by vertical lines. The construction of the NMP curve does not require preliminary “manual”

marking of the analyzed speech signal of the phrase. Signal segmentation into voice regions (see Section 7) is carried out automatically on the basis of information about the presence of periodicity in the signal (voice) with the simultaneous presence of a sufficiently high signal amplitude - A0 (t).

Below the graph of the NMP curve is the minimum and maximum values of the F0 phrase in Hertz, as well as the text of the phrase in which the nuclear vowel is indicated by a **"+**" sign.

The red column (to the left of the melodic curve) displays the range of the pitch frequency change in the phrase, expressed in octaves:

Octave's Range = (F0max / F0min) - 1.

Listening to the selected reference phrase is performed by pressing the "**Play reference sample**" button.

In fig. 5 shows the main window of the PC when selecting another method of screen displaying the analysis result: (**Show on screen - Source signals**).

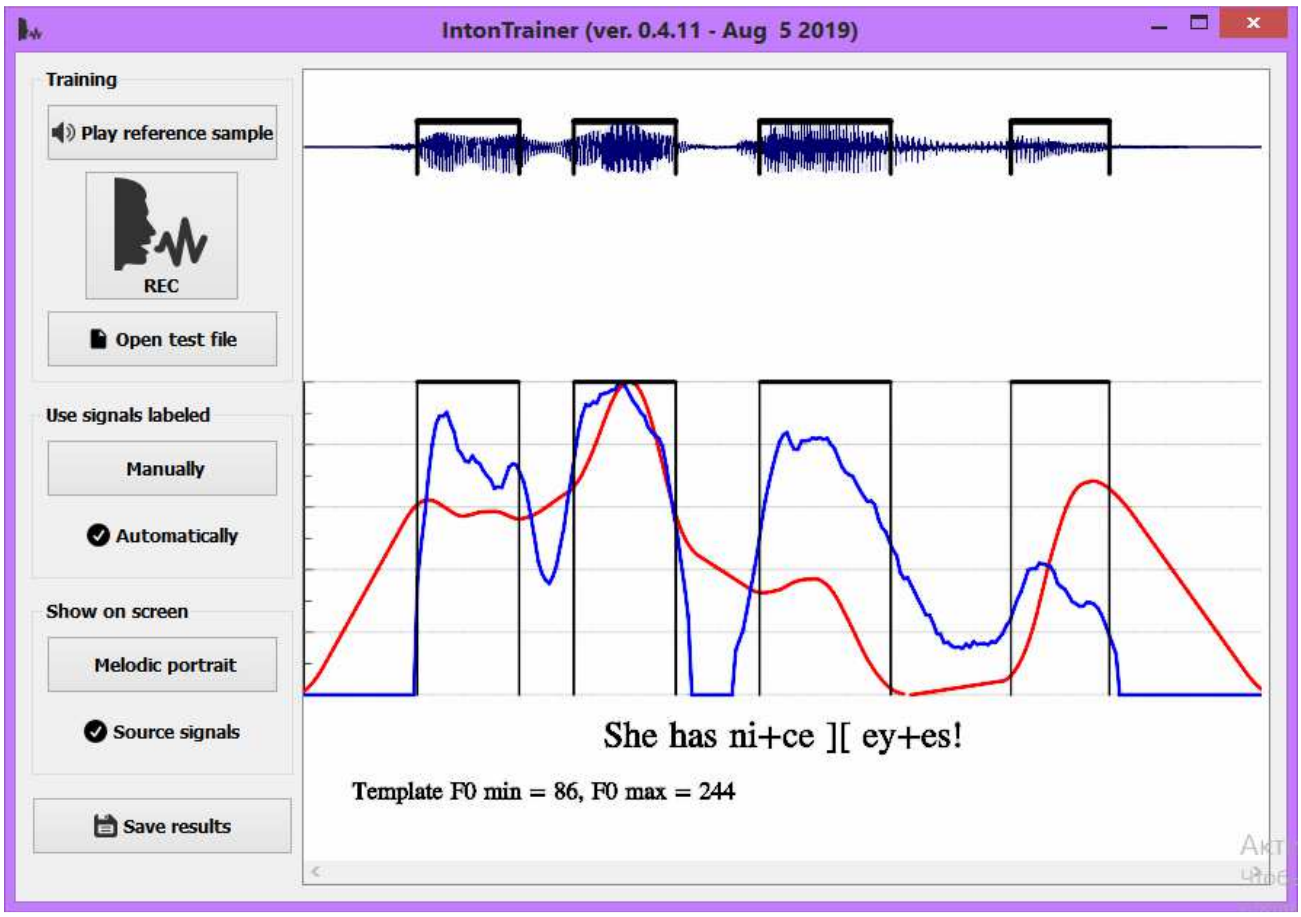

Fig. 5. Main window. Mode: "Automatically" - "Source signals"

In the upper part of the window, the waveform of the phrase is displayed, indicating the voice sections. In the middle part of the window, the F0 (t) curve is shown, as well as the signal amplitude change curve A0 (t). The SWM mode **"Source signals"** is used to control the analysis results in cases when doubts arise about the correctness of the markup on the voice sections necessary to adequately display the NMP phrase.

In fig. 6 shows the main window of the PC when choosing the following method of on-screen display of the analysis result: **(Use signal labeled - Manually)** and (**Show on screen - Melodic portrait).**

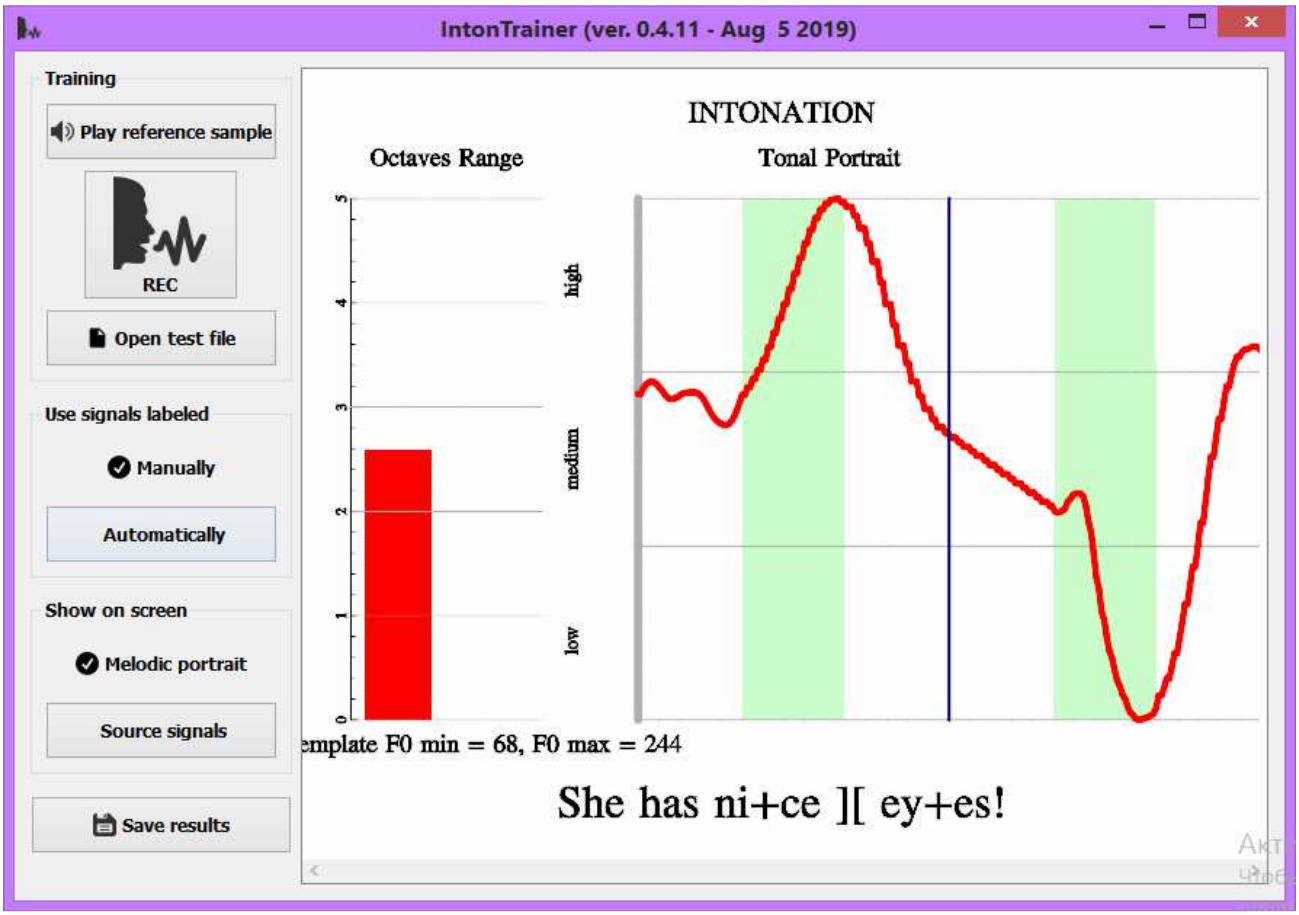

Fig. 6. Main window. Mode: "Manually" - "Melodic portrait"

In this mode, the F0 change curve in a phrase is displayed as a **Universal Melodic Portrait (UMP**), consisting of three time-normalized sections: pre-core, core, and post-core. The core is marked with frequent vertical lines. Fig. 6 shows the result of constructing the UMPs of the phrase "***She has nice eyes***!", consisting of 2 accent units (AUs). In Fig.6, their border is shown as a vertical line, and in the text, with an "][".

In contrast to the NMP, for the construction of the UMP, a preliminary manual marking of the phrase signal to the sections of the pre-core, core and post-core is required (see Section 7).

Below below the graphs are the minimum and maximum values of F0 for the selected phrase, as well as the text of the phrase in which the nuclear vowel is indicated by a "+" sign.

The main window view of the SWM when choosing the following method of on-screen display of the analysis result **(Show on screen - Source signals**) is shown in fig. 7.

