SECTION 1. CLINICAL MEDICINE

DOI: 10.46299/ISG.2023.MONO.MED.3.1.1

1.1 Spectral acoustic analysis of voice for diagnostics of dysphonia in children

Speech is a complex psychophysical process, and the voice is a unique tool for interpersonal communication, so the quality of the voice is of great importance for the social adaptation of the child in society [1, 2, 3]. In modern society, the frequency of pathologies of the larynx in children is increasing every year. Dysphonia can be congenital or acquired, as a result of injuries, operations, neuropsychiatric, infectious diseases, and the current psychophysical state of the child can cause significant changes in the speech of children [4]. At dysphonia, a change in the voice of children occurs in the form of hoarseness of varying severity from a deaf voice up to a complete lack of voice, which arise from various diseases from acute laryngitis, up to severe lesions in the form of respiratory papillomatosis, congenital membranes of the larynx and milder diseases of the larynx (nodules of the vocal ligaments, unilateral paresis of the vocal cords). Dysphonia is a psychotraumatic disease [5, 6], often causing of illness in the child. Traditionally, invasive methods such as flexible and rigid fibroscopic laryngoscopy, video stroboscopy, electromyography are used to diagnose dysphonia. However, laryngoscopy is an endoscopic examination method that is prescribed to patients with complaints of difficulty swallowing, pain in the throat and neck, partial or complete loss of voice. Currently, the diagnosis and effective treatment of dysphonia is ensured by the introduction of modern objective methods for assessing the state of the voice function of the larynx into the practice of a phoniatrist. The method of computer spectral acoustic analysis of the voice is a method for diagnosing voice dysphonia, which is based on the subjective and objective data of the patient's voice [7, 8], and the acoustic parameters of the voice objectively reflect the qualitative and quantitative characteristics of the voice [9, 10]. Acoustic analysis is an objective, noninvasive modality for the evaluation of voice quality in patients with voice disorders such as laryngitis, laryngospasm, laryngeal tumors, spasmodic dysphonia, and vocal cord paralysis.

The aim of the study is to study the features of the functional characteristics of the voice in children with dysphonia using the method of acoustic analysis of the voice. The objective of this study is to analyze the spectrograms of the voice to differentiate various pathological conditions of the larynx.

In the study, 47 children with dysphonia aged 7 to 10 years were examined, voices were recorded with the pronunciation of the vowels "A", "I", "U" with the same voice power and the corresponding spectrograms were obtained. In the control group, healthy 20 children aged 7 to 10 years old without voice pathology were examined, voices were recorded with the pronunciation of the vowels "A", "I", "U" and the corresponding spectrograms were obtained. Among the patients, 16 children with functional dysphonia and 12 children with organic dysphonia were diagnosed on the basis of laryngoscopy.

The degree of dysphonia was assessed according to the N. Yanagihara scale (1967) - a classic method for determining the degree of dysphonia: 0 - normal voice; 1 - deaf voice; 2 - slight degree of hoarseness; 3 - medium degree of hoarseness; 4 - severe degree of hoarseness; 5 - aphonia. The study group consisted of 47 patients - patients after operations on the larynx, of which 42 were boys, 5 were girls. The patients were after the following operations on the larynx: after operations to remove the nodes of the vocal cords - 20, after operations for recurrent respiratory papillomatosis - 5, after operations for cicatricial deformity of the larynx - 22. According to the degree of dysphonia, the patients were the following degree: 2nd degree - 18 patients, 3rd degree - 15 patients and 4th degree - 11 patients, 5th degree - 3 patients.

Acoustic examinations were performed in a quiet room with a sound field less than 40 dB. The microphone was positioned 25–30 cm away from the patients’s mouth. The microphone was used to record the sound of a voice on a computer system (Intel Core i7-11800H (2.3GHz - 4.6GHz) (8-Core; 16-Threads); GeForce RTX3060 6GB/192Bit/GDDR6, RAM 16GB DDR4). The audio signals of children's voices were recorded on a computer in *wav* format with a sampling frequency of 44 100 Hz, 16 bits [11] and a mono sound recording type using the software "Audacity" and "Adobe audition 2020" was used. Adobe Audition is a open-source audio production software suite that includes a surprising array of editing tools and recording systems including dispaying the sound waveform and spectral analysis of sound.

