

=VoiceEmoMeter= (VEM)

VEM version from 2023-07-18

USER GUIDE

1. Functionality

The **VoiceEmoMeter (VEM)** software is designed to measure **the degree of emotionality** of a person's voice on a scale from 0° to 100°, divided by the degree of emotionality into three conditional zones:

- Low degree from 0° to 30° - "*You are calm, relaxed and in control of the situation*" (reflections, memories, oral reading etc.);
- Average degree from 30° to 70° - "*You are active and confidently control yourself*" (dialogue, speech, lecture, etc.);
- High degree from 70° to 100° - "*You are agitated and unable to control the situation.*" (anger, hysteria, aggression, etc.).

VEM is adapted to analyze the emotionality of male and female voices. Measurement of the degree of emotionality of the voice can be carried out both directly from the microphone by the user, and from pre-recorded audio files from various sources.

VEM is implemented as a standalone application for Windows (7, 8, 10).

2. Scope of application

VEM can be used to control the degree of voice emotionality in medicine, psychiatry, education, telephony, radio and TV broadcasting, in teaching, medicine, psychiatry, as well as in other public spheres of life.

VEM can be useful in everyday family and social life to monitor the emotional atmosphere at home or at work.

VEM can become a personal tool for people of various professions seeking to improve EQ of their so called "emotional Intelligence".

3. Measurement method

To measure the degree of emotionality of the voice, certain parameters of the histogram of the distribution of the fundamental tone frequency (F0) of the analyzed

sound samples are used. Measurements are carried out separately in the ranges of F0 for male and female voices.

Figure 1a shows two F0 histograms of male voice samples in the range of 60–330 Hz: with a low degree of emotionality (blue curve) and with a high degree (brown curve), and 1b shows (in the range of 130–400 Hz) similar histograms of a female voice.

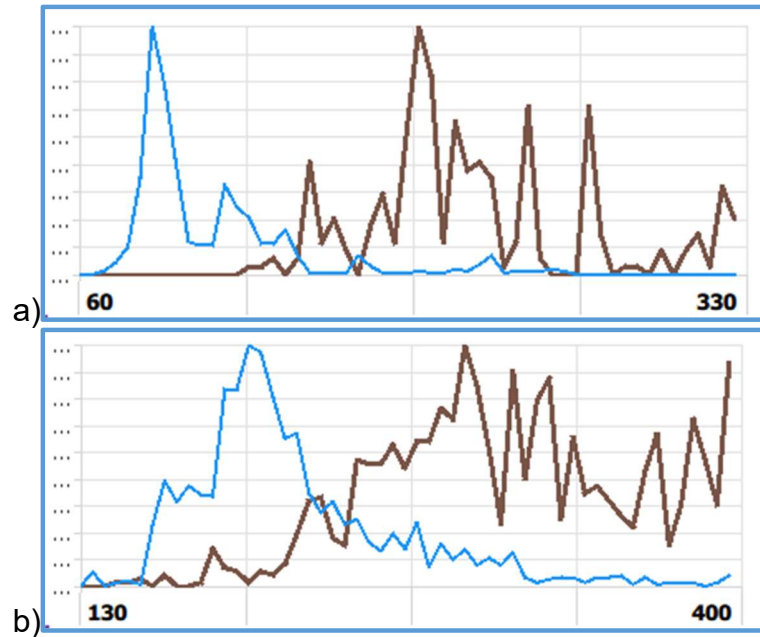


Fig. 1. F0 – histograms of male and female voices

Visual comparison of F0-histograms with low and high degree of emotionality presented in fig. 1a and 1b, clearly shows their significant difference, both for male and female voices.

The following two parameters are used to evaluate the integral differences between F0-histograms with low and high degree of emotionality:

- the center of "severity" of the histogram - C,
- effective "width" of the histogram – B.

The numerical values of both parameters vary from the smallest values - Cmin, Bmin - for a voice with a low degree of emotionality, to the largest values - Cmax, Bmax - for a voice with a high degree of emotionality.

The degree of emotionality E° of the voice is determined on a scale from 0° to 100° according to the formula:

$$E^\circ = \{[(C+B) - (C_{\min} + B_{\min})] / [(C_{\max} + B_{\max}) - (C_{\min} + B_{\min})]\} \times 100^\circ. \quad (1)$$

Here:

C, B - the values of these parameters obtained in the process of measuring the actual emotionality of the voice of a particular person in a particular situation.

Cmin, Bmin are the values of the parameters of the basic personal voice sample (with a low degree of emotionality) obtained at the preparatory stage (see paragraph 5 below).