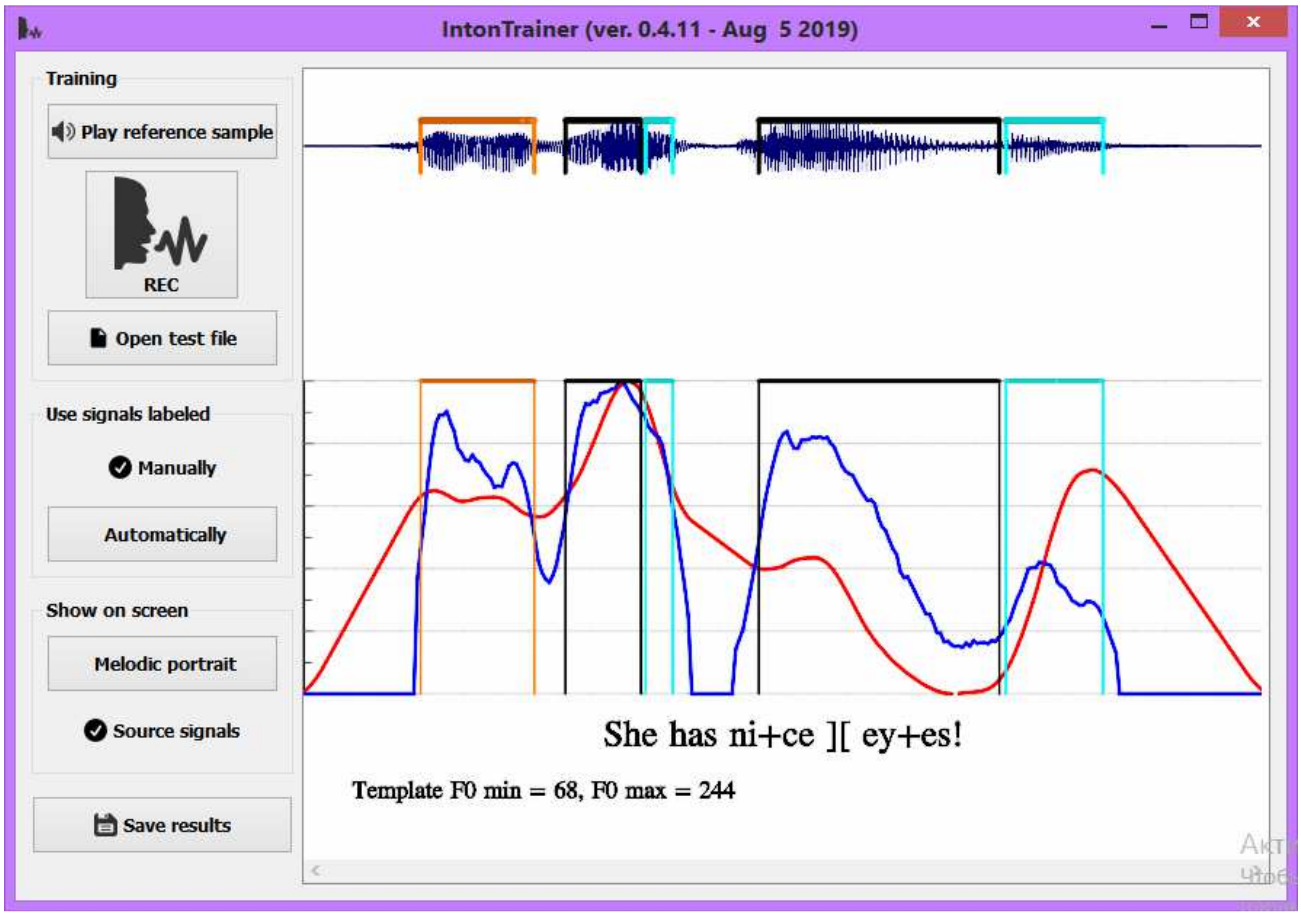

Fig. 7. Main window. Mode: "Manually" - " Source signals"

In the upper part of the window, the waveform of the phrase signal is displayed with indication of the pre-core marks (red line), the cores (black line) and behind-core marks (the blue line). In the middle part of the window, the curve of F0 in real time is shown in red. On this curve are indicated the sections of the pre-core, core and behind-core, from which the perpendicular magnetic field is formed, shown in fig. 7. The blue color represents the signal amplitude change curve - A0 (t).

The function “Source signals” is used to control in cases where there are doubts about the correctness of the manual markup for obtaining the UMP phrase.

Thus, the user, looking through a structured list of reference phrases in a manner described above, carries out a primary study of the features of the implementation of the basic intonation constructions of English speech.

4. Individual intonation training for correct pronunciation

When using the **SWM** for individual intonation training the user must use an extended or built-in microphone. In this case, the user should press the **"Rec"** button, wait for a short "beep-signal" and pronounce the phrase into the microphone. The text of the phrase is indicated in the lower part of the window (see: Fig. 4 - 7). After recording (to the **"RECORDS"** folder) and processing of the entered speech signal, the user will hear the second "beep-signal", and the image in the graph window shown in Fig. 6 will be replaced with the image shown in Fig. 8.

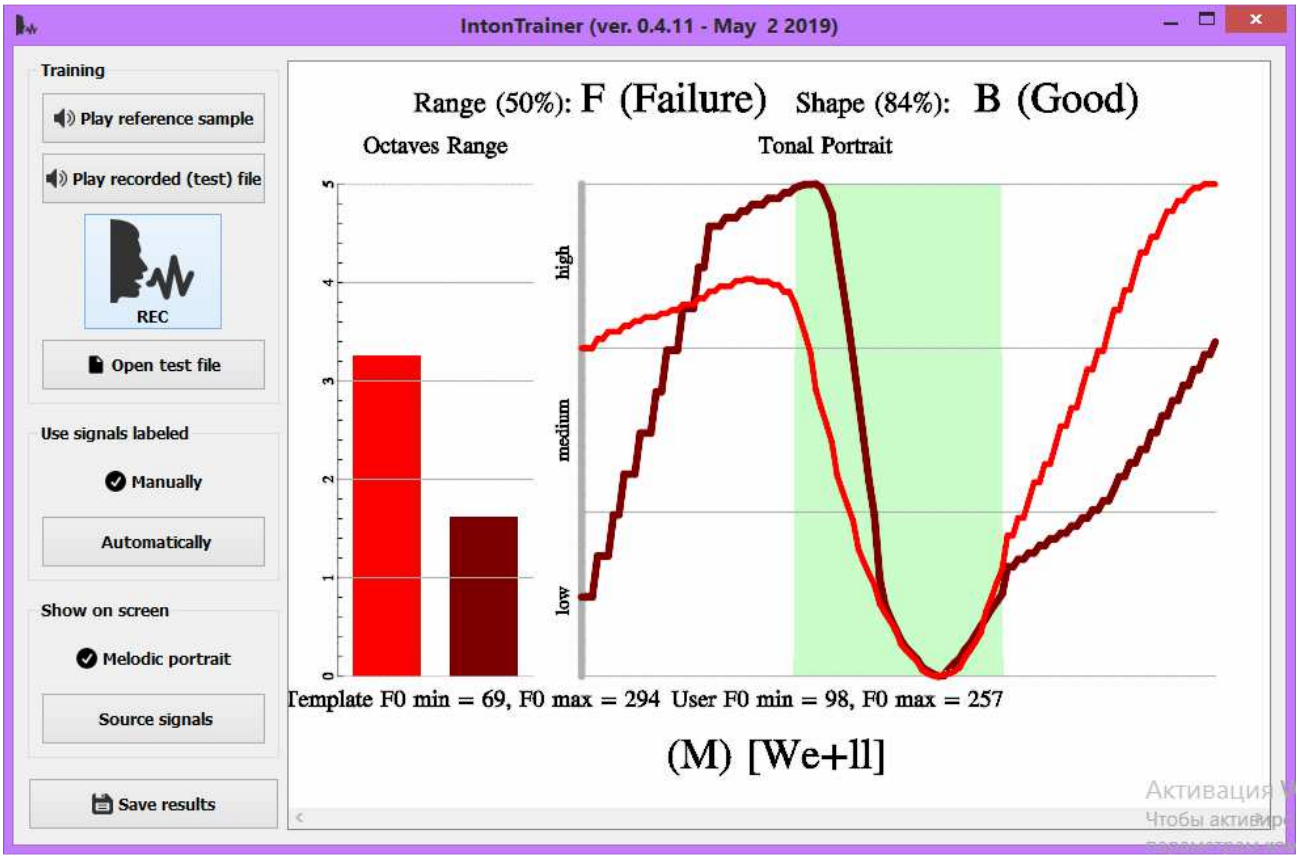

Fig. 8. The window displaying the results of analysis and comparison. Mode: "Manually" - "Melodic portrait"

The upper part of the window shows the results of comparison of the reference **(red line)** and spoken **(brown line)** phrases: **Range (50%)** - proximity to the range of changes F0 and **Shape (84%)** - proximity to the shape of the trajectory F0. Near the percentage of proximity assessments, estimates **school marks** can be made (if necessary), as shown in Fig. 8.

In Fig 8 **the red column** on the left shows the range of F0 change of the reference phrase, and **the brown** one – the pronounced phrase. Below below the graphs are the minimum and maximum F0 values of the reference and spoken phrases.

To listen to the spoken phrases click the **“Play reference sample**” button and to the pronounced phrase click the **“Play recorded (test) file”** button (Fig. 8)**.**

In fig. 9 shows the results of intonation analysis of the same phrase when choosing the display mode: (**Use signal labeled - Manually**) and (**Show on screen - Source signals**).

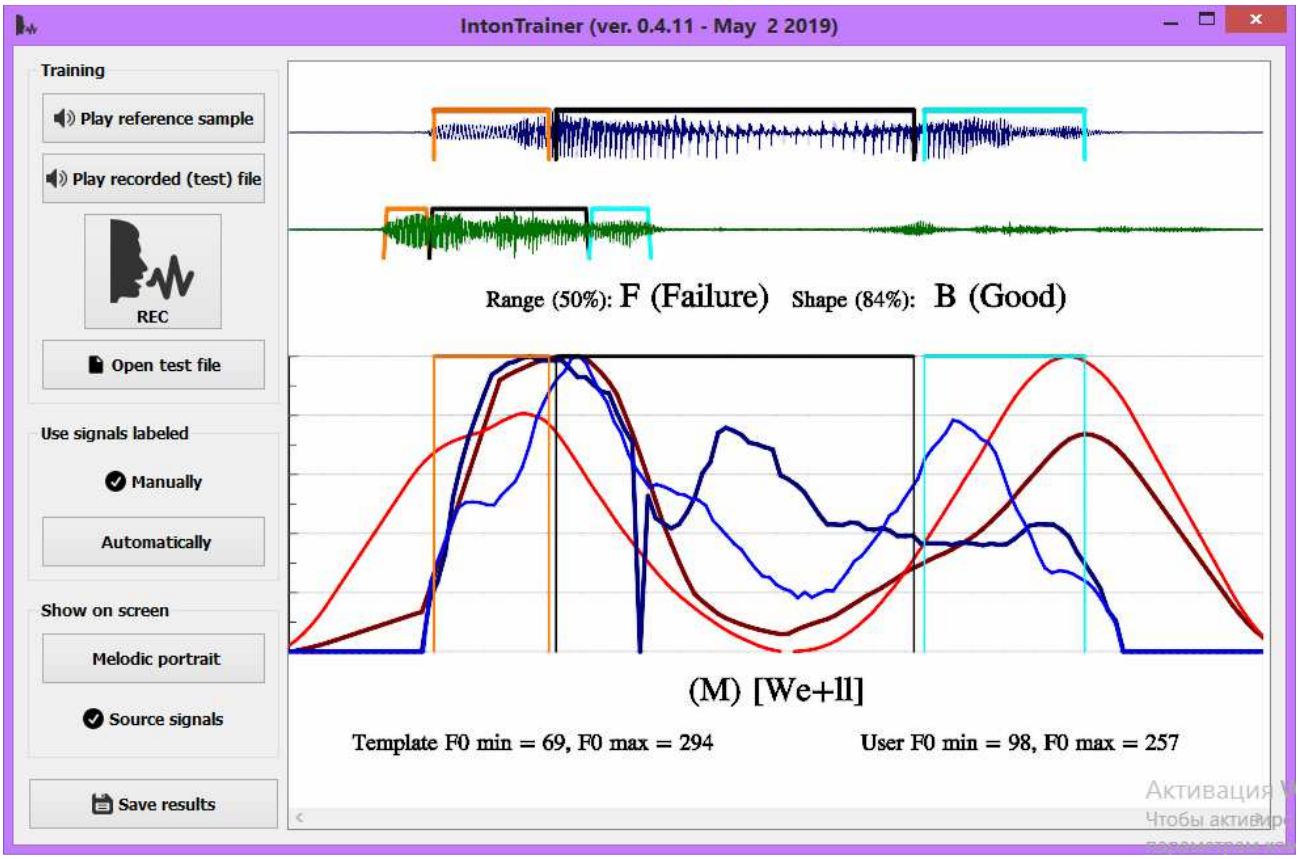

Fig. 8. The window displaying the results of analysis and comparison. Mode: "Manually" - " Source signals "

