

# Visualization of a Person's Emotional Tension by a Speech Signal

Y.G. Gorshkov<sup>1</sup>

Bauman Moscow State Technical University, Moscow, Russian Federation

<sup>1</sup> ORCID: 0000-0003-0483-4603, [y.gorshkov@npo-echelon.ru](mailto:y.gorshkov@npo-echelon.ru)

## **Abstract**

The article presents the results of research on the assessment of human emotional tension by a speech signal using the developed technology of multilevel wavelet analysis. The disadvantages of polygraph systems during psychophysiological inspections of personnel are given. The data obtained by domestic and foreign experts allow us to make a confident conclusion that the characteristics of a person's oral speech correlate with changes in their condition. The characteristics of the hardware and software complex «Icar Lab» (Russia) in assessing the emotional state of the speaker, as well as the computer voice stress analyzer CVSA (USA) are considered. A new contactless technology is proposed for assessing the level of emotional tension of a person using stress biomarkers obtained on the tonal parts of speech. The WaveView-VSA program for obtaining high-precision time-frequency characteristics of a speech signal has been developed. The possibilities of the developed program for visualizing the level of emotional tension are presented. A database of audio recordings of 6th-year students who are in conditions with increased emotional stress during the exam and the defense of the graduation project has been formed and processed. The research conducted by the Scientific and Educational Medical and Technological Center of the Bauman Moscow State Technical University showed the fundamental possibility of obtaining real-time acoustocardiography data and the level of emotional tension in the examined students during express cardiognostics. The developed technology provides biomarkers of stress based on a speech signal lasting several seconds.

**Keywords:** emotional tension, speech signal, multilevel wavelet analysis, biomarkers of stress.

## **1. Introduction**

Assessment of the psychophysiological or emotional state of a person is carried out mainly using polygraphs. When using them, a number of difficulties arise due to both the shortcomings of testing methods and the low quality of the equipment used [1, 2]. The solution to the problem of obtaining a more objective assessment of the functional state of the subjects in world practice is carried out by improving the methods used and both hardware and software.

Almost all known modern polygraph systems during inspections do not include an assessment of a person's condition based on the characteristics of his speech. If there are paths for recording speech in polygraphs, then when processing audio recordings, as a rule, the energy of the speech signal or the sound volume level is estimated. At the same time, acoustic, linguistic and psychophysiological studies have established that the characteristics of a person's oral speech correlate with changes in his condition.

Significant results on the identification of objective signs of emotions in an acoustic signal based on the provisions of the theory of speech formation were obtained by V.I. Galunov [3]. At the same time, it was noted in [4] that no fundamentally new practically significant results have appeared in recent decades. Publications devoted to the analysis of speech signal

characteristics in solving the problem of stress analysis by voice do not always contain quantitative research results that allow formalizing this relationship. Most papers do not provide technical characteristics of the speech recording tools used, as well as recording conditions, which makes it difficult to evaluate and compare the data obtained [5-7].

The purpose of this study is to develop new solutions for determining the emotional tension of a person based on a multi-level wavelet analysis of a speech signal. The first results on wavelet analysis and visualization of emotional speech were published by the author in the journal «Special Technique» in 2006 [8].

## 2. Speech model

It is known that the speech model includes several levels. If we analyze each of them from the point of view of self-control of speech, they differ in many ways. The physiological and emotional levels cannot be controlled, while the identification level is partially controlled. Therefore, it is believed that the verbal and nonverbal components of human oral speech are sufficiently reliable to assess the reliability of the information received (Fig. 1) [9].

		<b>УРОВНИ</b>	
КОНТРОЛИРУЕМЫЕ	<b>С</b> МЫСЛОВОЙ	<i>содержание речи</i>	
	<b>И</b> НТОНАЦИОННЫЙ	<i>чувственный – способ передачи информации</i>	
КОНТРОЛИРУЕМЫЙ ЧАСТИЧНО	<b>И</b> ДЕНТИФИКАЦИОННЫЙ	<i>отражает наличие устойчивых признаков отождествления</i>	
НЕКОНТРОЛИРУЕМЫЕ	<b>Ф</b> ИЗИОЛОГИЧЕСКИЙ	<i>отражает все спонтанные события, происходящие в ЦНС человека</i>	
	<b>Э</b> МОЦИОНАЛЬНЫЙ	<i>обусловлен психо-эмоциональным состоянием человека</i>	

Fig. 1. The structure of the speech model: the main levels of speech

The identification of the reliability of information is facilitated by the fact that a person, due to the specifics of the perception of his own speech, tries to disguise false information by attempts to control his own voice, often not very successful. Signs of insincerity in answering questions include: changes in the tempo and timbre of the voice, intonation, excitement, the appearance of trembling, uncharacteristic pauses, quick answers to questions implying mental processing, the appearance of uncharacteristic turns and expressions in speech or their sudden disappearance, focusing on minor points to hide the true attitude to them [9].

