MATHEMATICAL MODEL OF VOCAL SIGNALS FOR THE TASKS OF HUMAN VOCAL APPARATUS DIAGNOSTIC

Vasyl Dozorskyy, Leonid Dediv, Oksana Dozorska

Ternopil Ivan Pul'ui National Technical University, Ternopil, Ukraine

Address for Correspondence: Leonid Dediv, Ph.D,

Institutional affiliation: Ternopil Ivan Pul'ui National Technical University, Ternopil, Ukraine

E-mail: dediv@ukr.net

Abstract. The necessity of voiced fricative sounds processing is grounded for the problem of human vocal apparatus diagnosis.The analysis of possibility of fricative sound modeling as the determined and stationary random process is conducted. The inadequacy of these models is found out for the task of vocal organs diagnostics on the early stages of disease, because they do not take into account the signal periodicity, chance of forms of pronunciation violation, related to the functional state of vocal organs. Grounded mathematical model of fricative sound as the periodically stochstic random process with the use of energy theory of stochastic signals, which takes into account a time structure, combination of chance and periodicity of signal.

Keywords: vocal signals; fricative sound; periodically correlated stochastic process; human vocal apparatus; diagnostic.

Introduction. according to statistics of Ministry of Health of Ukraine and the World Health Organization, every year there is a tendency to increase the number of people with diseases of the vocal apparatus organs. Therefore, an important task of modern medicine is timely diagnosis of pathological changes in the vocal apparatus in the early stages of their emergence and development.

Pathological changes of the vocal apparatus lead to irregularities in their work [1]. It finds a clear reflection in voice signals - sonorant consonants, due to the specifics of their creation, and as a result, leads to them noise component. Early diagnosis allows to detect changes in the functional state of the vocal apparatus by proper processing of vocal signals and to conduct preventive measures or to choose medical treatment. For an objective diagnosis in medicine are used indirect methods based on system-signal concept, according to which the voice signal is treated as a physical process that extends from the object and is a means of transferring information about this object. The efficiency of the diagnostic system functioning is

crucially determined by the mathematical model of the signal, thet is at its basis, and shall include in its structure an informative sign changes in the vocal apparatus. It is necessary to substantiate the algorithms of measurement and processing characteristics of voice signals and interpreting the results.

The actual problem is the grounding of voice sounds mathematical model choice and development of methods of them analysis for automated computer diagnostic systems, oriented to the problem of early diagnosis of human vocal apparatus functional state by introducing a new class of informative signs, that are indicators of pathological conditions of the human vocal apparatus.

Materials and Methods. Mathematical model of the vocal sound signals should be adequate as research problems and their physical nature and structure. For the diagnosing problems is necessary to processing the class of voice signals, which are the most sensitive to changes in the functional state of the vocal apparatus. According to the logopedic statistics [1-3], in the presence of disruption in the work of vocal tract, the irregularities in the pronunciation of voiced fricative sounds (VFS) are observed. For the research the sound [l] was selected, the creation of which utilized almost all organs of human vocal apparatus, so this voice signal contains much information about the work of vocal apparatus.

Analyze the process of VFS creating by humans vocal apparatus [4,5]. With the creation VFS in the flow of exhaled air (Fig. 1 (1)) source of signals form the sound of signals with a typical repeatability - the basic tone (Fig. 1 (3)) generated by the vocal folds, which are excited quasi-periodic sequence of nerve impulses p(t) (Fig. 1 (2), Fig. 2). Articulation apparatus forms the phonetic structure of signals x(t) (Fig. 1 (4), Fig. 2). So voice signal y(t) can be represented as a complex amplitude modulated signal y(t)=p(t)·x(t), where y(t) - signals VFS (sound [l], shown in Fig. 2) - message; p(t) - carrying signal, characterizing the work of signals source; x(t) - the envelope of the voice signals in the time domain, which describes the behavior of the speech apparatus in time. Analysis of envelope and carrier signals of VFS in the time and frequency domain will enable estimate of the source of signals and speech apparatus as a whole and its organs in particular. Accordingly mathematical model must take into account the frequency, caused by the vocal folds work, and have a means of analyzing time-phase structure, that reflects the behavior of the vocal tract during creation VFS [l] and the need to identify the moment of occurrence malfunction of these organs.

Grounding the VFS mathematical model. Fig. 3 shows the experimentally registered sounds [l] from the patients in a state of medical norm and pathology.