In the upper part of the window, the waveform of the phrase signal is displayed with indication of the pre-core marks (red line), the cores (black line) and behind-core marks (the blue line). In the middle part of the window, the curve of F0 in real time is shown in red. On this curve are indicated the sections of the pre-core, core and behind-core, from which the perpendicular magnetic field is formed, shown in fig. 7. The blue color represents the signal amplitude change curve - A0 (t).

The function “Source signals” is used to control in cases where there are doubts about the correctness of the manual markup for obtaining the UMP phrase.

Similar (but not identical) results (see fig. 7, 8) can also be obtained in the mode of automatic segmentation and marking of signals: (**Use signal labeled - Automatically**).

5. Comparison of the phrase intonation from various sources

As already mentioned above, this SWM can also be used as an instrument in a number of scientific and practical studies. For example, the SWM can be successfully used in experimental phonetic studies, during which it becomes necessary to compare the reference intonation with the intonation of the phrases studied from various sources (for example, when comparing the intonation of natural and synthesized speech). In this case, instead of using an external or built-in microphone, the "**Open Instance File**" button in the left section of the window is used. When using this button, a speech signal of the same content is selected from a specially created "TEST" folder, but obtained from another source, for example, for example, from a speech synthesizer, or a phrase spoken by another speaker or in another emotional state.

An example of the result of comparing the intonation of a reference phrase ***"How?"*** uttered by a female speaker (red line) and a male speaker (brown line) is shown in Fig. 8.

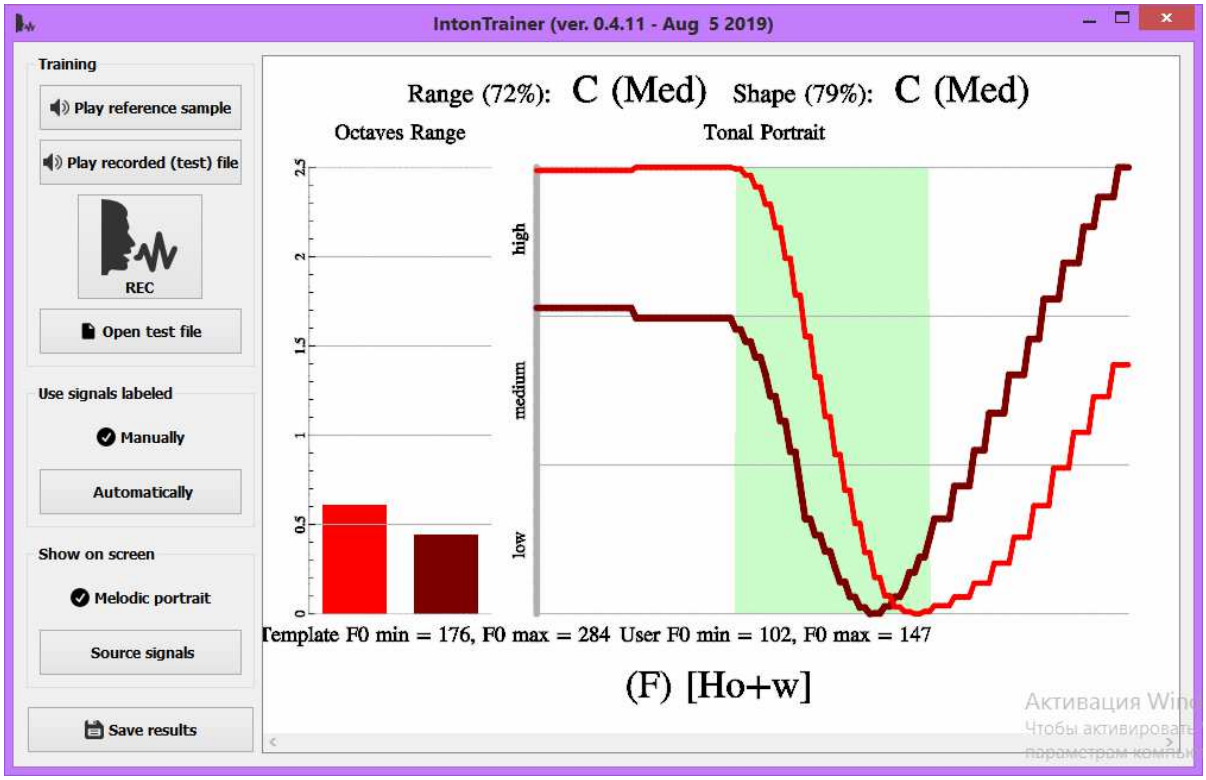

Fig. 8. The results of the UMP comparison of two speakers: female and male voices

In the upper part of the window the results of comparison of the natural and synthesized phrase are shown: **Range** (36%) - proximity to the range of changes F0 and **Shape** (88%) - proximity to the shape of the trajectory F0. In more detail, numerical similarity estimates are given in Fig. 8 on the left side of the window, as shown below:

6. Additional settings by using the button - " Acoustic Speech Database "

The software package **“Inton@Trainer”** is an open system which allows for various modifications. First of all, this refers to the used set of reference data (patterns), which can be supplemented or formed anew in accordance with the task in hand.

When you click the "**Acoustic Speech Database**" button (see fig. 3), a window (see Figure 9) opens, containing a structured list of reference phrases.

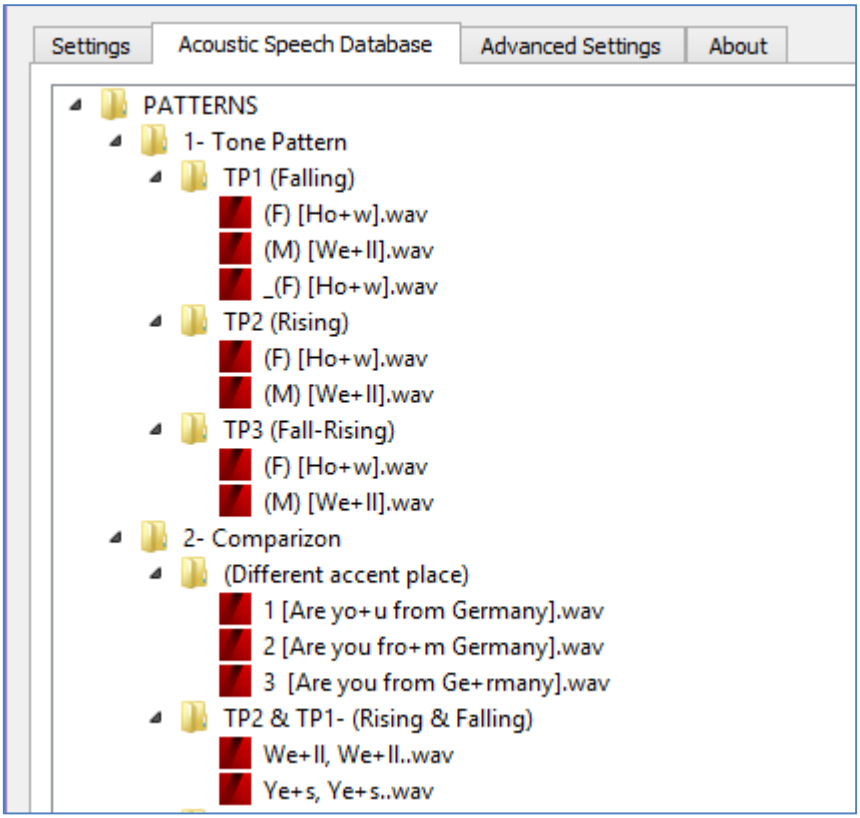

Fig. 9. Window displaying a structured list of reference phrases

Double-clicking on this list opens a file or explorer window. When you right-click a folder, the drop-down menu opens (see Figure 10):

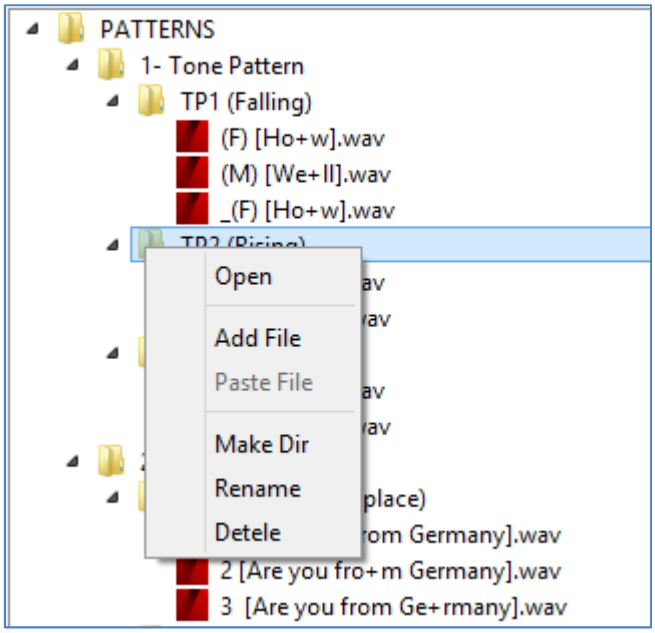

Fig. 10. Drop-down menu when right-clicking on the folder TP2

You can open the chosen folder (**Open**), add the file from the outside (**Add file**), create a new folder (**Make Dir**), rename (**Rename**) or delete (**Delete**) the folder.
When right-clicking the mouse on the file, a drop-down menu opens (see Figure 11):