## 3. Assessment of the dynamics of a person's emotional state using the hardware and software complex «Icar Lab»

The method of isolation and analysis of acoustic characteristics of speech to assess the altered psychophysiological state of the speaker was developed by specialists of LLC «Center for Speech Technologies» (St. Petersburg) in 2009. It is intended for experts using the hardware and software complex «ICAR Lab» [10], which is part of a diagnostic system that includes data recording tools, as well as monitoring the dynamics of a person's

psychophysiological state. The methodology complements the «SIS-6 Program User's Guide.X» APK «ICAR Lab» for solving problems of obtaining and analyzing acoustic characteristics of speech in order to assess the dynamics of a person's psychophysiological or emotional state.

Speech analysis to assess the state of emotional tension includes the following stages: drawing up a protocol of speech messages; identification of psycholinguistic signs; measurement of acoustic and temporal characteristics; analysis with the definition of signs of emotional speech. The processing technology, in general, is quite time-consuming and requires a highly qualified expert. Time costs are determined from a ratio of 1 to 10 (i.e. 10 times the duration of the studied fragment of speech). The method uses the principle of psychological and psychophysical scaling.

As a rule, experts who have experience with emotional speech, when analyzing, accurately determine the state of emotional tension and the level of its severity. A five-point scale is used to assess the severity of the state of emotional tension (Table 1).

Table 1. Scale of assessments of the degree of severity of the state of emotional tension

Number of points	The degree of severity of the state of emotional tension
1	Absent
2	Weak
3	Average
4	High
5	Maximum

The expert also uses the information obtained during the instrumental analysis of speech in the form of the dynamics of the kepstrogram (the values of the pitch period).

#### **4. CVSA Lie Detection System based on Voice stress Analysis**

In 1971, Olof Lippold, a scientist from University College London, in an article «Physiological Tremor» («Physiology of tremor») published in the journal «Scientific American» analyzed the results of the discovery made by Martin Halliday and Joe Redfearn based on research performed at the National Hospital of London. Scientists have found that when a person is agitated, an arbitrary muscle contraction is accompanied by a tremor in the form of small vibrations. In addition, it was found that most of the physiological tremor consists of vibrations of a special reflex mechanism that controls the length and tension of muscles in the frequency range from 8 to 12 hertz [11].

In 1988, the National Institute for Truth Verification - NITV USA (National Institute of Truth) introduced the development of Computer Voice Stress Analyzer – CVSA (computer Voice Stress analyzer), which has found wide application in law enforcement agencies.

Since 1991, the CVSA system, implemented on the basis of a powerful multifunctional laptop, has been supplied to government agencies and units of the US Armed Forces. The latest version of the CVSA II analyzer (Fig. 2) has gained a reputation as the most effective investigative tool that has been put into operation by US law enforcement agencies over the past three decades [11].

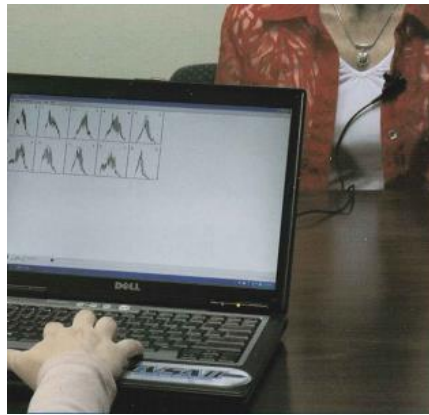


Fig. 2. Appearance of the CVSA II analyzer

## 5. Voice stress Analysis program WaveView-VSA

The NPO «Echelon» has developed algorithms and mathematical support for obtaining sonograms, based on data from multilevel wavelet analysis, displaying a detailed time-frequency structure of the signals of the elements of the vocal tract, rebuilt due to neuromuscular actions on the orders of the brain. The developed WaveView-VSA research program implements several algorithms for processing speech signals, as well as obtaining wavelet sonograms and biomarkers of stress [12, 13].

### 5.1. Examples of a wavelet sonogram

Below (Fig. 3-9) are examples of a wavelet sonogram of speakers' speech in the absence of emotional stress, as well as at various levels of tension. Figure 3 shows the sonogram of the words [Ilya Olegovich] - the name and patronymic of a 6th-year student, - a registered test control phrase.