Cmax, Bmax are the expected values of the parameters of an emotionally excited personal voice sample (with a high degree of emotionality). The calculation of the numerical values of these parameters is carried out according to the formulas:

$$C_{\max} = C_{\min} * dC, \quad (2)$$

$$B_{\max} = B_{\min} * dB.$$

Here, the factors dC and dB are determined on the basis of statistical studies. They are always greater than 1.0 and are presented in the settings section separately for male and female voices.

4. Preparation for E° measurement

The VEM start window that opens after starting the program is shown in Fig. 2.

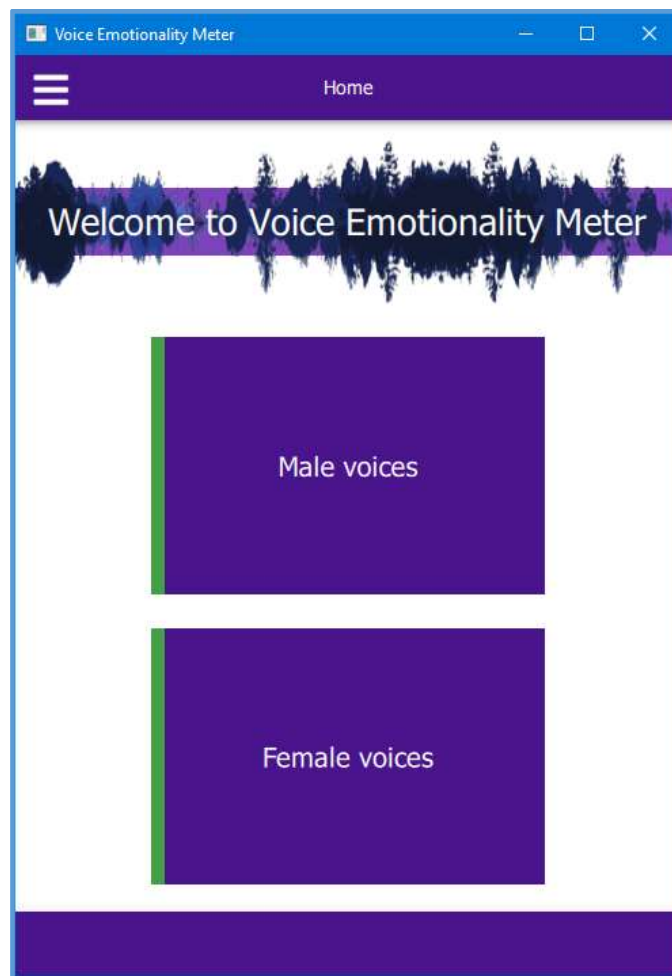


Fig. 2. VEM start window

After starting the program, the user is prompted to select the gender of the tested voice: male or female.

4.1. After specifying the Male voice, a preparatory window opens - Fig. 3a. A similar window opens when you select the Female voice option - Fig. 3b.