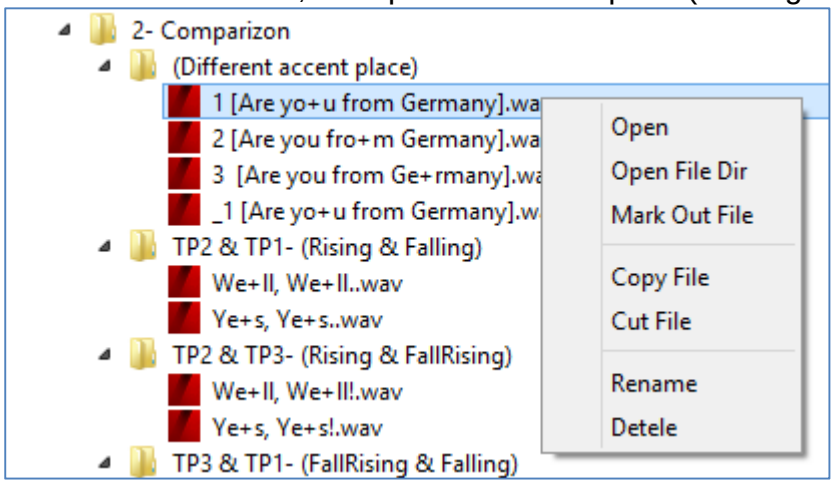

Fig. 11. Drop-down menu when right-clicking on the file: {1 [Are yo+u from Germany].wav}

In this window you can open the file (**Open**), or open the entire directory (**Open File Dir**), and you can also **copy, cut, rename** or **delete**. Thus, the set of reference data (acoustic database) used can be supplemented or formed anew in accordance with the task.

In addition, it is possible to automatically mark and design any file out the voice areas (**Mark Out File**). As a result, a new file with the same name is generated, but with the addition of the underscore to the beginning of the name**:** {_1 *[Are yo+u from Germany*].wav} (see Figure 11).

7. Prosodic marking of reference phrases of the acoustic database

An important factor in the formation of the acoustic database of the studied phrases is their prosodic marking by the areas (regions) of pre-nucleus, nucleus and post-nucleus. Currently this operation is performed manually using the standard SWM **"Sound Forge"**, but in the future it is supposed to be automated.

The speech signal of the phrase is recorded in a "wav" format with a sampling of 8 kHz, 16 bits and is labeled into regions P1 (pre-nucleus), N1 (nucleus), T1 (post nucleus) as shown in Fig. 12 for a single-nucleus (mono-accented) phrase: ***"I am saying till fi+ve.",*** pronounced by a male voice. The result of the construction of the UMP of this phrase, obtained on the basis of such a markup, is shown in Fig. 13.

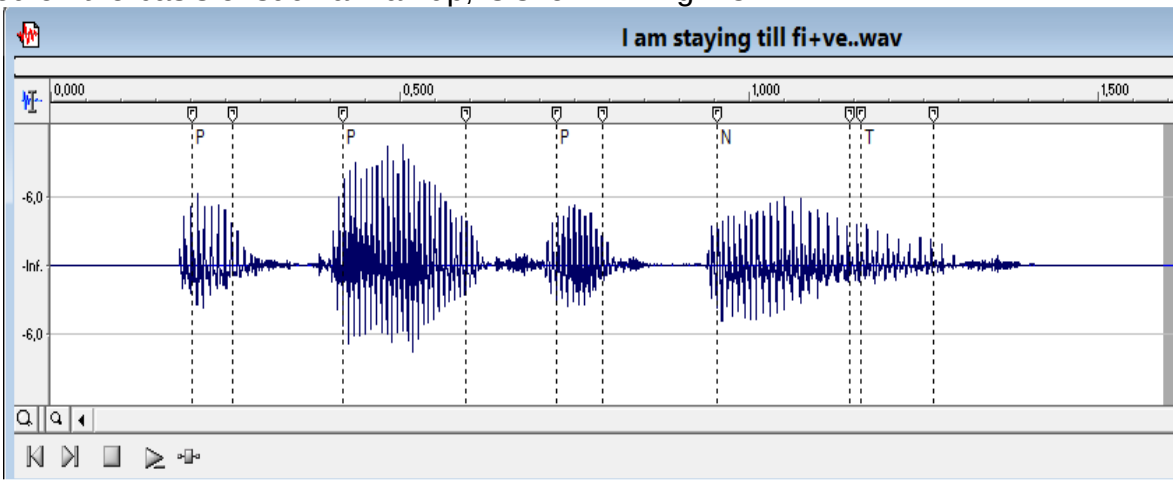

Fig. 12. Example of marking speech signal of mono-accented phrase "I am saying till fi+ve."

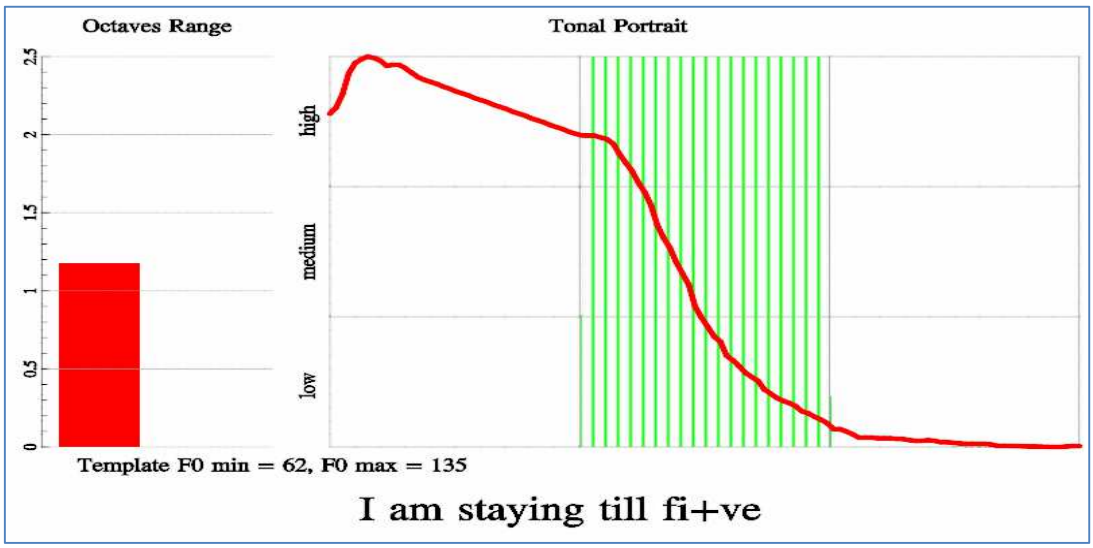

Fig. 13. The result of the construction of the UMP phrase: "I am saying till fi+ve."

In figure 14 an example of marking of two-accented phrase: *"**Befo+re you open the do+or, …"*** is shown. The result of the construction of the UMP of this phrase, obtained on the basis of such a markup, is shown in Fig. 15.

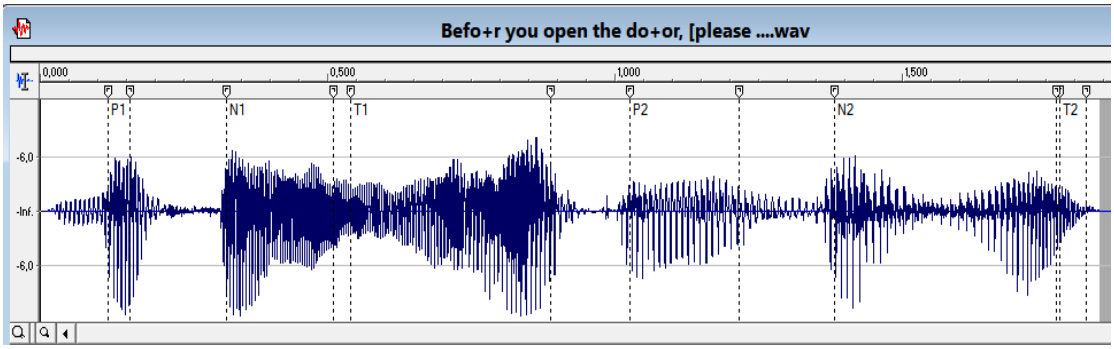

Fig. 14. Example of marking of two-accented phrase: "Befo+re you open the do+or, …"

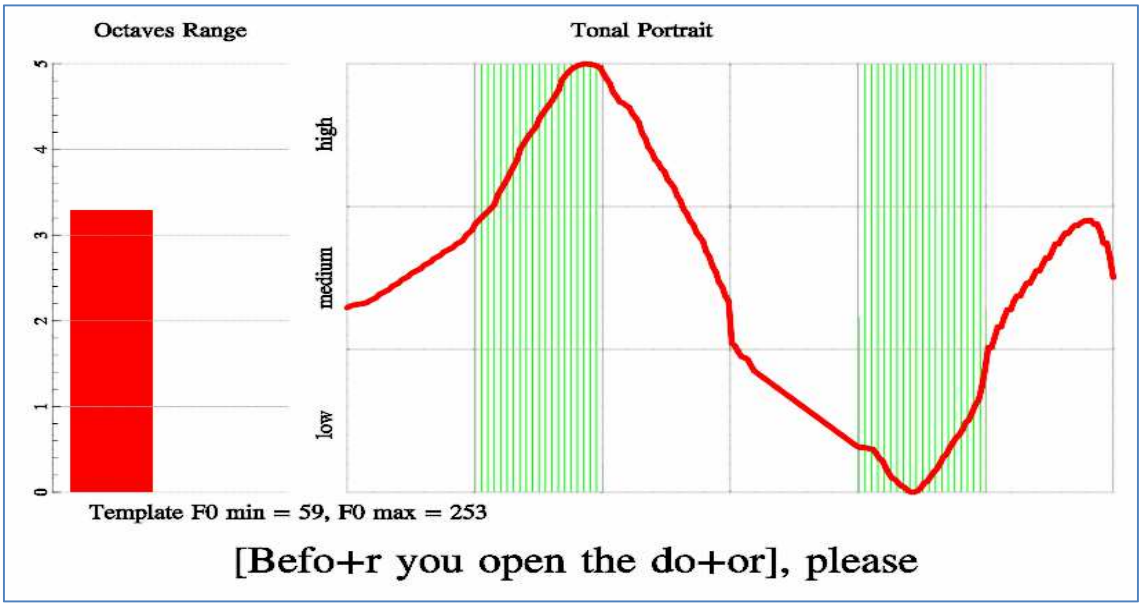

Fig. 15. The result of the construction of the UMP phrase: "Befo+re you open the do+or, …"

If, for some reason, it is difficult to determine the position of the elements of the accent structure of the phrase (P, N, T), then all the voice regions of the phrase can be assigned one the same index N. In this case, it is considered that each of the voice regions represent as nuclear. Fig. 16a (above) shows an example of manual marking waveforms (using the Sound Forge program) of the one-accent phrase "*Are yo+u from Germany*?" on (P, N, T) are regions. Fig 16b (below) shows an example of automatic waveform mapping, when each of the voice sections of a phrase is assigned the same index N. The way of creating such a file (as already mentions in section 6) is performed by selecting the operation (**Mark Out File**) in the **Advanced Settings** section - "**Acoustic Speech Database**" (see Fig. 11).