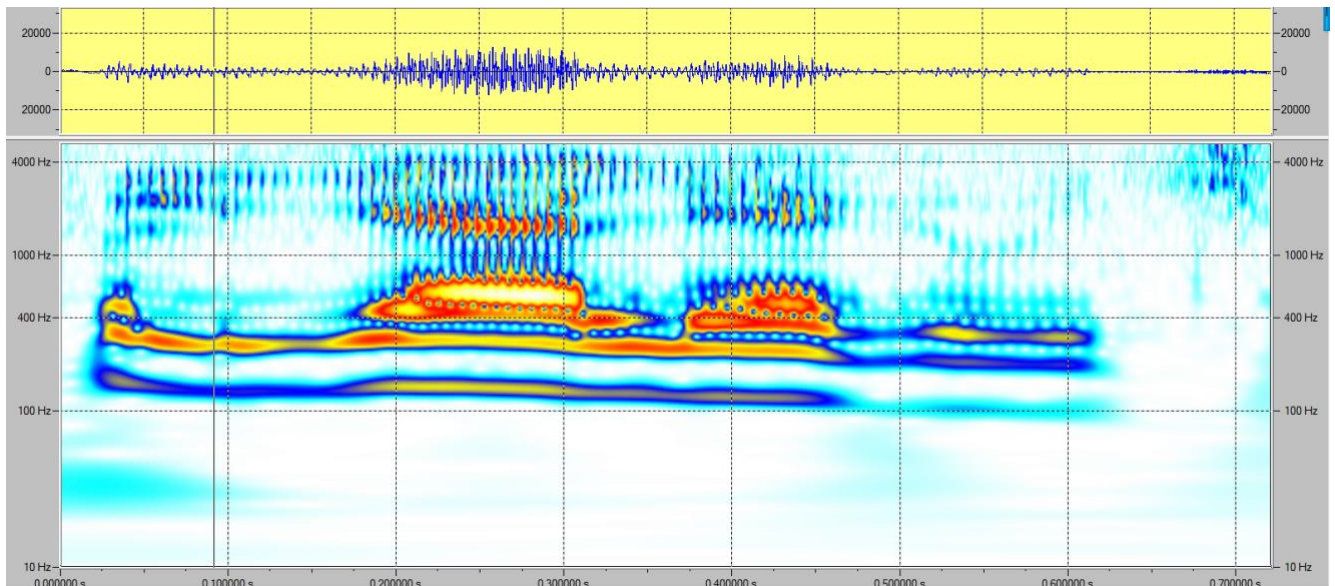


Fig. 3. A wavelet sonogram of the words [Ilya Olegovich], a control (emotionally neutral) phrase uttered by a 6th-year student. In the tonal areas of vowel sounds, a sequence of pulsations of the vocal folds is visible, characterizing the stability of the value of the pitch period; in the low-frequency region, there are no biomarkers of the «tremor» of the voice, characterizing emotional tension

Figure 4 shows a wavelet sonogram of the words [Victoria Igorevna] - the name and patronymic of a 6th-year student, - a registered test record.



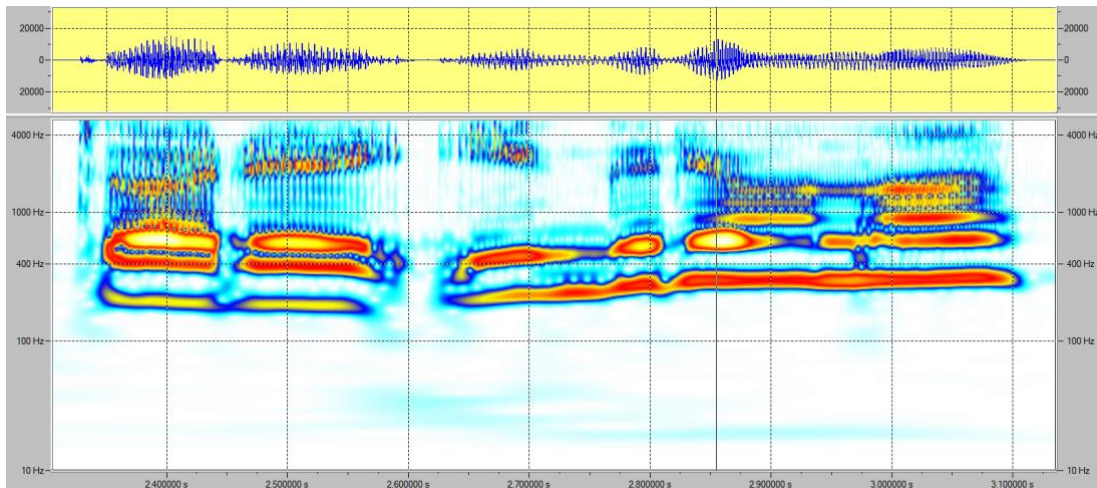


Fig. 4. A wavelet sonogram of the words [Victoria Igorevna], a control (emotionally neutral) phrase uttered by a 6th-year student, a registered test record when performing laboratory work. On the tonal parts of vowel sounds, a sequence of pulsations of the vocal folds is visible, characterizing the stability of the value of the period of the main tone; in the low-frequency region there are no signals of «tremor» of the voice - biomarkers characterizing emotional tension

Figure 5 shows a wavelet sonogram of the speaker's speech signal experiencing stress (the response of a 6th-year student on the exam).

