SpeechRateMeter+ (SRM+)

SRM+ version from 2021-08-02

In the current version of SRM+ from 2021-08-02, unlike the previous SRM from 2021-04-05, a new important function is added to detect "FILLER" sounds. Its use makes it possible to assess the degree of "pollution" of the analyzed speech by sounds of this kind.

USER GUIDE

- ## **Functional purpose**

The SpeechRateMeter+ (SRM+) software is designed to assess the complex of prosodic characteristics of the speech tempo, namely:

1) Speech rate – Rs – total speed of reading the text (*calculated as the number of words per minute during the actual pronunciation time, i.e. taking into account the duration of interphrase pauses*);

2) Articulation rate – Ra – phonemic stream generation rate (*calculated as the number of words per minute for the total time of articulation, i.e. without taking into account the duration of interphrase pauses*);

3) Pauses score – Tp – the mean duration of the inter-phrase pauses (*calculated as the RMS value of the duration of all pause sections*);

4) Filler score – Tf – the relative duration of "filler" vocal insertions (*calculated as a ratio in pae cents of the RMS to the average-median duration of vowels*)

- ## **Areas of use**

SRM+ is implemented as a separate application for **Windows (7, 8, 10).**
The use of SRM+ can be useful in the following practical cases:
- When preparing speeches and presentations for self-control and self-study of the desired rate of speech;
- To improve the skills of oral reading in the native and foreign languages;;

- As a means of self-control to eliminate speech defects associated with various physiological disorders and diseases;
- For individual training of oral and speech skills in a number of professions such as: call center operators, radio, TV announcers, etc .;
- As a means of monitoring the rate of speech during the work of call-center operators, radio and television announcers in real conditions

- ## **Getting started**

The **SRM+ start window** that opens after starting the program is shown in Fig. 1.
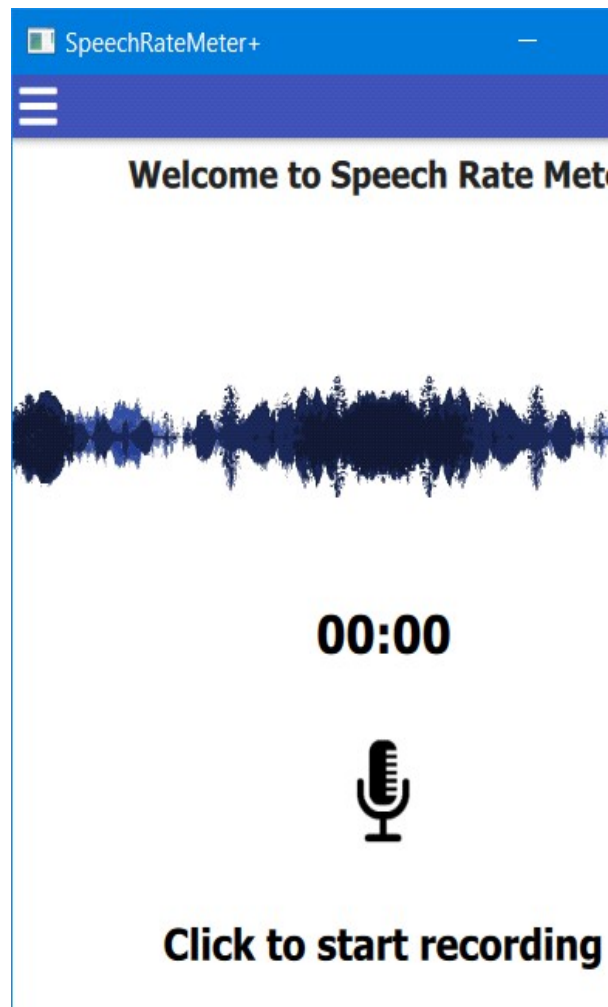

Figure: 1. SRM+ start window

After starting the program, the user is provided with the following options:

1) Recording through a microphone a sufficiently long segment of the analyzed speech (preferably at least 10 seconds);

2) Calling one of the previously prepared test audio files stored in the "data-tests" folder.