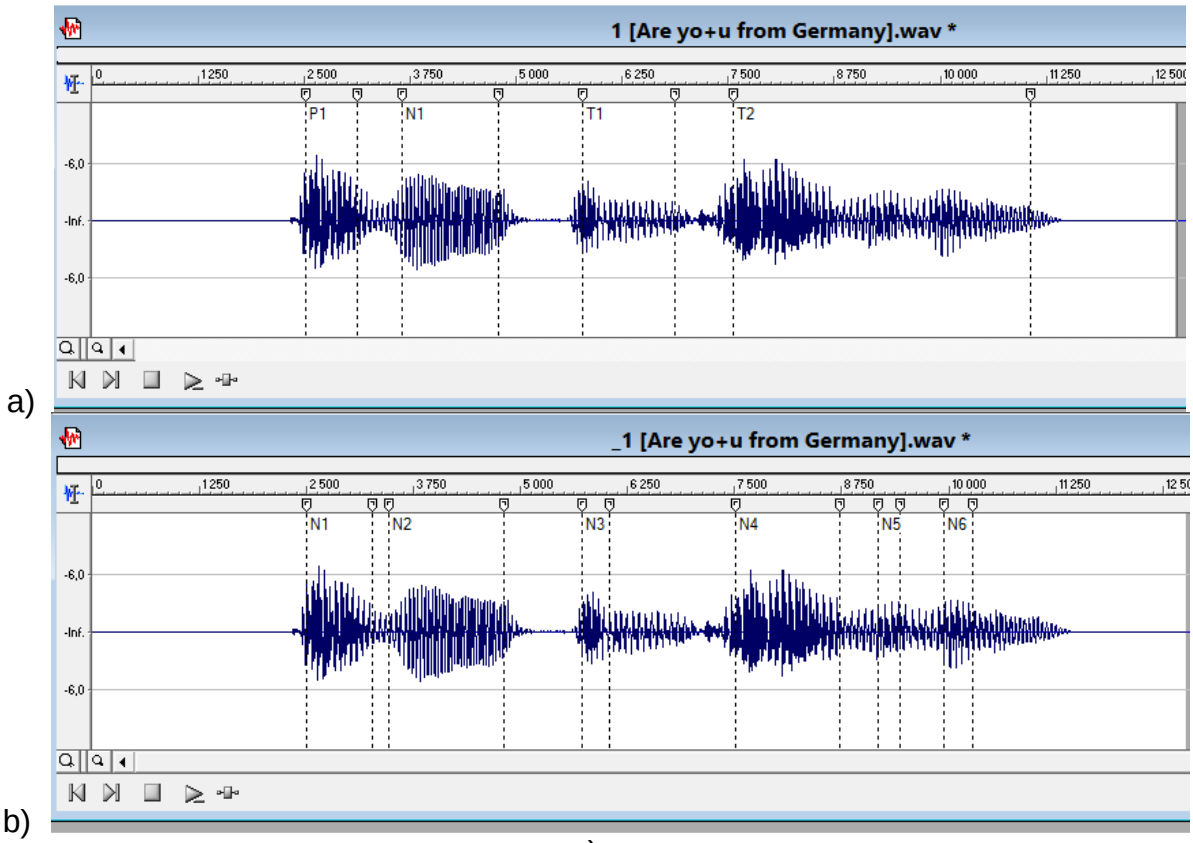

Fig. 16. Examples of manual marking of one-accent phrase (above) and its automatic marking (below)

Fig. 17 shows the UMP of the phrase "***Are yo+u from Germany?***" built in the presence of markup on P, N, T - regions, and in Fig. 18 shows the trajectory of F0 in the case when each of the voice regions of the phrase is assigned the same index N.

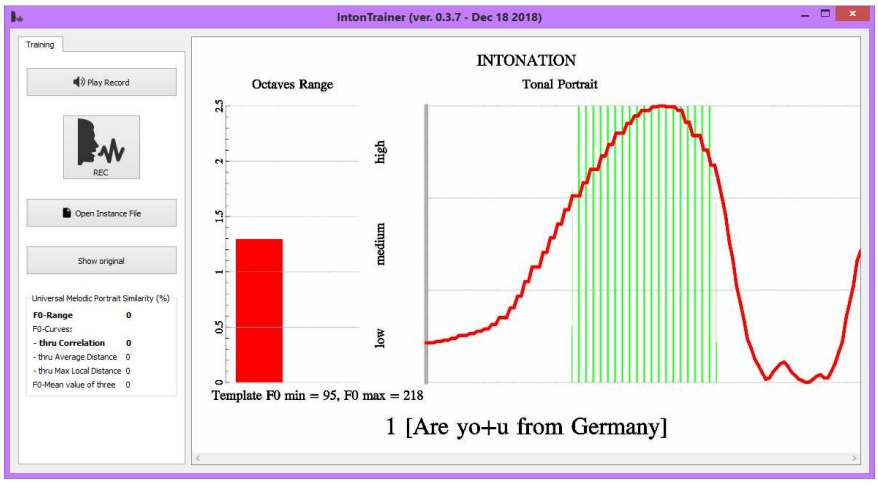

Fig. 17. An example of displaying the trajectory F0 (UMP) by manual marking on the P, N, T - regions

8. New additional functionality of the SWM

8.1. Extended parametric display of melodic intonation portraits

The standard graphic of the UMP presented earlier in Fig. 4, obtained when selecting in the **Settings** section **Display** (see fig. 3) **Show UMP** and **Show F0.** In case if the option **Show derivative F0** is selected, then we will get a joint image of the UMP and its derivative (see fig. 19). In the Settings section of the Display, there is also another possibility to display the UMP and its derivative. With an additional selection of the **Show phase plane** mode, we obtain the mapping in the coordinates “F0 - dF0”, shown in fig. 20.

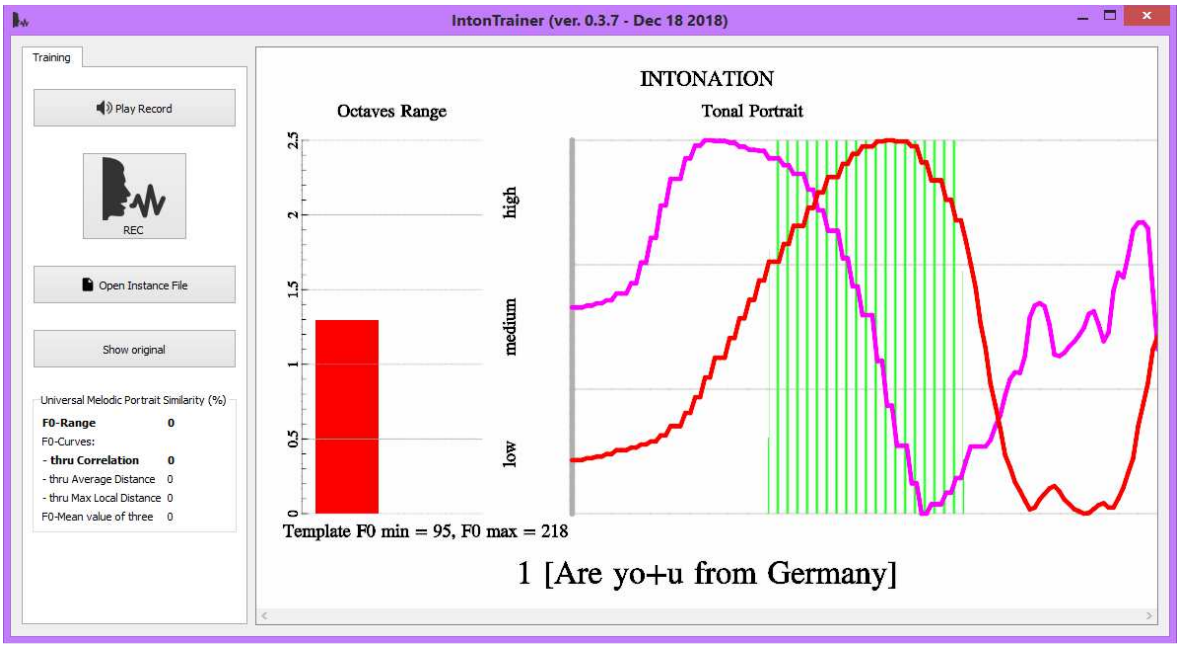

Fig. 19. Joint display of the UMP and its derivative

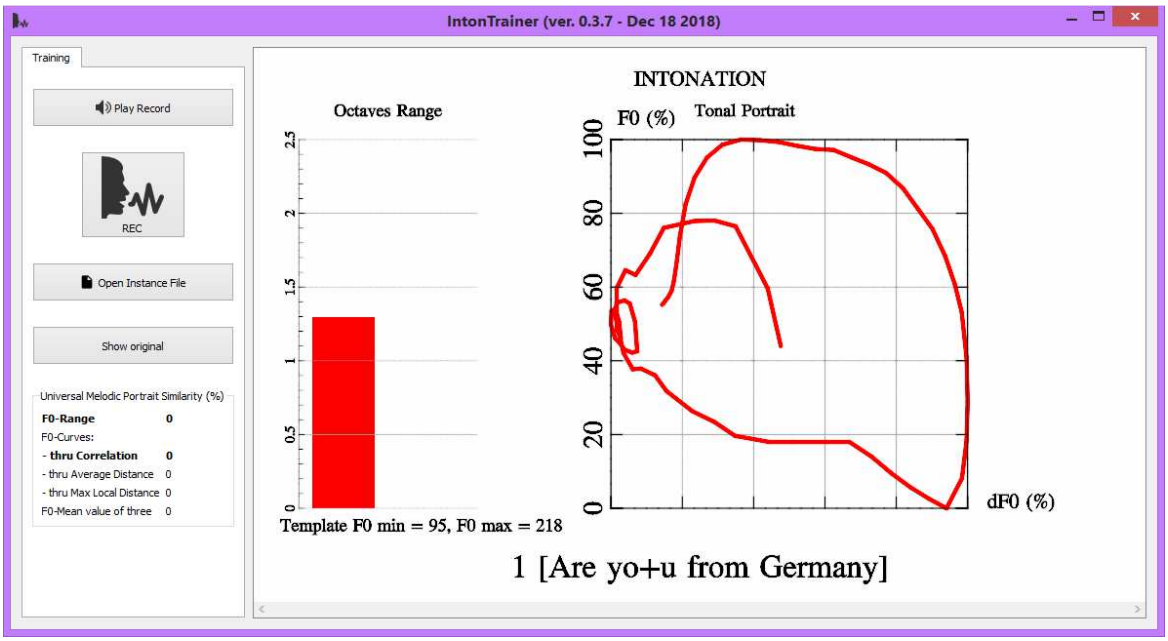

Fig. 20. Joint display of the UMP and its derivative in the coordinates “F0 - dF0”

In fig. 21, 22 shows the results of display, analysis and comparison of the reference and spoken phrases in the two modes described above.