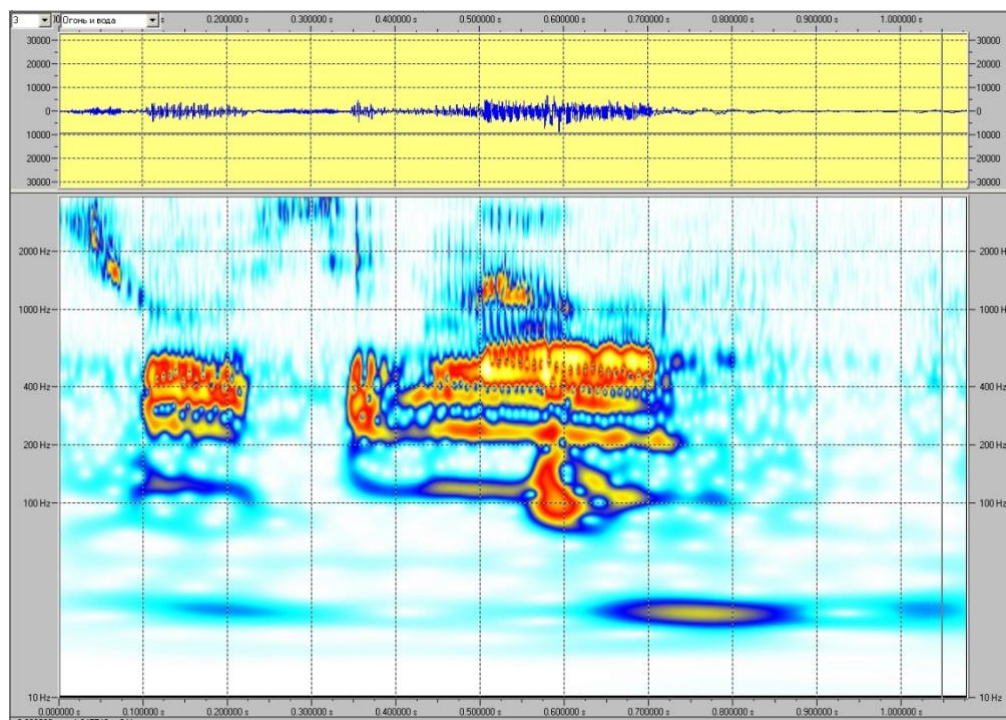


Fig. 5. Wavelet sonogram of the speaker's speech signal experiencing stress (the response of a 6th-year student on the exam)

Signs of stress on the voice (biomarkers) manifested on the sonogram are: - «destruction» of the spectral-temporal structure of vowel sounds; - «micro-trembling» of the speaker's vocal folds on the tonal sections of vowel sounds; - the appearance of oscillations with a frequency of 24-28 Hz in the low-frequency part of the spectrum.

Figure 6 shows a wavelet sonogram of a fragment of the speech signal of a 6th-year student's response to an exam also with a high level of emotional tension.

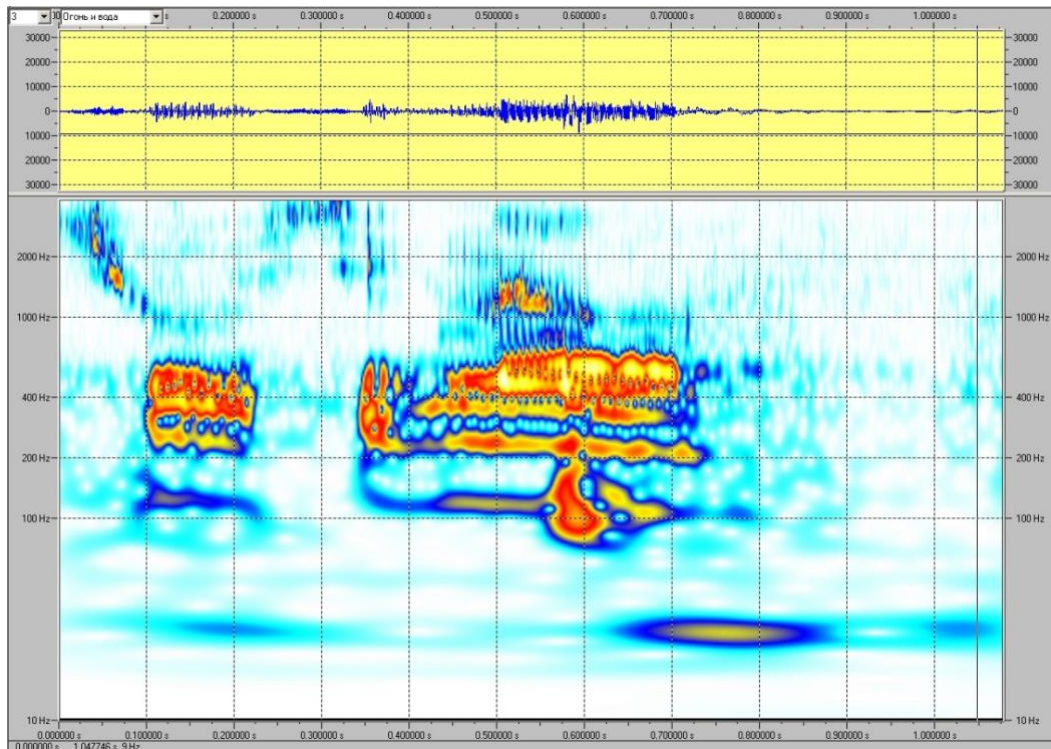


Fig. 6. A wavelet sonogram of a fragment of a speech signal of a 6th-year student's response to an exam; a low-frequency signal of 20-30 Hz characterizes a significant level of stress. On a plot of 1 sec. there is a «destruction» of the spectral-temporal structure of the vowel sound. At a frequency of 50 Hz, a low-level power supply background signal is noticeable

Figure 7 shows a wavelet sonogram of the speech signal of the response to the exam of a 6th-year student.