Due the requirements of the VFS mathematical model, defined physical nature and task of diagnostic, namely, taking into account the such signals frequency and its time-phase structure, the simplest model of VFS is the deterministic model as a mixture of periodic functions [6]. Informative and invariant signs of the signal are the morphological features of signals time structure and amplitude spectra, evaluation of which characteristics underlie the methods of harmonic and formant analysis.
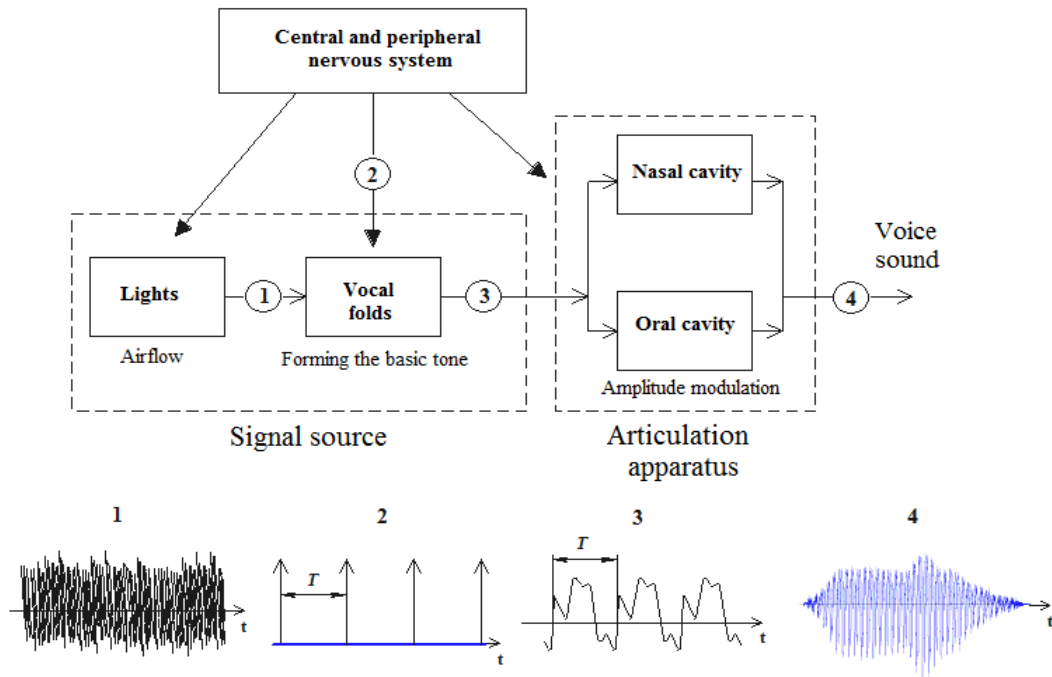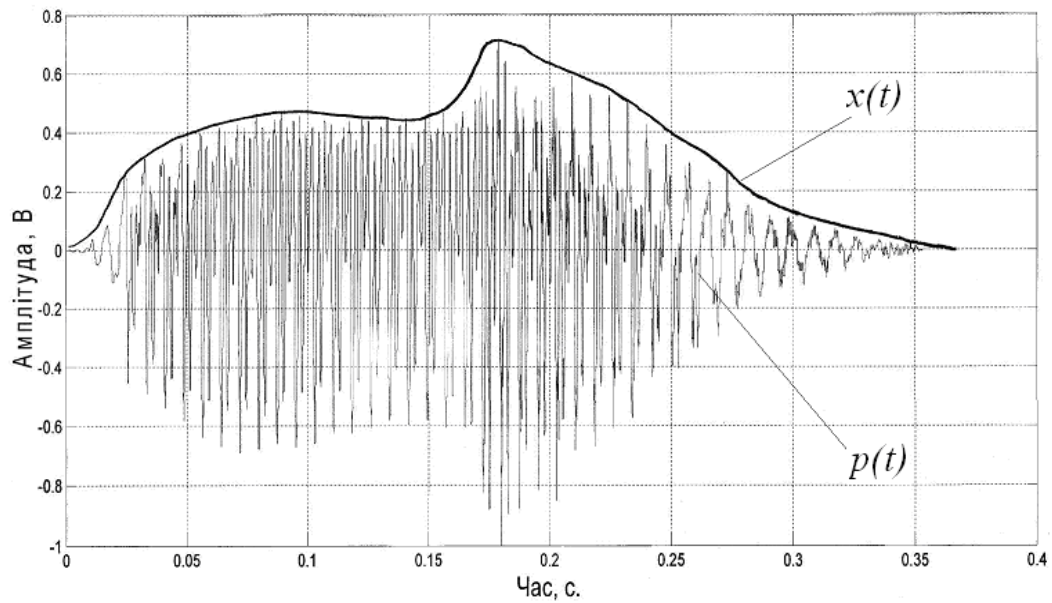
Fig. 1. The process of VFS creating