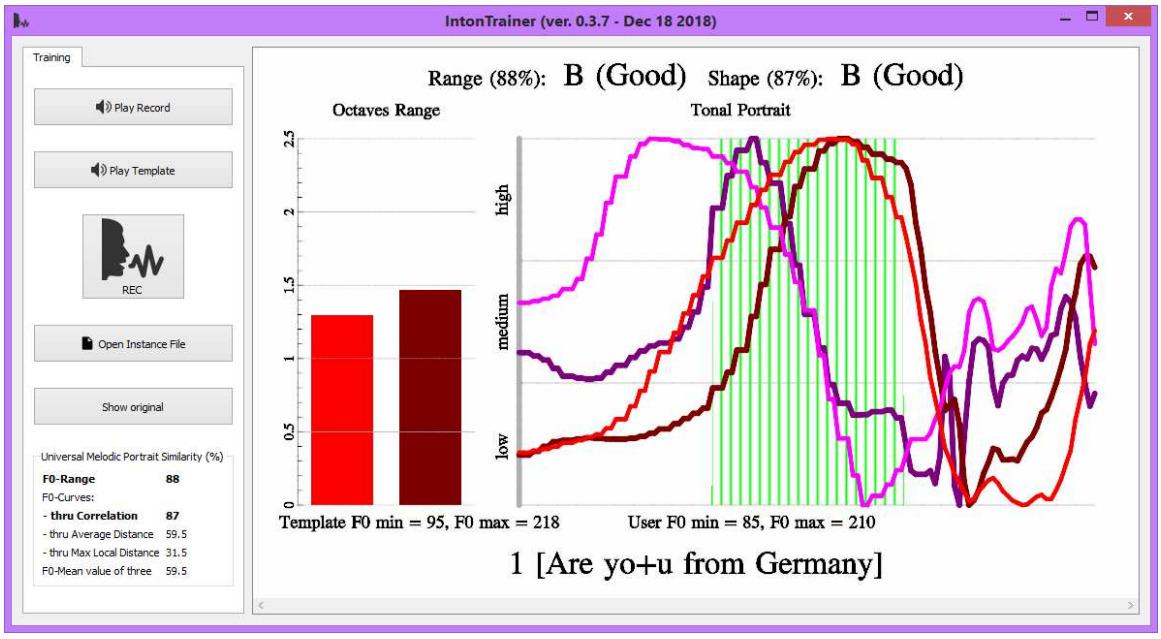

Fig. 21. Results of displaying and comparing reference and spoken phrases and their derivatives

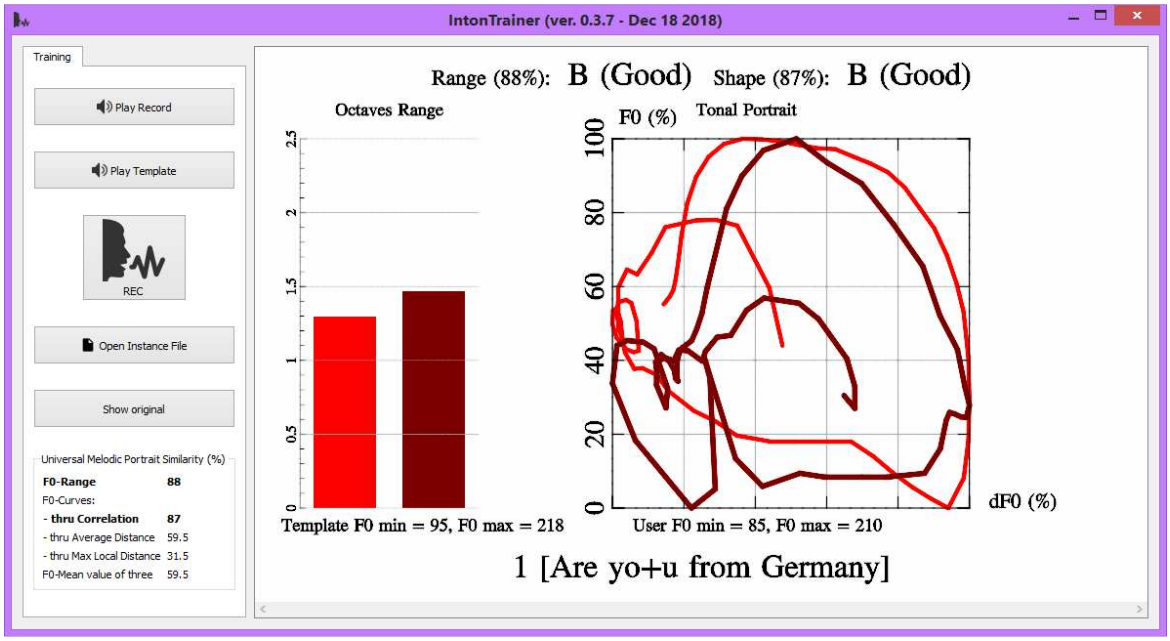

Fig. 22. Results of displaying and comparing reference and spoken phrases and their derivatives in the coordinates "F0 - dF0"

The extended parametric display of melodic portraits of intonation by the joint analysis of the UMP and its derivative (see Fig. 21) allows one to take into account more subtle differences in the intonation of the reference and spoken phrases. Adding the ability to display the UMP and its derivative in the coordinates “F0 - dF0” (see Fig. 22) made it possible to make a visual comparison of the intonation of the reference and the spoken phrases more vivid as compared with Fig. 21.

8.2. Automating of the procedure for marking the analyzed signals into voice regions

The **IntonTrainer** software package is an open system and the user is provided with ample opportunities for correction, updating and creating new reference databases for various languages and various practical SWMs. At the previous version there was the only possibility of preliminary automatic marking of reference sound files into voice regions, namely: by selecting the operation **Mark Out File** in the advanced settings section - **“Acoustic Speech Database**”.

In this version of **the SWM**, there is an additional possibility of direct automatic marking of reference signals. To do this it is necessary to select in the section **Autosegmentation** (see Fig. 3) the **Auto Marking** mode. Signal segmentation into voice regions is carried out on the basis of information about the presence of periodicity in the signal while the signal amplitude is present at sufficiently high amplitude - A0 (t).

For auto segmentation, the result of the current smoothing (integration) of the signal amplitude - **A0 (n)** is used. Smoothing is performed on the interval **(+, -) N**, where **N** is the number of samples in which the values of the **A0 (n)** curve are averaged at every nth time count. As a result of smoothing, the current values of the curve **Th rel(n)**, which sets the relative segmentation threshold, are formed. The value of N is selected by the user in the settings as the parameter **Th N** (see Fig. 3). In addition to the relative threshold, the user also sets the absolute segmentation threshold **Th abs**, Moreover, the common threshold **Th (n) = max [Th abs, Th rel (n)].**

The user in the **AutoSegmentation** section is given the opportunity to independently specify the absolute threshold values - **Th abs (%)** and relative - **Th N**.

A graphic representation of the actual shape of the A0 (t) signal and the results of the segmentation of the phrase signal into voice regions (see Fig. 23) can be obtained by pressing the **Show Original** button and setting the **Show A0** mode in the **Display** section (see Figure 3).

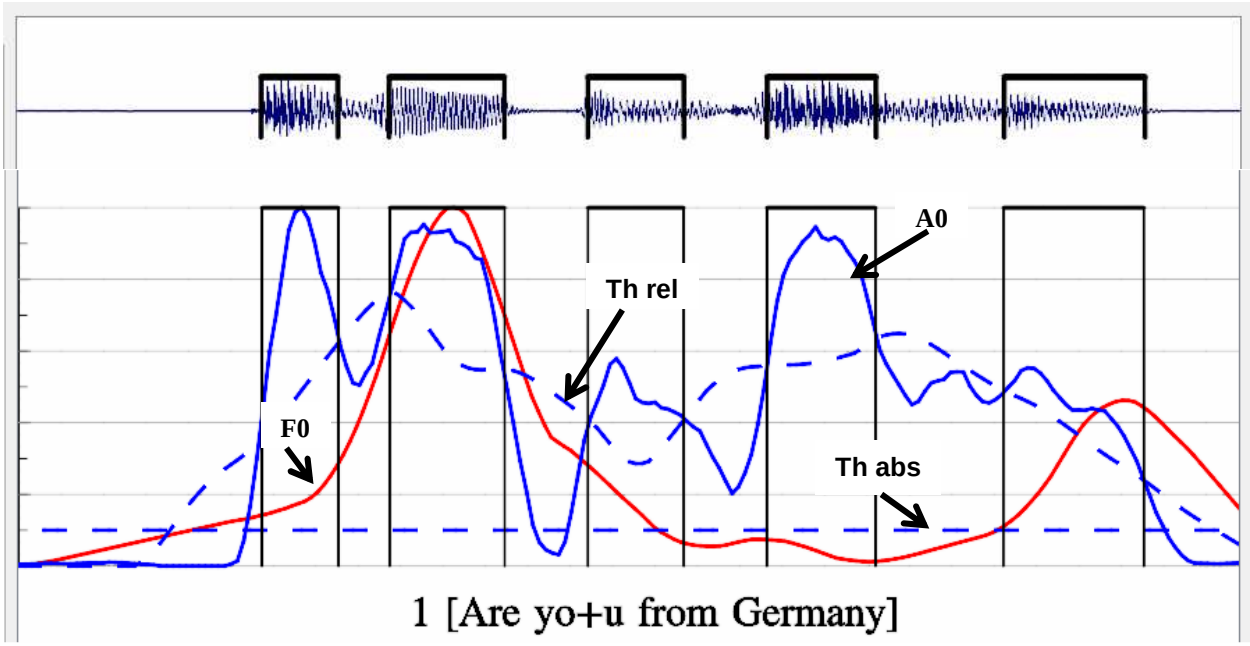

Fig. 23. Graphical representation of the results of segmentation of the signal

8.3. Calculation and visualization of prosodic signs of the analyzed phrases (On the example of emotional intonation)

The initial SWM window after downloading "**Emotions (English) IntonTrainer Demo"** is shown in fig. 24.

Fig. 24. Initial window

After clicking the **“Start”** button, the main window of the Program opens, containing a structured list of reference phrases indicating **the name of the announcer, the name of the emotion and the text of the phrase** in which it is reflected (Figure 25).