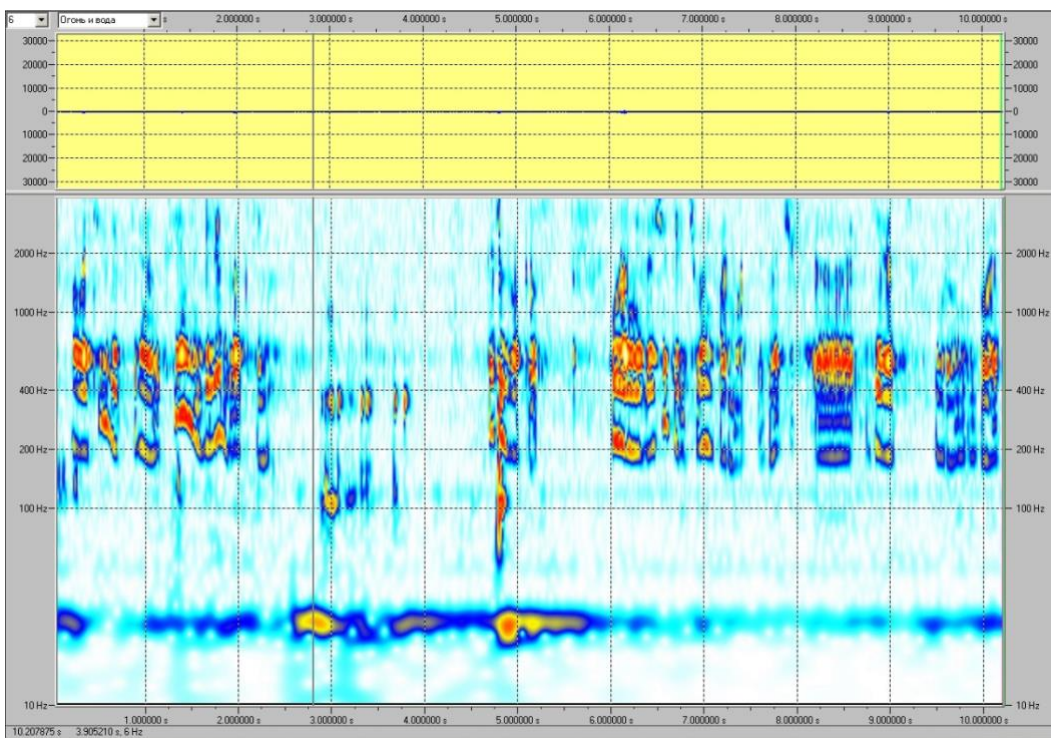


Fig. 7. A wavelet sonogram of the speech response signal at the exam of a 6th-year student; a stress biomarker - a low-frequency «tremor» of the speech tract with a frequency of 24-28 Hz is also observed in speech pauses



Figures 8 and 9, respectively, show the wavelet sonograms of the speech signal of the 6th-year student «A» when passing the exam and defending the thesis project.

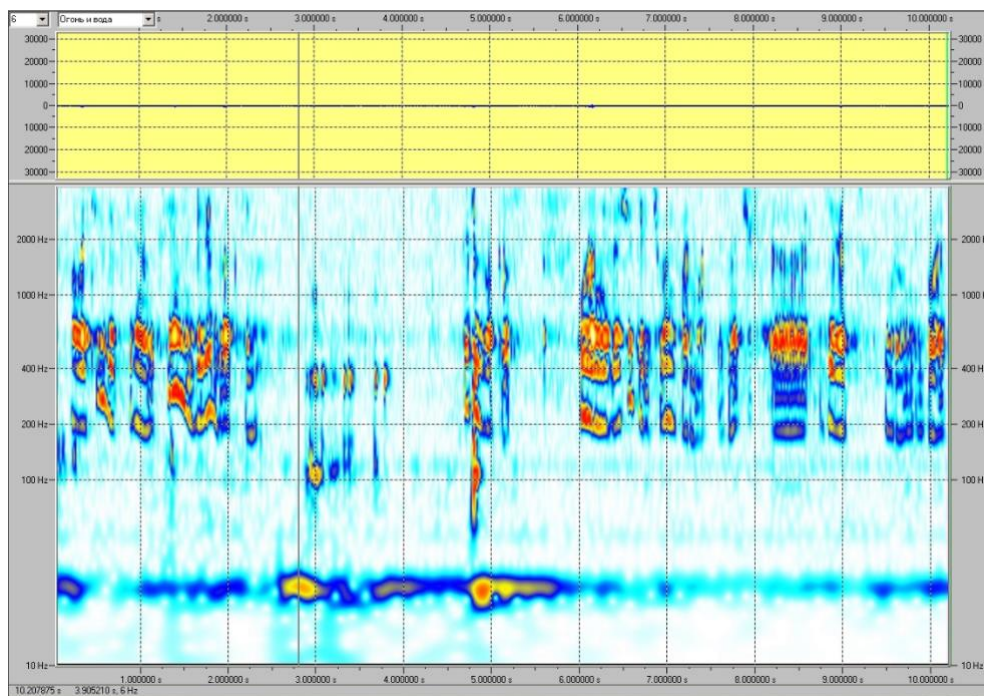


Fig. 8. Passing the exam by the student "A" of the 6th year. The VSA graph is shown at the bottom of the sonogram. The maximum value of the stress level by voice is 2%. The degree of severity of the state of emotional tension is 2 points (weak)