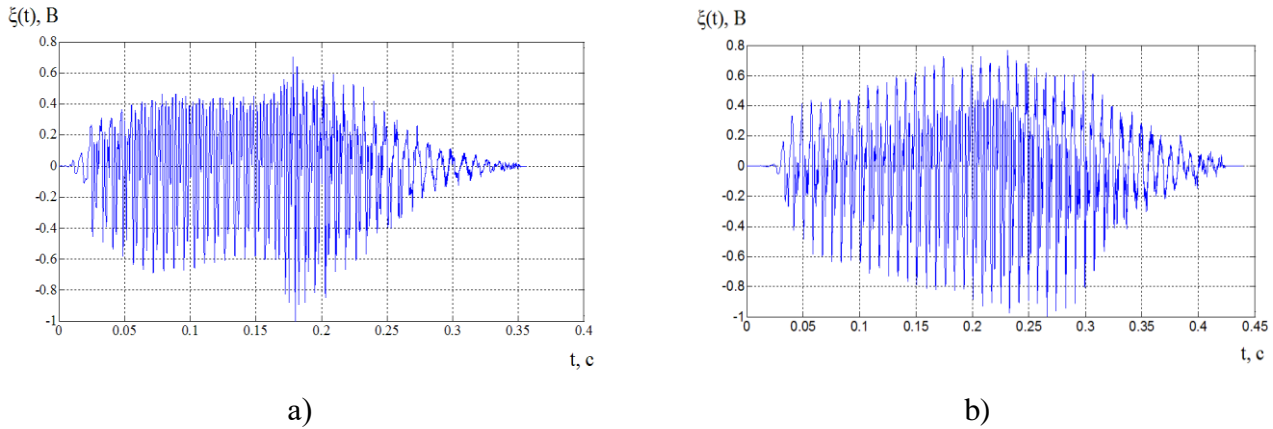Fig. 2. VFS, as a result of the speech formation process

a)   b)

Fig. 3. Experimentally registered sounds [*l*] (a) - the state of the vocal tract is normal; (b) - the state of the vocal tract is pathology

The estimates of amplitude spectra for the same volume of samples, taken from VFS signals are shown in Fig. 4.
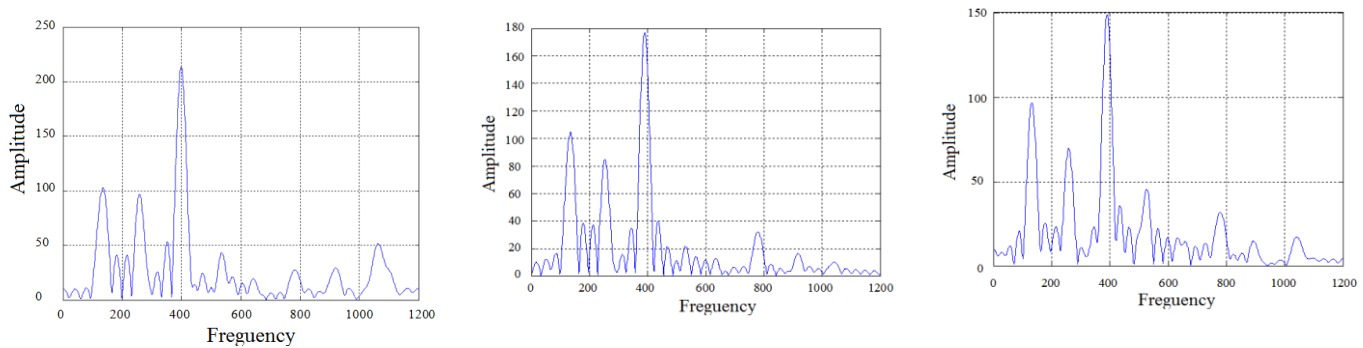

Fig. 4. Estimates of amplitude spectra of samples from experimentally registered sounds sound [*l*]. Duration sample - 0.03 sec

Results. The results of VFS harmonic analysis methods in presenting them as deterministic periodic process are confirming that the resulting amplitude spectra VFS reviews (Fig. 4) are variable. This fact indicates the presence of stochastic component in such signals.

In the case of probabilistic approach to modeling VFZ the stationary model is known [6], which defines the methods of spectral-correlation analysis. In this case, the informative and invariant features of its signal are the probability characteristics and their distributions. This mathematical model reflects the complexity VFS in the spectral power distribution, but it does not reflect the time-phase structure, which is important to identify the time moments of occurrence of changes in the vocal apparatus functioning In terms

of energy theory of stochastic signals [7] these requirements satisfies a model in the form of periodically correlated stochastic process (PCSP), which has a means to take account of connectedness the signal frequency and probability of its characteristics changes.