In the experiments, the voices of children with dysphonia were recorded when they pronounce the vowel sounds “A”, “I”, “U” with the same voice power and the corresponding spectrograms were obtained.

In the computer spectral analysis of the voice of patients, the following 5 indicators of the acoustic parameters of the voice were considered: Energy of discrete signals, Entropy, Maximum signal value, Minimum signal value, Zero number points.

The main voice indicators, on the basis of which it was possible to analyze the speech and non-speech signal of both a sick child and a healthy child, were the following 5 most informative voice indicators [12, 13]:

1) The energy of discrete signals plays an important role in detecting useful signals against the background of noise, since useful signal components have more energy than random noise. The energy of a discrete signal is formed as the sum of the squares of readings on a given segment:

$$E = \sum_{k=1}^{N-1} X^2(t) \qquad (1)$$

Energy characterizes the power expended on the production of voice.

2) Entropy is a measure of useful information in the processes of signal transmission over wires. Signals that are useful and useful to the receiver are information, and signals that are not useful are noise and interference. If the signal at the output of the communication channel is an exact copy of the signal at the input, then, from the point of view of information theory, this means the absence of entropy, and the absence of noise means maximum information. As a measure of the amount of information, Claude Shannon [14] proposed to consider a function that he called entropy:

$$H = -\sum_i P_i \cdot \log_2 P_i \qquad (2)$$

where: P_i is the probability of the i-th outcome occurring.

The measure of entropy does not depend on the volume of the recording, that is, it will not depend on the volume of the recording and on the strength with which the voice is pronounced - entropy is a personal individual characteristic of the voice.

3) Maximum signal value (max SV) is the largest instantaneous signal value during a given time interval. The value $f(x_0)$ of the function y = f (x) at the point $x_0 \epsilon X$ is called the greatest value of the function f (x) on the set X if the inequality $x \epsilon X$ is satisfied for any point $f(x) \leq f(x_0)$.

4) The minimum signal (min SV) value is the smallest instantaneous signal value during a given time interval. The value $f(x_0)$ of the function y=f (x) at the point $x \epsilon X\ is$ called the least value of the function f (x) on the set X if the inequality $f(x) \geq f(x_0)$ is satisfied for any point $x \epsilon X.$

5) The number of zero points (Zero number points) is an acoustic parameter containing information about the number of zero crossings (zero amplitude). From the point of view of program development, the concept of a counter is introduced, which fixes the number of zero points and the counter is increased by one if positive and then negative values are found in the vector of signal values, or vice versa. It is an indicator of the noise level of both speech and non-speech signals.

Spectrograms of the voice of functional dysphonia (in the absence of pathological changes in the larynx) and organic dysphonia (laryngitis, vocal nodules, cicatricial changes) are shown in Fig. 1. On the spectrograms, the abscissa axis shows the real phonation time in seconds, and the ordinate axis shows the strength of the converted signal in millivolts.