These capabilities provide numerical estimates of the four tempo characteristics of the analyzed speech indicated above.

The first option is implemented by clicking on the microphone icon:

and, after the entry icon appears:

recording of the spoken speech begins.

At the same time, a time counter is initiated showing the current duration of the spoken speech in seconds and minutes:

To end the recording, you must press the "Record" icon again.

The spoken speech signal is recorded in the **"data/records"** folder and prepared for further processing and display of the results.

The second option is implemented by clicking on the icon:

It is designed to call test audio files from the "data/tests" folder (see p. 5).

- ## **Graphic display of results**

After the analyzed speech signal is entered from the microphone, or from the **"data/tests"** folder, the parameters of the tempo of the analyzed speech are calculated and the **main SRM** window opens with the screen display of the calculation results (see Fig. 2).
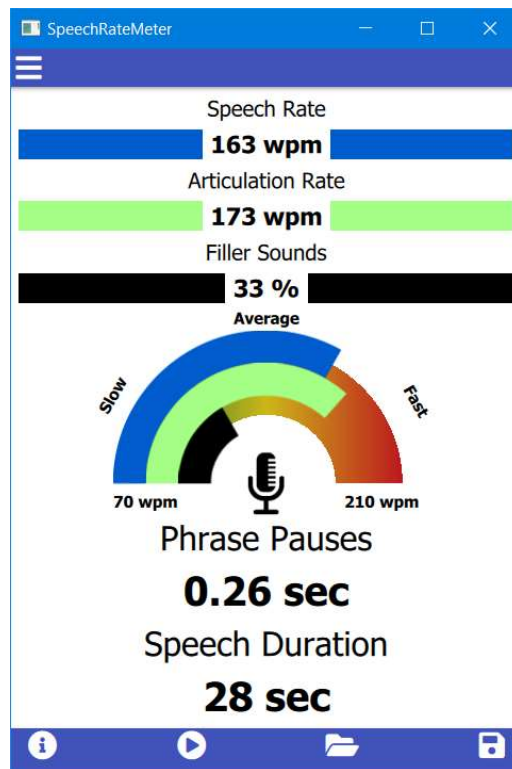
Fig. 2. Main window of SRM

The calculated numerical values of tempo parameters are displayed in the upper part of the main window:
- **Speech Rate (wpm),**
- **Articulation Rate (wpm),**
- **Filler score (%).**

In the lower part of the working window, the calculated numerical values of the estimate of the average duration of interphrase pauses and the duration of the analyzed speech segment are displayed:
- **Phrase Pauses score (sec),**
- **Speech Duration (sec).**

In the middle part of the window, the **Speech Rate (Rs), Articulation Rate (Ra)** and **Filler Sounds (Rf)** parameters are graphically shown in the form of two moving arcs of different colors: **blue – Rs, green – Ra and black - Rf**. The length of these arcs is normalized relative to the user-specified **minimum and maximum values** of these parameters.

The choice of the minimum and maximum values is determined by the specific conditions of using the SRM, under which the measured value of the speech rate will be considered: **Slow - Average - Fast**. The selected minimum and maximum **wpm values** are shown at the ends of the semicircle (see Fig. 2).
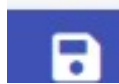
In the lower blue bar of the main window, besides the icon for calling test audio files from the "**data/tests**" folder, there are also two other icons.

Using the icon

you can listen to speech recorded through the microphone or recalled from the "**data/tests**" file.

Using the second icon

the results of measuring the parameters of the tempo of the analyzed speech are recorded in the file "**results**".

There is an icon in the upper blue bar of the main window on the left:

When you click on it, an additional **information window** opens (see Fig. 3),

Fig. 3. Information window

By choosing the icon:

the contents of the **file "results"** are called (see Fig. 4).