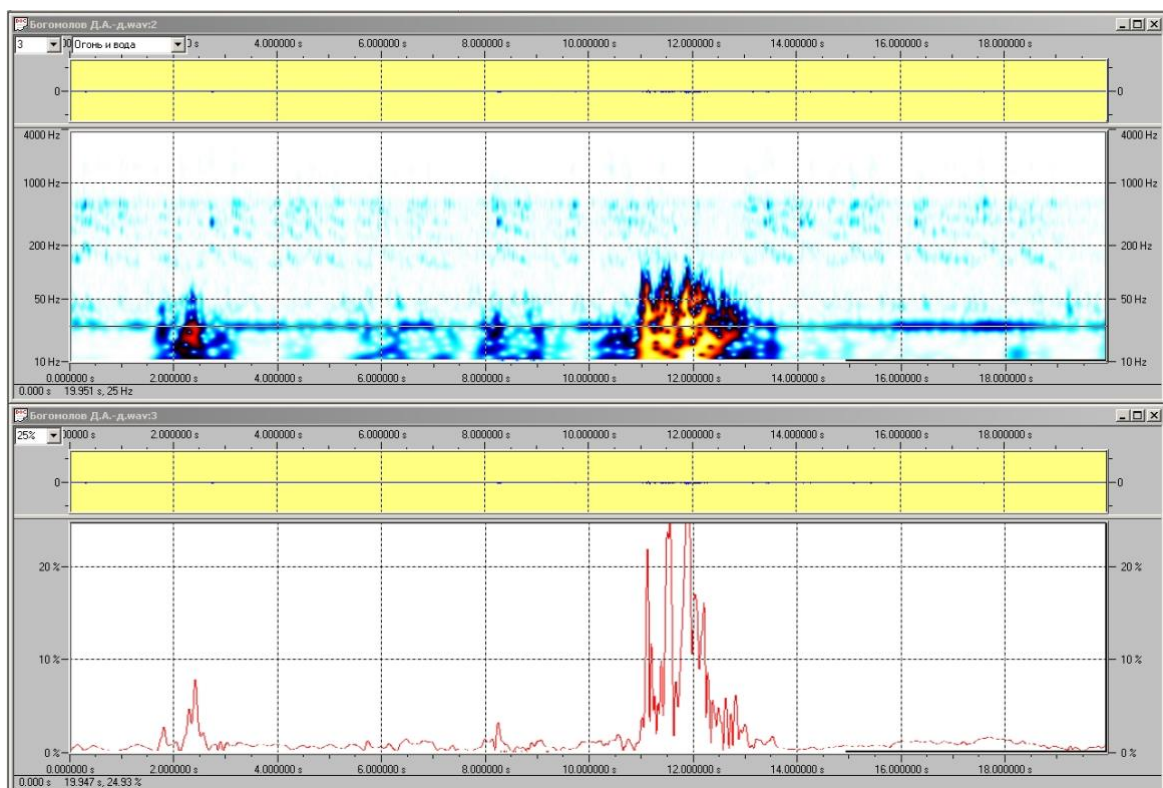


Fig. 9. Defense of the thesis project by the student «A» of the 6th year. According to the VSA schedule, the stress level on the voice reaches 25%. The degree of severity of the state of emotional tension is 3 points (average)

Studies on the assessment of the emotional state of the speaker by voice have been conducted at the Bauman Moscow State Technical University and the Echelon NPO since 2002 [14-19]. A database of audio recordings of 870 6th-year students who are in conditions with increased emotional stress during the exam and the defense of the thesis project was formed and processed. As control (emotionally neutral) recordings, test phrases of students' speech were used, registered by them independently when performing laboratory work. Out of the total number of students: men - 703, women - 129; students of the faculty «Head educational, research and Methodological Center for vocational Rehabilitation of persons with disabilities (hearing impaired)»: men - 27, women - 11.

## **6. Methods of stress analysis by voice**

The technique of voice stress analysis based on a multilevel wavelet transform includes a stage of high-precision recording of a speech signal; selection of audio recording sections for analysis; obtaining, using the WaveView-VSA program, wavelet sonograms; identification of signs characterizing emotional tension – biomarkers of stress.