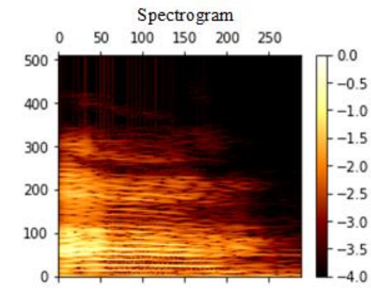
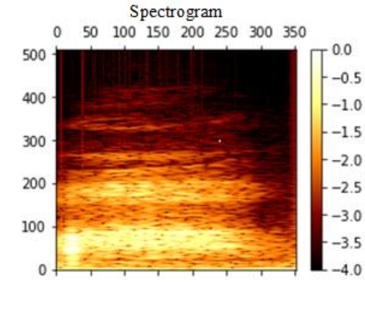
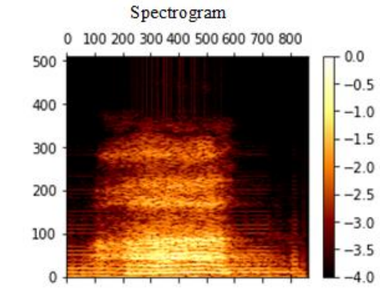
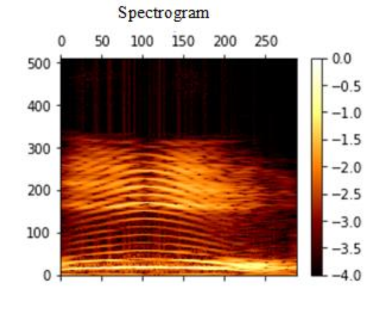
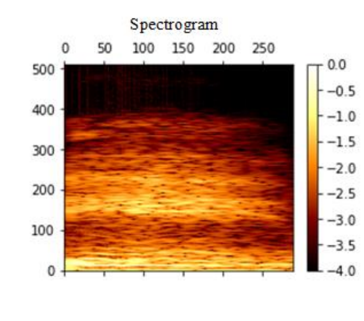
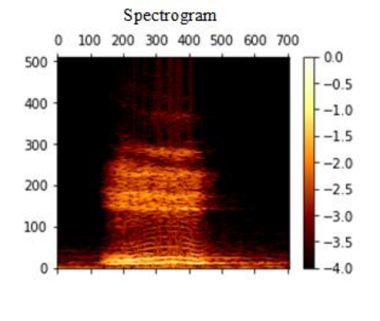
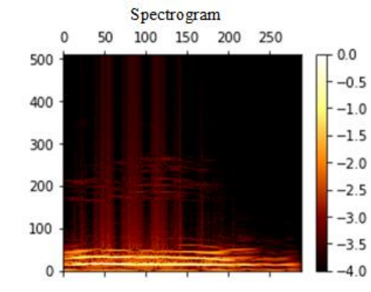
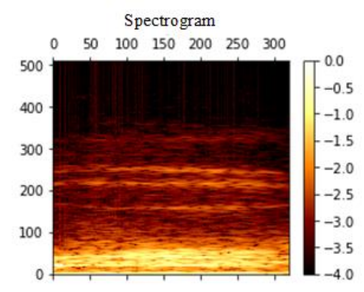
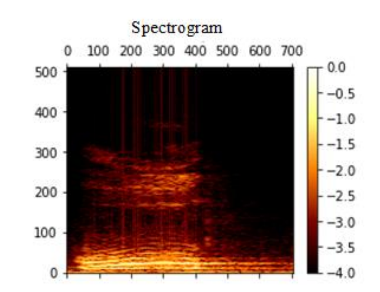
Sounds	Spectrograms 1	Spectrograms 2	Spectrograms 3
A			
I			
U			

Figure 1. Voice spectrograms: spectrograms 1 - normal voice; spectrograms 2 - nodules of the vocal cords (2 degrees of dysphonia); spectrograms 3 - cicatricial changes (3, 4 degrees of dysphonia)

The results of the analysis of the spectrogram (Figure 1) in functional and functional dysphonia showed that in spectrogram 1 (healthy child) no changes in the tone of the vocal folds were detected when pronouncing the sounds "A", "I", "U" (pubertal dysphonia). From spectrograms 2 and 3 it can be seen that a change in the voice of children with organic dysphonia is observed with asynchronous movements of the vocal cords and their irregular nature. Analysis of the spectrograms with vocal nodules and overstrain of the voice showed that there was a decrease in the strength of the voice by 10-15 dB, the time of maximum phonation (PMF) was shortened to 8-12 seconds, the fundamental tone frequency (FOT) was reduced to 8-10 Hz, and instability of the FOT was observed. by the amplitude in the short and long-term periods of tone pertubation or a decrease in voice trembling (Jitter 1.53 + 0.32%) and the degree of amplitude instability of the voice (Shimmer 9.2 + 1.1%), as well as an increase in the level of noise components in the spectrum: index of the ratio of noise spectral energy to harmonic (NHR 0.16 + 0.53%) and turbulence index (VTI> 0.06). It can be seen from Figure 1 that with grosser changes in the organic nature, voice tension is observed, due to the fact that in order to pronounce the sounds “A”, “I”, “U”, dysphonic children strain in order to remove vowel sounds and increase the strength of the voice and tone muscles, however, there is an increase in the soft phonation index (SPI) from 1.5-2 times to 27.6 + 2.7%, and the Jitter indicators increase - 1.56 + 0.23%, Shimmer - 7.95 +2.31%, SF0 - 1.35+0.34% and VTI>0.07, which indicates a significant overstrain of the muscles of the larynx. At the same time, an increase in the strength of the voice by 20 dB was noted, at which the VMF shortens to 10 seconds and the frequency response decreases to 20-25 Hz.