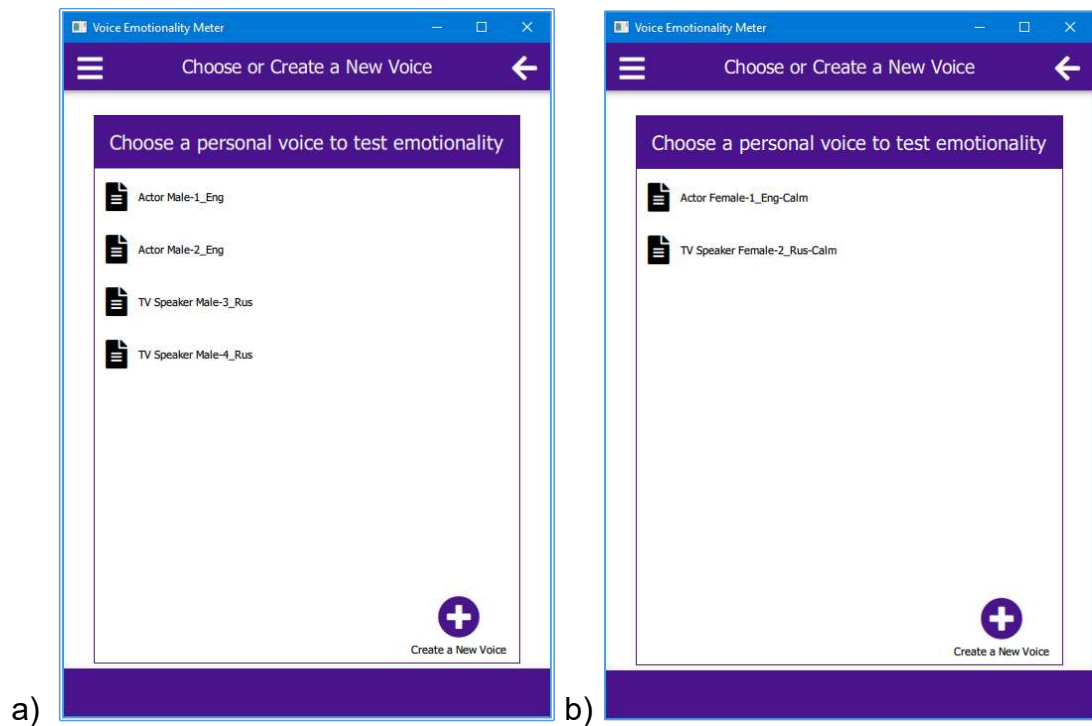


Fig.3. Preparatory window (a - male, b - female voice)

This window provides a choice of pre-created base voice samples, and also allows you to create new personal base samples. The preparatory mode of operation for the male and female voices is identical.

As examples, the *data/patterns/male* VEM folder contains samples of male voices: **Actor Male-1**, **Actor Male-2** from “*The Ryerson Audio-Visual Database of Emotional Speech and Song*”, as well as voice samples **TV Speaker Male-3**, **TV Speaker Male-4** from *Russian TV announcers* - (see Fig. 3a).

It is important to note that in order to measure the degree of emotionality of the voice of a particular person, the user must first create or borrow from other known sources his **basic personal voice sample with low degree of emotionality**.

4.2. To create a new emotionally neutral personal voice corresponding to a low degree of emotionality, use the icon (see Fig. 3).



When it is activated, an additional window opens (Fig. 4), which allows you to record a sound sample of a person's voice, his name and personal data, as well as save the sound in wav format.

Fig. 4. Additional window

4.3. Recording a sound sample of a personal voice is carried out through a microphone by clicking on the microphone icon:



Then, after the icon



appears, voice recording starts.

You must click this icon again to end the recording.

The executable voice fragment is recorded in the “*data/patterns/male(female)*” folder and displayed on the preparation window (see Fig. 5, **Boris_12-07-23**). It is then used as a base emotionally neutral sample when measuring how emotional this person's voice is in various emotional states.

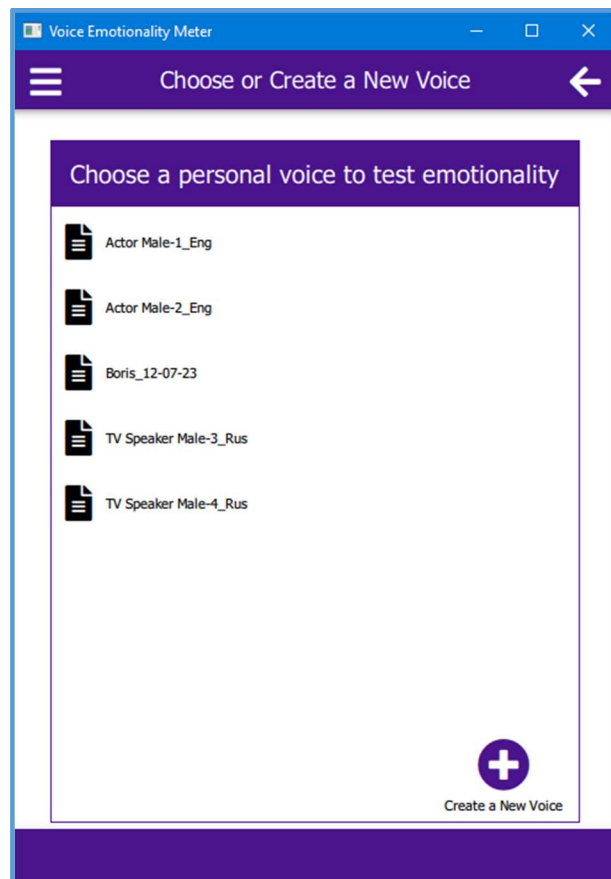


Fig. 5. Display the result of sample recording

The described procedure can be continued many times to record voice samples of other personalities.

The desired duration of the recording should be about 10 sec.

5. Operating mode of measuring E^o

After a set of personal sound samples has been created (see paragraphs 4.2, 4.3), you can switch to the E^o measurement operating mode.