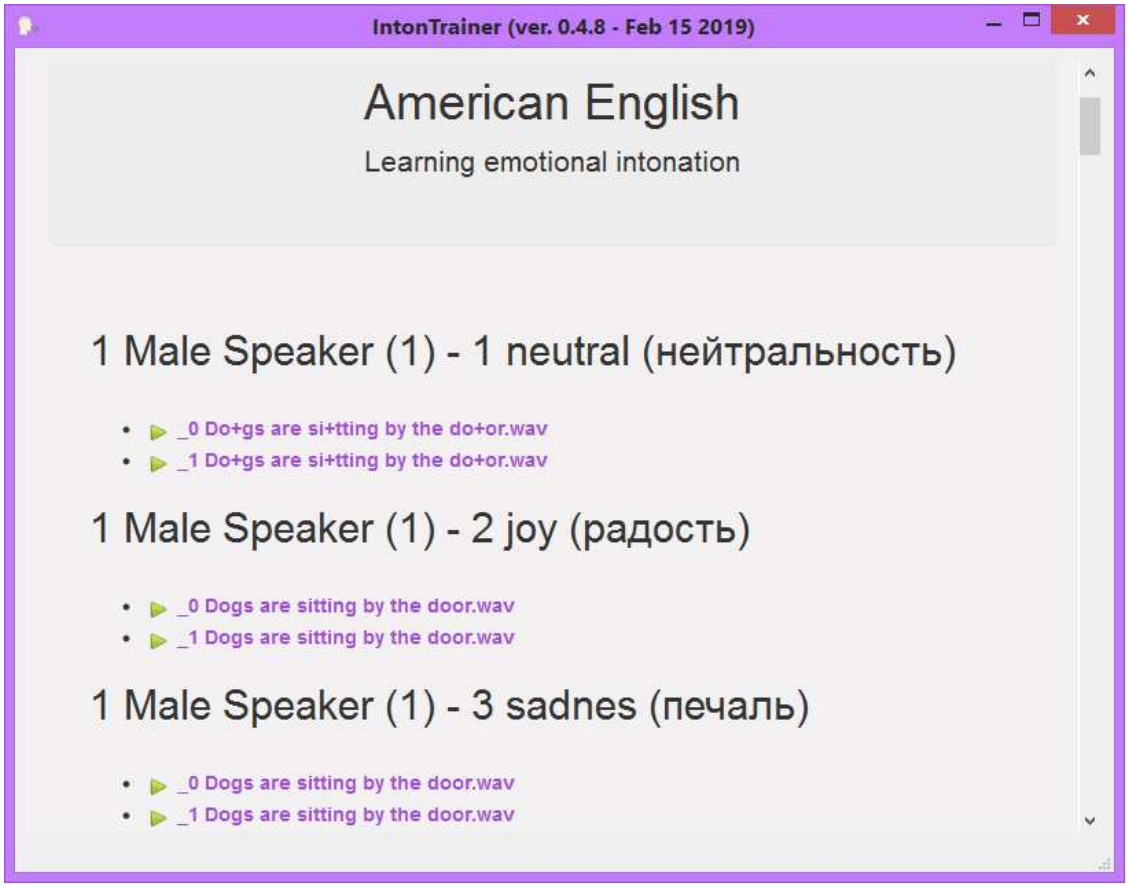

Fig. 25. Main window

By selecting the desired directory with the cursor, for example:

“1 Male Speaker (1) - 1 neutral _1 *Do+gs are si+tting by the do+or*”

a window (Fig. 26) opens in which the results of the intonation analysis of this phrase are graphically displayed.

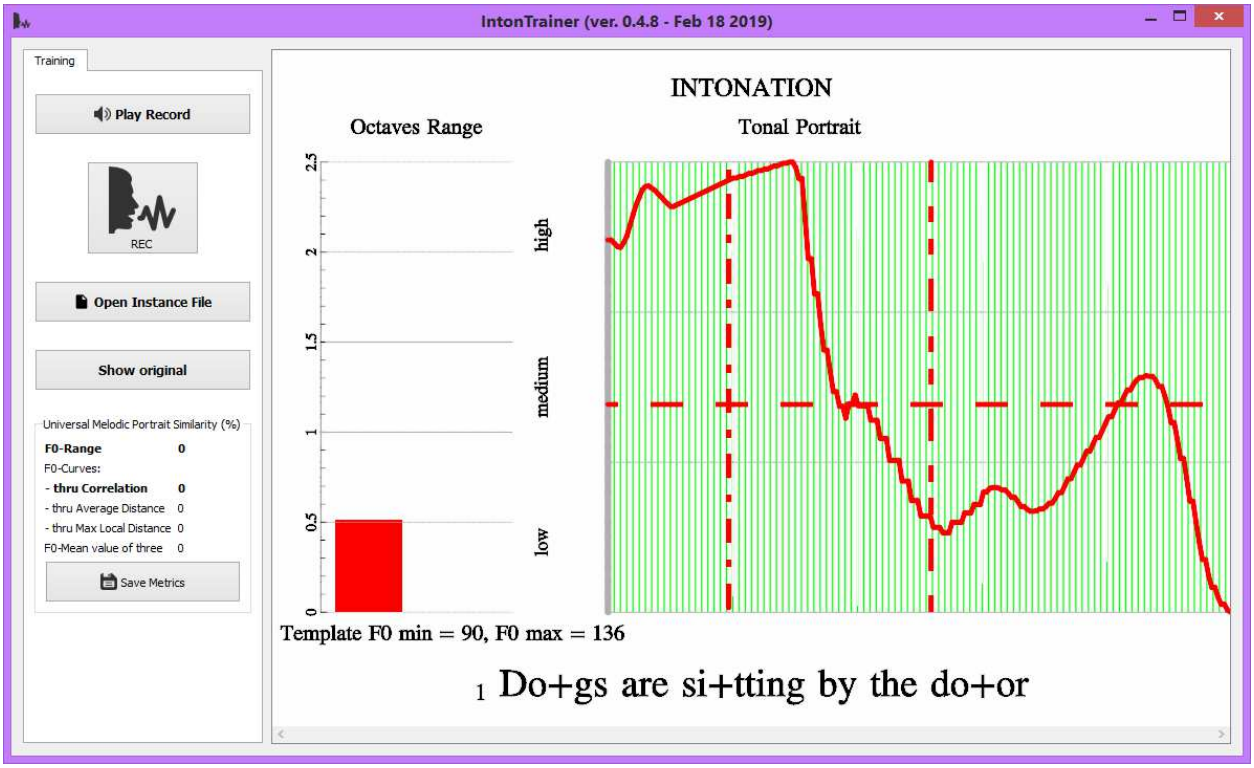

Fig. 26. Window displaying the NMP curve of the phrase "Do+gs are si+tting by the do+or"(Neutral emotion)

The continuous curve in fig. 26 displays the trajectory of F0 change on the voice sections of the phrase and is presented in the form of the Normalized Melodic Portrait (NMP). Segmentation for voice signal regions is carried out automatically (by selecting the section Setting Auto Markin mode). Segmentation is based on the information about the presence of periodicity in the signal (voice), while the presence of a sufficiently high signal amplitude - A0 (t).The construction of the NMP curve, in contrast to the Universal Melodic Portrait of the UMP, does not require "manual" marking of the phrase into sections of the pre-core, core and post-core.

The horizontal dashed line on the NMP curve shows its average value. Two vertical lines characterize the position of the center of the NMP curve and its width on the normalized time axis. The height of the column (to the left of the NMP) shows the range of variation of the F0 in octaves.

In the left part in fig. 26 shows control buttons with which the following functions are available:

- **"Play Template**" - listening to reference phrases.
- **"Rec"** - quick recording of user phrases through a microphone,
- **"Open Instance File"** - call test phrases from the "TEST" folder,
- **"Show original"** - view source signals,
- **Save "Save Metrics"** - saving data on measured prosodic signs.

When you click the **"Save Metrics**" 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.

Table 1. ***Prosodic features of the phrase*** *"Do+gs are si+tting by the do+or"* ***(Neutral emotion)***

i№	Names of Prosodic Features	Results
1	Diapason F0	1,51
2	Register F0 [Hz]	113,00
3	Mean Value of the curve NMP	46,15
4	Center of the curve NMP	35,65
5	Width of the curve NMP	32,50
6	Mean Value of the Derivative curve NMP	64,05
7	Center of the Derivative curve NMP	45,49
8	Width of the Derivative curve NMP	53,75
9	Voiced Sounds Level	0,19
10	Voiced Sounds Duration	214,00

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. A comparative view of these curves is presented in Fig. 27. 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.

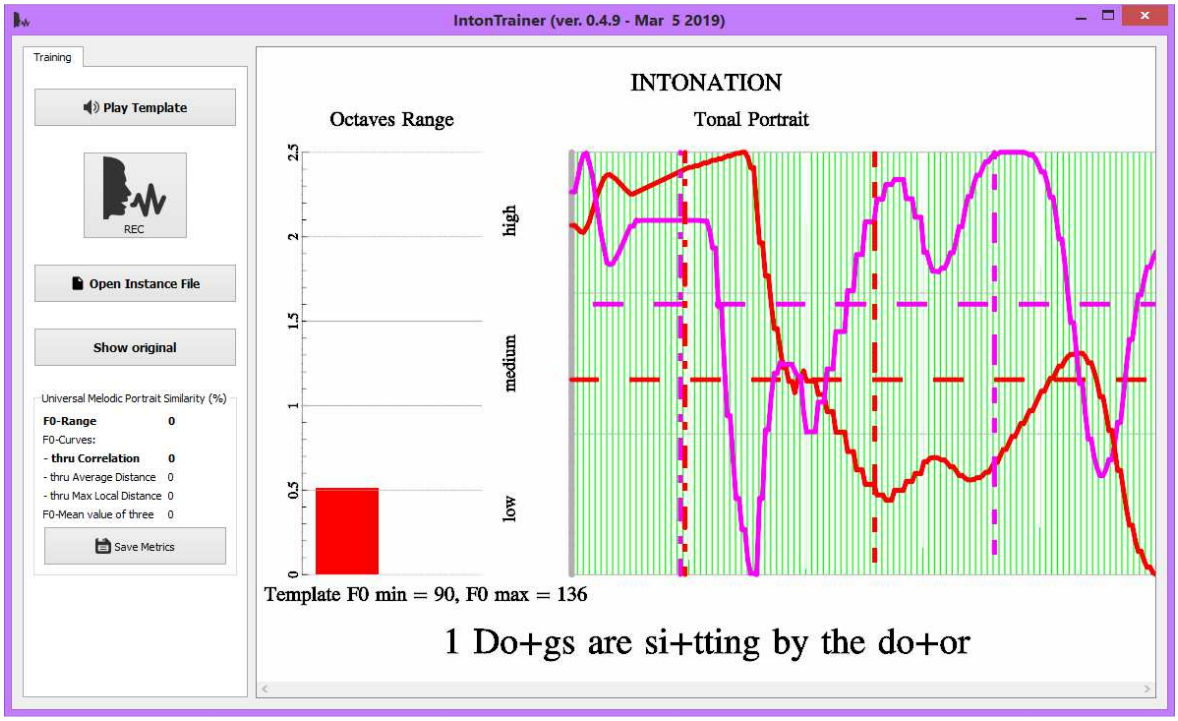

Fig. 27. Example of displaying NMP and d / dt (NMP) curves: *"Do+gs are si+tting by the do+or"* ***(Neutral emotion)***

After clicking the "Save Metrics" button, an additional icon appears and when it is pressed, a page opens in EXCEL, on which a complete set of prosodic features of the reference and compared phrases is written (see Tables 2, 3, 4).