The analysis of acoustic characteristics showed that entropy in dysphonia of the 2nd degree and dysphonia of the 3rd, 4th degree did not have strong differences. However, its «min SV» and «max SV» components differed from the control group: for example, in the group with more pronounced impairments (dysphonia 3, 4 degrees) - the indicators differed from 3.2 to 1.7 times, in the group with dysphonia of the 2nd degree differed from 3.8 to 2.02 times. At the same time, the «zero» indicator (zero number points) was proportional to the possibility of the presence of noise when pronouncing the sounds “A”, “I”, “U” due to wave-like movements of the vocal cords. The index of zero points (number of zero points) characterizes useful and unhelpful movements of the vocal cords, and the more proportional these movements, the lower the index of zero points. Thus, in dysfunction of the movement of the vocal cords caused by violations of oscillations in the dysfunction of the fibers in the vocal cords, the index of zero points was 3.9 times higher than this index in the control group of healthy children. Sick children with cicatricial changes had a "zero" index 1.6 times greater than in the control group, which indicates that "zero" index is an indicator of the movement of vocal cord vibrations, especially their synchronous movement. At the same time, asynchronous oscillations of normal vocal cords with a violation of the formation of a sound wave form a large noisiness with an increase in the “zero” indicator. Also, the indicators of the minimum and maximum values of the voice, and especially the indicator of the energy spent on pronouncing the sound, are more informative, while the more energy expended, the higher the energy indicator, and the greater the pathology of dysphonia and the worse the voice of the sick patient (Fig. 2).