Phonogram recording facilities should provide recording of both speech and low-frequency biomedical acoustic signals in the range of 10 Hz - 100 Hz. Recommended hardware and software tools for recording audio recordings: Logitech USB Desktop Microphone digital microphone; Logitech USB Headset headset; specialized voice recorder «Protection» (Telesystems, Russia), belongs to a new class of digital voice recorders [20, 21]. The essential advantages of using the dictaphone «Protection» include the fact that its recordings can be used in court as evidence.

## **7. Conclusion**

Over the past few years, an experimental database of audio recordings of 6th-year students of the Faculty of Computer Science and Management of Bauman Moscow State Technical University who are in conditions with increased emotional stress during the exam and the defense of the thesis project has been formed and processed. The voice stress analysis technique is based on the technology of multilevel wavelet transformation of non-stationary signals. It has shown high efficiency in visualizing the sounds of the heart, lungs [22, 23], biomedical signals in telemedicine systems [24], power supply network interference in mobile electrocardiography systems [25], as well as solving problems of forensic investigation of phonograms [26, 27].

The approbation of the developed technology for assessing the level of emotional tension of the speaker by voice was carried out on the materials of audio recordings of 870 speakers (720 men and 150 women), with a total volume of more than 14 hours. The Scientific and Educational Medical and Technological Center of Bauman Moscow State Technical University implemented a pilot project on express cardiognostics of the examined students with an assessment of the current emotional state. The results of the study showed the fundamental possibility of obtaining real-time acoustocardiography data and biomarkers of stress by voice [28]. In addition, when identifying neurological diseases, as well as their possible causes, processing of brain potentials using the technology of multilevel wavelet analysis will allow obtaining new additional diagnostic information.

The materials of the article may be of interest to developers of lie detection systems by voice, as well as new promising solutions for home telemedicine.

## **References**

1. Komissarova Y.V., Myagkih N.I., Pelenicyn A.B. Poligraf v Rossii i SSHA: problemy primeneniya. «YUrlitinform». M.: 2012. 224 p. [in Russian].
2. Pelenicyn A.B., Soshnikov A.P. Osnovnye trudnosti i problemy ispol'zovaniya poligrafa v pravoohranitel'noj deyatel'nosti i kadrovoj rabote i rekomenduemye puti ih preodoleniya.



Polikonius. Sovremennye tekhnologii detekcii lzhi. [https://www.polyconius.ru/company/library/article\\_8.php](https://www.polyconius.ru/company/library/article_8.php). [in Russian].

3. Galunov V.I. Issledovanie variativnosti rechevogo povedeniya cheloveka: Avtoref. diss. dokt. biol. nauk. L.: 1975. 38 p. [in Russian].

4. Galunov V.I. O vozmozhnosti opredeleniya emocional'nogo sostoyaniya govoryashchego po rechi // Rechevye tekhnologii. 2008. № 1. pp. 60-66. [in Russian].

5. Gorshkov Y.G., Dorofeev A.V. Sravnitel'naya harakteristika sistem detekcii lzhi na osnove analiza rechevogo signala na vyyavlenie stressa (VSA) // Sb. trudov X Vseros. nauch. konf. «Problemy informacionnoj bezopasnosti v sisteme vysshej shkoly». M.: 2003. p. 51. [in Russian].

6. Gorshkov Y.G., Dorofeev A.V. Rechevye detektory lzhi kommercheskogo primeneniya // INFORMOST. M.: 2003. № 6. pp. 13-15. [in Russian].

7. Gorshkov Y.G., Efremenkov S.V., Barinov E.V. Specializirovannye sredstva programmno-tekhnicheskogo kompleksa registracii rechevogo signala i kriminalisticheskogo issledovaniya fonogramm // Materialy XXVI Vserossijskoj nauchnoj konferencii «Informatizaciya i informacionnaya bezopasnost' pravoohranitel'nyh organov». M.: Akademiya upravleniya MVD Rossii. 2017. pp. 218-222. [in Russian].

8. Gorshkov Y.G. Novye resheniya rechevyh tekhnologij bezopasnosti // Special'naya tekhnika. M.: 2006. № 4. pp. 41-47. [in Russian].

9. Osyshnaya D.P., Samohina M.A., Ivanov L.N. K voprosu o «golosovom poligrafe». CHast' 1, 2 // Laboratoriya MMPYAiP SGU im. N.G. CHernyshevskogo. Saratov: 2011. pp. 1-5. [in Russian].

10. Centr Rechevyh Tekhnologij. IKAR Lab: Kompleks kriminalisticheskogo issledovaniya fonogramm rechi. [in Russian].

<http://www.speechpro.ru/product/analysis/criminalistic/ikarlab>.

11. National Institute for Truth Verification. The World Leader in Voice Stress Analysis. <http://www.cvsa1.com/History.htm>.

12. Gorshkov Y.G. Sredstva mnogourovnevnogo vejjvlet-analiza stressa po golosu // Sb. trudov XXIV Vseros. nauch. konf. «Informatizaciya i informacionnaya bezopasnost' pravoohranitel'nyh organov». M.: Akademiya upravleniya MVD Rossii. 2015. pp. 169-173. [in Russian].