When you select one of the samples (for example, the basic voice sample “**Actor Male-2_Eng**”, see Fig. 5), the initial working window opens (see Fig. 6) showing the histogram of the F0 distribution of the basic sample.

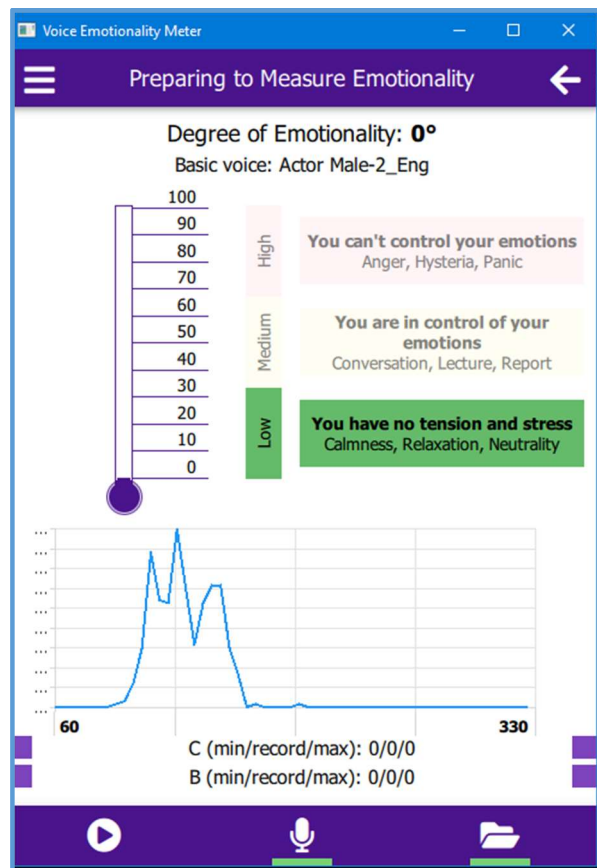



Fig. 6. Initial working window of VEM



Then, using the icon, , a sample of the same personal voice - "**Actor Male-2_Eng-Angry**" - is called from the "*data/tests*" folder, with a clearly expressed emotion of anger. As a result, the final working window VEM appears (see Fig. 7), displaying the results of measuring the degree of emotionality E_o of this voice.

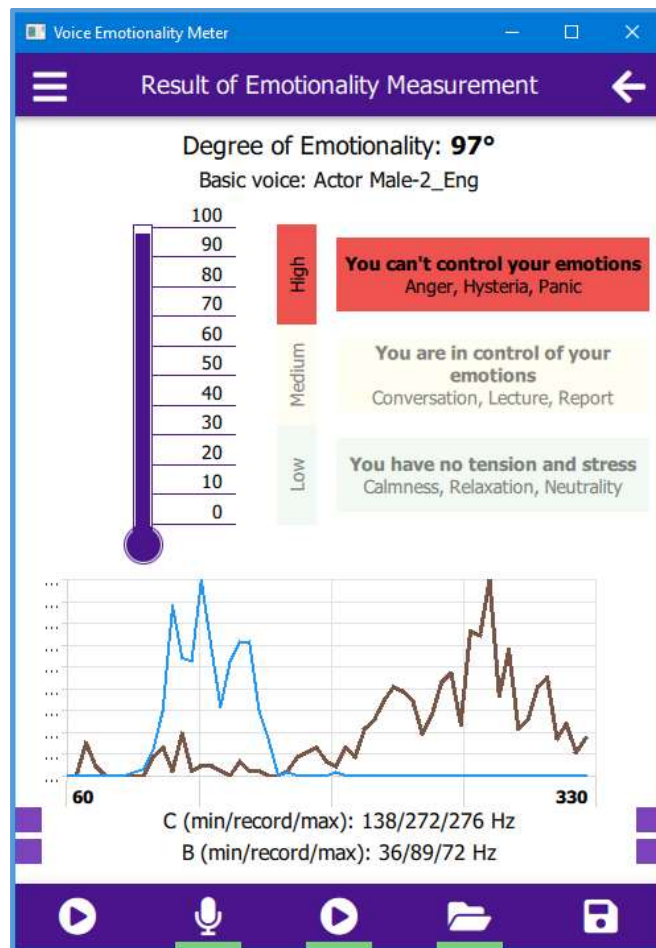


Fig. 7. Final working window (basic voice **Actor Male-2**)

In the lower part of the working window, the histogram of the distribution of F0 values of the basic voice with a neutral emotion is shown in blue, and the voices with the emotion of anger are shown in brown.

Below the images of the histograms are the corresponding values of the parameters C and B.