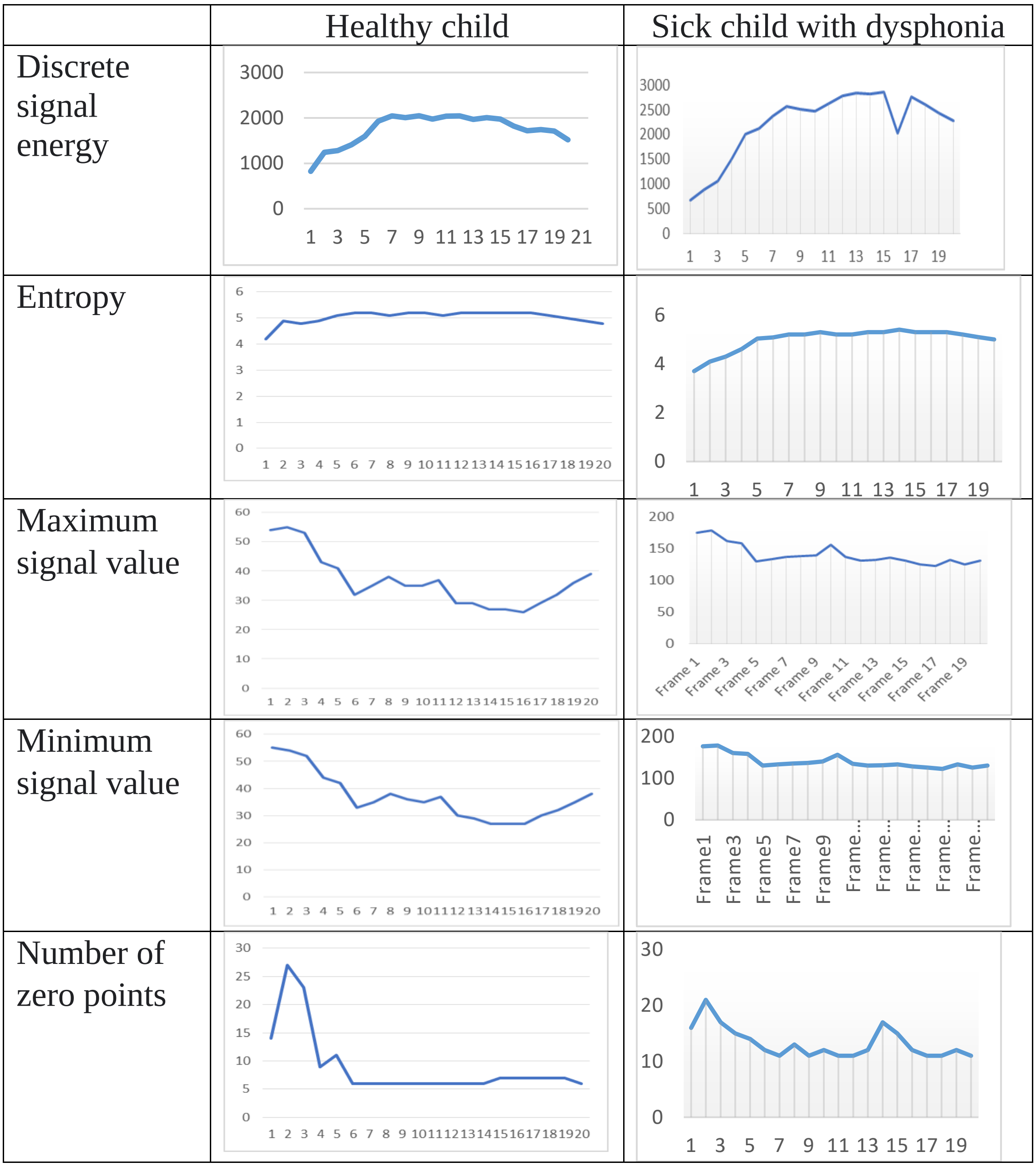


Figure 2. Diagrams of the main acoustic characteristics of the voice in the control group (healthy children) and in the study group (children with dysphonia).

Diagram of the indicator Energy of discrete signals is shown in Fig.3.