13. Gorshkov Y.G., Kaïndin A.M., Markov A.S., Cirlov V.L. «WAVEVIEW VSA» Analiz stressa po golosu (programma dlya EVM). Svidetel'stvo o registracii RU 2017662095, 27.10.2017. Zayavka № 2017619125 ot 08.09.2017. [in Russian].

14. Gorshkov Y.G. Apparatno-programmnye sredstva ocenki emocional'nogo sostoyaniya cheloveka po akusticheskim signalam // Instrumental'naya detekciya lzhi – 15 let na strazhe zakona: itogi projdennogo i perspektivy razvitiya. Mezhdunarodnaya nauchno-prakticheskaya konferenciya specialistov-poligrafologov pravoohranitel'nyh organov, 21 – 24 sentyabrya 2009, Kazan'. [in Russian].

15. Gorshkov Y.G. Issledovatel'skij kompleks chastotno-vremennogo analiza rechevogo signala s ispol'zovaniem vejjvlet-tekhnologii // Vestnik Moskovskogo gosudarstvennogo tekhnicheskogo universiteta im. N.E. Bauman. Ser. «Priborostroenie». 2011. № 4. pp. 78 - 87. [in Russian].

16. Gorshkov Y.G. Mnogourovnevyy vejjvlet-analiz akusticheskikh signalov pri reshenii zadach fonoskopicheskoy ekspertizy // Materialy XX Mezhdunarodnoj nauchnoj konferencii «Informatizaciya i informacionnaya bezopasnost' pravoohranitel'nyh organov». M.: Akademiya upravleniya MVD Rossii. 2011. pp. 379-387. [in Russian].

17. Gorshkov Y.G. Ocenka emocional'nogo sostoyaniya cheloveka na osnove mnogourovnevnogo vejjvlet-analiza rechi // Biomedicinskaya radioelektronika. M.: 2014. № 10. pp. 64-70. [in Russian].

18. Gorshkov Y.G. Obrabotka rechevyh signalov na osnove vejjvletov // T-Comm: Telekommunikacii i transport. M.: 2015. № 2, tom 9. pp. 46-53. [in Russian].

19. Gorshkov Y.G. Analiz stressa po golosu na osnove mnogourovneвого vejvlet-preobrazovaniya // Special'naya tekhnika. M.: 2015. № 4. pp. 32-41. [in Russian].
20. Gorshkov Y.G., Dorofeev A.V., Markov A.S., Cirlov V.L. Ustrojstvo ocenki emocional'noj napryazhennosti cheloveka po golosu. Patent na poleznuyu model' RUS 165114 14.09.2015. Zayavka № 2015138941/14. Byul. № 28. [in Russian].
21. Telesistemy. Miniaturnye cifrovye diktofony. [in Russian].  
<http://www.telesys.ru/products/recorders>.
22. Gorshkov Y.G. Vizualizaciya zvukov serdca // Elektronnyj zhurnal «Nauchnaya vizualizaciya». Nacional'nyj Issledovatel'skij YAdernyj Universitet «MIFI» № 1, tom 9, kvartal 1, 2017. pp. 97-111. [in Russian].
23. Y.G. Gorshkov. Visualization of Lung Sounds Based on Multilevel Wavelet Analysis. Scientific Visualization, 2022, volume 14, number 2, pp.18 – 26. (DOI: 10.26583/sv.14.2.02)
24. Gorshkov Y.G. Novye resheniya vizualizacii biomedicinskih signalov v sistemah telemediciny // Elektronnyj zhurnal «Nauchnaya vizualizaciya». Nacional'nyj Issledovatel'skij YAdernyj Universitet «MIFI» № 2, tom 11, kvartal 2, 2019. pp. 56-72. (DOI: 10.26583/sv.11.2.05) [in Russian].
25. Gorshkov Y.G. Vizualizaciya pomekh seti pitaniya v telemedicinskih sistemah mobil'noj elektrokardiografii // Elektronnyj zhurnal «Nauchnaya vizualizaciya». Nacional'nyj Issledovatel'skij YAdernyj Universitet «MIFI» № 1, tom 13, kvartal 1, 2021. pp. 44 – 53. (DOI: 10.26583/sv.13.1.04) [in Russian].
26. Gorshkov Y.G. Vizualizaciya mnogourovneвого vejvlet-analiza fonogramm // Elektronnyj zhurnal «Nauchnaya vizualizaciya». Nacional'nyj Issledovatel'skij YAdernyj Universitet «MIFI» № 2, tom 7, kvartal 2, 2015. pp. 96-111. [in Russian].
27. Gorshkov Y.G. Obrabotka rechevyh i akusticheskikh biomedicinskih signalov na osnove vejvletov / Nauchnoe izdanie. M.: Radiotekhnika. 2017. 240 p. [in Russian].
28. Gorshkov Y.G., Volkov A.K., Voinova N.A. et al. Acoustocardiography with Assessment of Emotional Tension from the Voice. Biomed Eng 53, pp. 383–387 (2020). (<https://doi.org/10.1007/s10527-020-09948-8>)