Star DateTime - 2021-8-1 9:2 Star DateTime - 2021-8-1 17
End DateTime - 2021-8-1 9:2 End DateTime - 2021-8-1 17:
Speech Rate - 127 wpm Speech Rate - 111 wpm
Articulation Rate - 133 wpm Articulation Rate - 133 wpm
Phrase Pause - 0.12 sec Phrase Pause - 0.34 sec
Speech Duration - 5 sec Speech Duration - 16 sec

Fig. 4. Content of the file "results"

By choosing the icon:

the window for setting the **minimum and maximum values of Rs (Ra) [wpm]** and the average duration of "filler" voice inserts – **Rf [%]** selected by the user is called (see Fig. 5).

Fig. 5. Window for setting the selected values: minRs, maxRa

4. Testual audio files

The user, based on his own scientific or practical interests, is given the opportunity to record an unlimited number of speech audio files from various sources. To do this, he needs to open the folder **"SpeechRateMeterBin (SRM)"** downloaded from the site, select the folder "**data**", and then the folder "**tests**". The contents of the "tests" folder can be updated and voiced using existing tools.

As examples, the "**tests**" folder currently contains 10 natural speech samples from various sources (see Fig. 6).
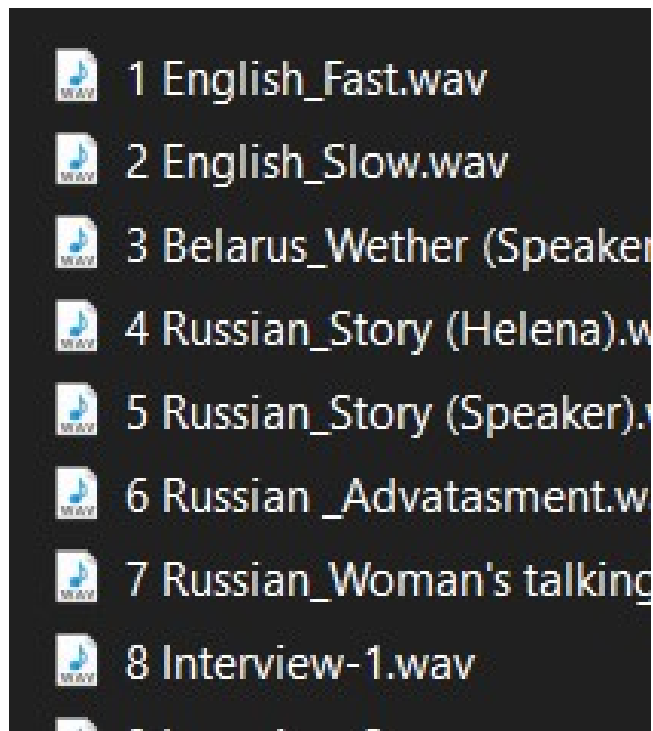
Fig. 6. Contents of the "tests" folder

The "tests" folder contains:

- **Examples (1, 2)** - samples of slow and fast speech in English;
- **Example (3)** - a sample of a weather broadcast in Belarusian;
- **Examples (4, 5)** - audio recordings of 2 identical fragments of the story by A.P. Chekhov's "Lady with the Dog" performed by a teacher of Russian and a professional actor;
- **Examples (6, 7)** - fragments of TV programs: advertising and a woman's monologue;
- **Examples (8, 9,10)** - excerpts from interviews with various persons in whose speech there are voice inserts - "fillers".

It should be borne in mind that the current version of SRM+ provides the ability to analyze the rate of speech only for signals in the **wav, mono, 8 kHz** format. All sorts of other formats should be adapted to this format using suitable software (eg Sound Forge).

CONCLUDING REMARKS

We encourage users to check our website from time to time, as there may be useful SRM updates on a regular basis.

Learn more about the theoretical foundations and programming code of SRM development see additional information posted on this site.