The upper part of the working window shows the numerical value of the measured degree of emotionality of the person's anger - $E^0 = 97^\circ$, and below the name of its base sample.

In the middle part of the window, a scale is visually presented in the form of a thermometer showing the measured degree of emotionality.

To the right of the scale are inscriptions explaining the situational conditions for the emergence of each of the three generalized degrees of emotionality - low, medium and high.

The procedure for measuring E_o for voice samples from "The Ryerson Audio-Visual Database", described above, remains the same for the case when the voice samples of a particular person are recorded through a microphone.

Figure 8 shows the result of measuring E_o during an expressive reading of an excerpt from A.S. Pushkin (basic voice Boris_12-7-23 with a neutral reading of the same passage).

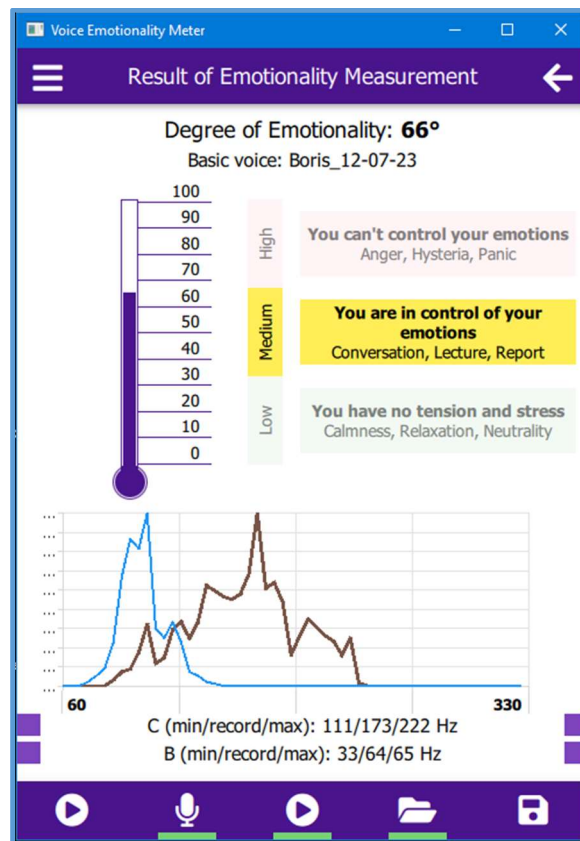







Fig. 8. Final working window (basic voice **Boris_12-7-23**)


6. Helper Icons

6.1. In the lower blue strip in Fig. 6 - 8 icon  is used to listen to the base voice, and the icon  in Fig. 7.8 - for listening to the voice, the degree of emotionality of which is measured.

6.2. In the lower blue bar of the working window in Fig. 7.8 is the microphone icon  used when recording a test audio file.

6.3. The icon  is used to call test files from the “data/tests” folder.

6.4. The icon  is used to save the test results.

6.5. In the upper blue bar of the working window on the left there is an icon , when activated, an additional information window opens (see Fig. 9).

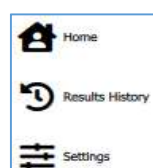


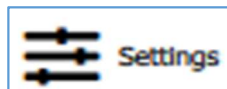
Fig. 9. Information window



6.6. When you select an icon the contents of the “results” file are called (see Figure 10).



Fig. 10. The contents of the “results” file



6.7. When you select an icon the settings window opens (Fig. 11), allowing you to set the dB and dC multipliers separately for female and male voices (see formula (2) in Section 3).

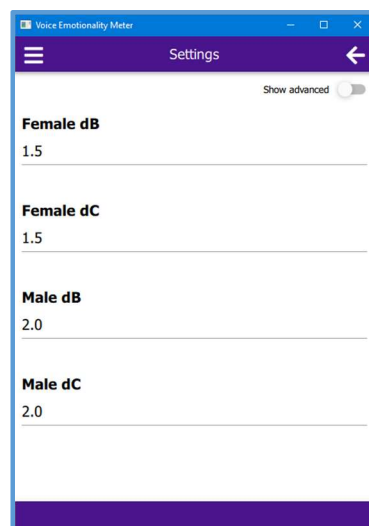
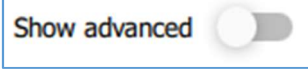


Fig. 11. Settings window - “Settings”

Figure 11 shows the default values of dB and dC multipliers, chosen experimentally by the developers.

ATTENTION!

Internal information triggered by a check mark in the upper right corner



,
is for the developer only.

Changing its contents by the user is highly undesirable!