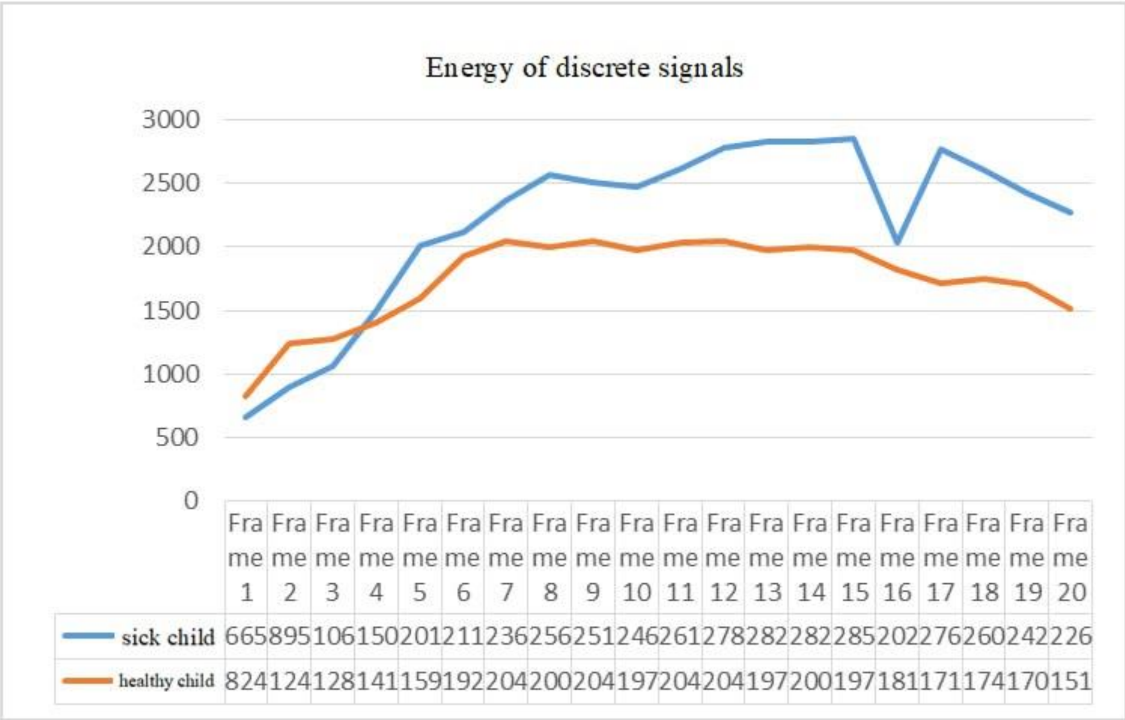
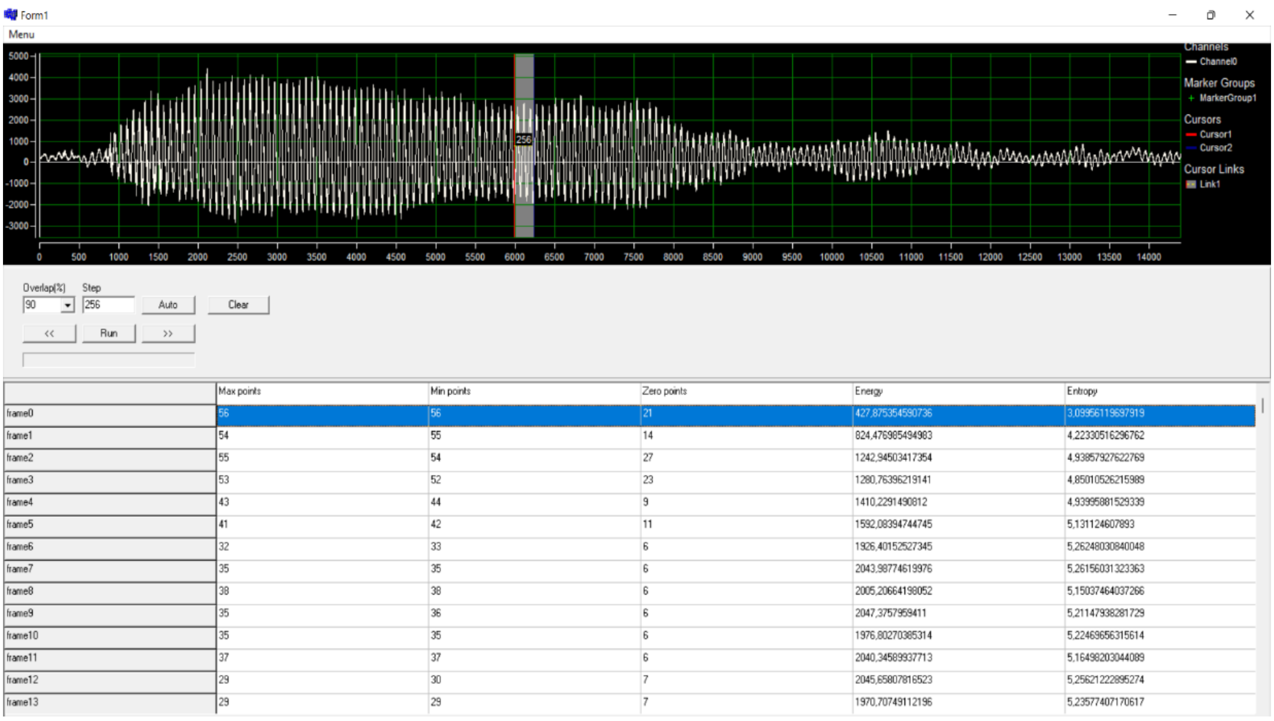
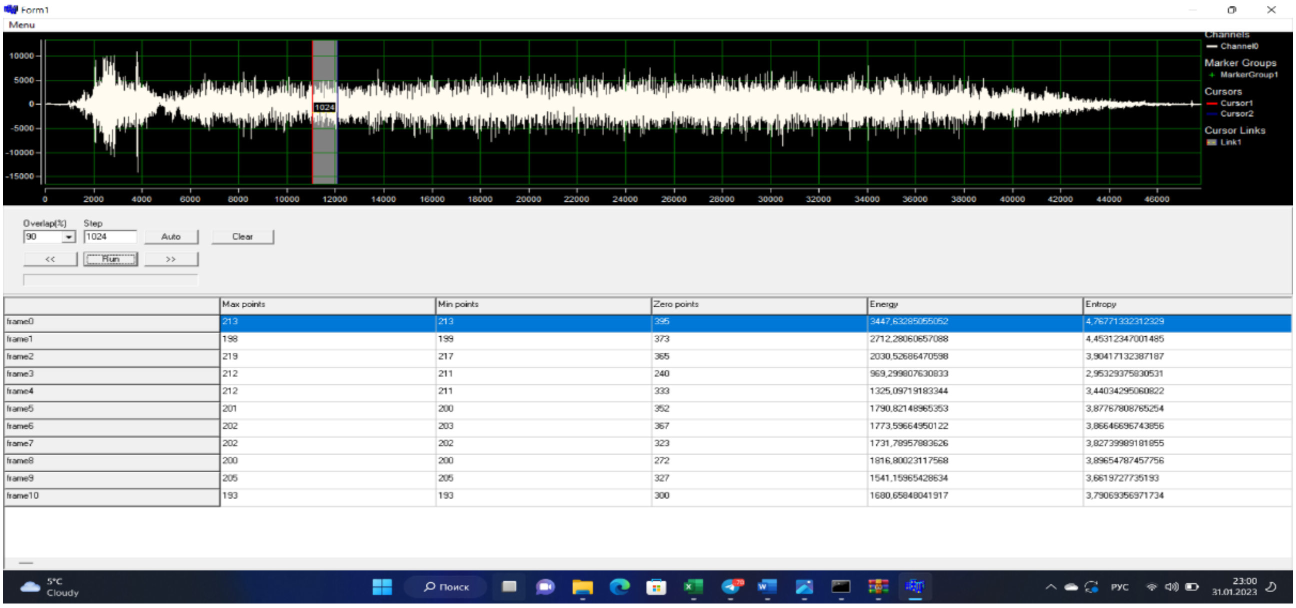