Table 2 shows an example of the results of calculating the numerical values of the prosodic signs of a phrase pronounced expressing the emotion “Anger”.

Table 2. ***Prosodic features of the phrase “****Do+gs are si+tting by the do+or****”***
(**Anger** ***emotion)***

i**№**	**Names of Prosodic Features**	**Results**
1	Diapason F0	1,83
2	Register F0 [Hz]	243,00
3	Mean Value of the curve NMP	40,59
4	Center of the curve NMP	37,66
5	Width of the curve NMP	29,39
6	Mean Value of the Derivative curve NMP	59,82
7	Center of the Derivative curve NMP	47,70
8	Width of the Derivative curve NMP	53,09
9	Voiced Sounds Level	0,32
10	Voiced Sounds Duration	294,00

Table 3 shows an example of the results of a calculation based on the data given in Tables 1 and 2 of the relative values for the prosodic features of a pair of phrases with “Anger / Neutrality” emotions expressed in decibels. The use of relative values allows the comparison of a pair of phrases with different emotions, using prosodic signs of different nature with evaluation in various units of measurement.

Table 3. ***Relative values for the prosodic features of a pair of phrases with “Anger / Neutrality” emotions)***

i**№**	**Names of Prosodic Features**	**Results**
1	Diapason F0	0,82
2	Register F0 [Hz]	3,33
3	Mean Value of the curve NMP	- 0,56
4	Center of the curve NMP	0,24
5	Width of the curve NMP	- 0,44
6	Mean Value of the Derivative curve NMP	- 0,30
7	Center of the Derivative curve NMP	0,21
8	Width of the Derivative curve NMP	- 0,05
9	Voiced Sounds Level	2,36
10	Voiced Sounds Duration	1,38

Table 4 also shows the results of calculating numerical measures of similarity and distances between two of emotional phrases presented (in this case, a pair: (Neutrality - Anger).

Table 4. ***Results of calculating measures of similarity and distance (Anger / Neutrality)***

№	Type of the proximity	Proximity	Distance
1	Proximity thru curve “Correlation”	53	47
2	Proximity thru curve “Integral”	61	39
3	Proximity thru curve “Local”	30	70
4	Average of the three above proximities	48	52
5	Proximity thru “Range”	62	38

8.4. Averaged assessment of the similarity of pronounced phrases to the reference phrase

In the mode of intonation training, the ability to set the desired number of test files (pronunciations) and to obtain an average score of their similarity to the reference phrase is added. To do this, in the initial settings window - **Settings** - in the section - **Test files limit** - the desired maximum number of test files is selected, for which an average score of their similarity with the reference phrase will be calculated (see Fig. 28).

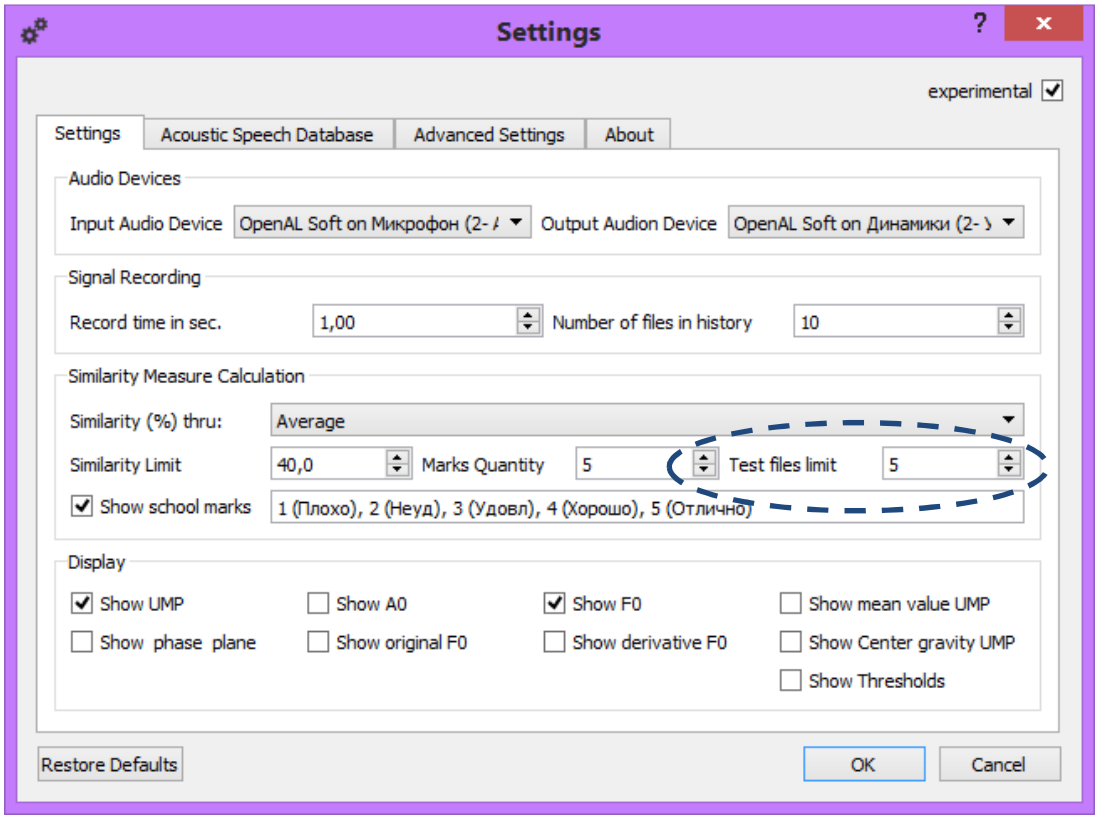

Fig. 28 Setting of “Test files limit”

Figure 29 shows the results of calculating the averaged estimates of the similarity of the 5 pronunciations of the phrase **“Well”** with the reference.

The left part of the figure (Octaves Range) graphically shows the results of calculating the similarities in the pitch frequency range for each pronunciation (thin columns, moreover, the last pronunciation corresponds to black) and the resulting averaged score (thick column). Above shows the numerical value of the average score - Range = 73% ..

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 (81%).

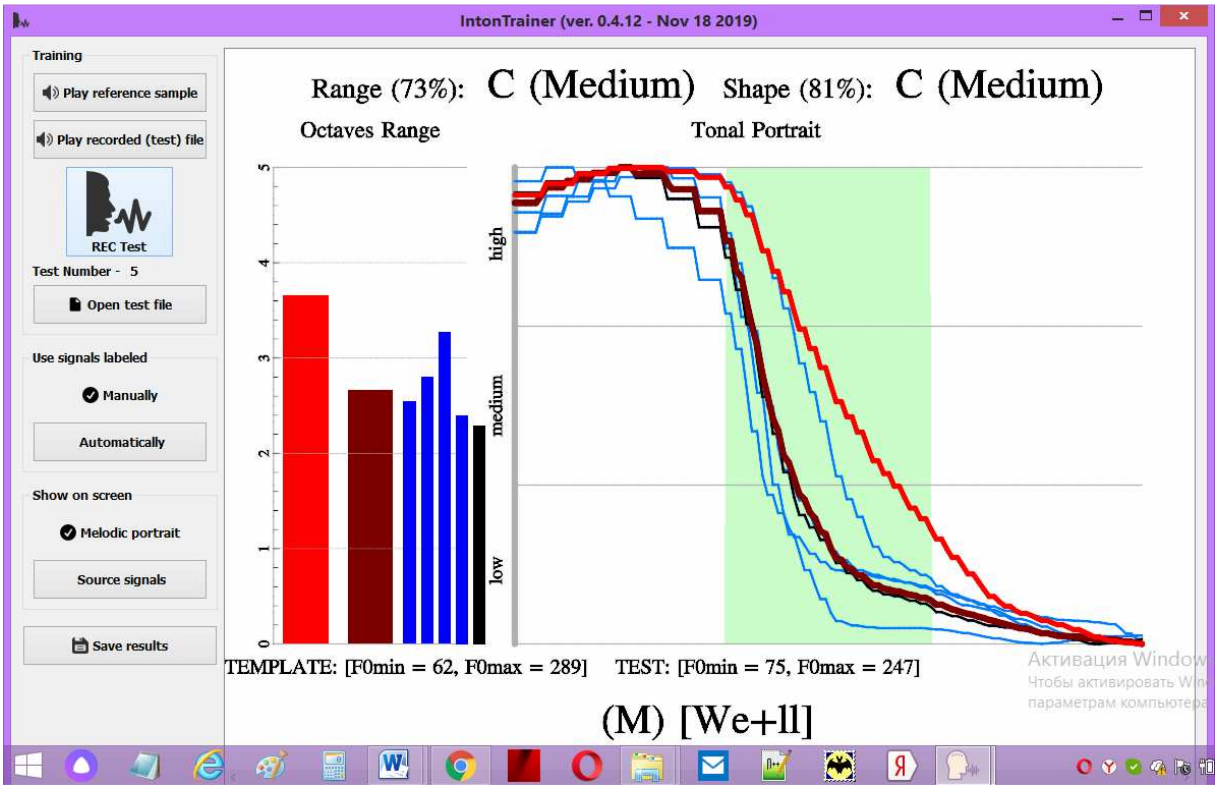

Fig. 27. Calculation results of averaged similarity ratings of 5 pronunciations

8.5. Service Software Module - Multi-Lingual Launcher

At present, 10 versions of **SWM** are loaded on the site **intontrainer.by**, which differ in language or function. The service software module - **Multi-Lingual Launcher** - was created to ensure the convenience of the user with the set of IntonTrainer SWMs he needs. To implement this feature, the user just needs to copy the set of SWMs he needs to the Applications folder.

Figure 28 show the view of the starting window of the **Multi-Lingual Launcher** program in the case when the user would need a full set of SWMs. Using this service program, the user has an ability to promptly call one or more SWMs of interest to it.

Fig. 28. The starting window of the Multi-Lingual Launcher

9. CONCLUDING REMARKS

We recommend that users from time to time look at our site, which can constantly appear useful updates to the program "IntonTrainer".

To learn more about the theoretical basis of the SWM, see the additional information posted on the site.

P.S. Sorry for some possible inaccuracies in the English text.

(Correct description of the SWM “IntonTrainer” see Russian version of the “User guide”)