Figure 3. Diagram of the indicator Energy of discrete signals (red line - standard (healthy child), blue line - (sick child with a scar on the larynx).

The audio signal parameters of the voice of a healthy child (A) and a child with dysphonia (B) are shown in Fig. 4.



A



B

Figure 4. Audio signal parameters of the voice of a healthy child (A) and a child with dysphonia (B): A – a healthy child without dysphonia; B - a sick child (cicatricial changes after the operation to restore the lumen of the glottis after cicatricial stenosis of the larynx).

Acoustic voice analysis is used to differentiate normal from abnormal voices in the diagnosis of dysphonia. Studies have shown that the method of acoustic voice analysis allows not only to reliably determine the degree of voice dysfunction, but also to assess the dynamics of its recovery in the course of treatment. The voice parameters obtained in the course of the study are of practical importance for assessing the voice function in childhood, including for dynamic monitoring of the effectiveness of the treatment of laryngeal diseases after surgery. Studies have shown that the acoustic spectral components of the voice, during the formation of vowel sounds, "A", "I", "U" significantly differ from the norm in various types of pathologies of the larynx. With a long, recurrent course of dysphonia, the results of the analysis of the stability of the fundamental tone and the duration of the phonation of various vowels help to diagnose the type of pathology of the larynx, predict the course of the disease, and assess the degree of probability of the transition of functional changes to an organic state.