PCSP of class π^T – is the process, which mathematical expectation is the periodic function, whose period equal to the period of correlation the process and the correlation function satisfies the conditions [7]:

1) periodicity acording to the shifts on the period of correlation $r_\xi(t+T, s+T) = r_\xi(t,s)$, $T > 0$ for all t, $s \in \mathbf{R}$;

2) finite the middle power during this period $\dfrac{1}{T}\displaystyle\int_0^T r_\xi(t,t)dt < \infty$.

Image the stochastic oscillating process (SOP) as PCSP should submit it via stationary components [7]

$$\xi(t) = \sum_{k \in \mathbf{Z}} \xi_k(t) \cdot e^{ik\frac{2\pi}{T}t}, t \in \mathbf{R}, \qquad (1)$$

where $\xi_k(t)$ – stationary and stationary associated components of SOP in presenting it as PCSP; $\mathbf{Z}$ – the set of all integers; T – signal period of correlation.

Components $e^{ik\frac{2\pi}{T}t}$ of expression (1) reflects the harmonious (fluctuating) structure of signals, and components $\xi_k(t)$ – its stochasticity.

However, the analysis of temporal structure of VFS shows that they are close in appearance to amplitude modulated oscillations. In this case, using the concept of beating from the vibrations theory, VFS advisable to interpreted as poly-PCSP, and as the most important for applications – bi- PCSP. Bi-PCSP has image [7]:

$$\xi(t) = \sum_{k,j \in Z} \xi_{kj}(t) e^{i\lambda_{kj}t}, \qquad (2)$$

$$\lambda_{kj} = k\Lambda_0 + j\Lambda, \ \Lambda_0 = 2\pi/T_0, \ \Lambda = 2\pi/T, \ k,j \in Z,$$

where T_0, T – the periods of correlation.

However, there are two ways to image bi-PCSP by the periodically correlated components:

1) $\xi(t) = \sum_{k \in Z} \xi_k^{(T)}(t) e^{ik\Lambda_0 t}$, де $\xi_k^{(T)}$, $k \in Z$ – bi- periodically correlated stochastic processes

with the same period correlation $T = 2\pi / \Lambda$;

2) $\xi(t) = \sum_{j \in Z} \xi_j^{(T_0)}(t) e^{ij\Lambda t}$ through periodically correlated components $\xi_j^{(T_0)}$, $j \in Z$, with

common correlation period $T_0 = 2\pi / \Lambda_0$. In this way the reduction of order of stochastic processes correlation and construction bi-PCSP to PCSP is implemented.

So as the model of VFS is used PCSP as a result of lowering the order of correlation, which features VFS envelope in the time domain will be included in the characteristics of its carrier.

The model of VFS in the form of PCSP defines the common methods of statistical processing of signals: synphase and component, and used to calculate the statistical assessments of their probability characteristics, which in the case of VFS processing are informative and indicators of the human vocal apparatus state.

Discussion. From the analysis of the VFS structure and described properties of periodically correlated stochastic processes implies that the mathematical model of such class of processes makes it possible to adequately describe the signal VFS, namely to consider a combination the stochasticity and periodicity of signals, and therefore develop the methods for determining invariant informative signs of sound, based on statistics of such signals for the task of diagnosis of human vocal apparatus diseases.

Conflict of interest statement: The authors state that there are no conflicts of interest regarding the publication of this article.

Author Contribution: Conceptualization: Vasyl Dozorskyy. Data curation: Vasyl Dozorskyy. Formal analysis: Vasyl Dozorskyy. Writing - original draft: Oksana Dozorska. Writing - review & editing: Leonid Dediv.

REFERENCES:

1. Jafek B, Stark A. ENT secrets. Philadelphia, PA: Hanley & Belfus; 1995.

2. Babiyak V, Nakatis Ya. Clinical otorhinolaryngology: a guide for doctors. St. Petersburg: Hippocrates; 2005.

3. Palchun V, Kryukov A. Otorhinolaryngology: a guide for doctors. Moscow: Medicine; 2001.

4. Fant G. Acoustic theory of speech production. The Hague: Mouton; 1970.

5. Sadaoki F. Digital speech. Processing, synthesis and recognition. Tokyo: Tokyo institute of technology; 2000.

6. Rabiner LR, Shafer RW. Digital processing of speech signal. New Jersey: Prentice-Hall; 1978.

7. Dragan Ya. The energy theory of linear models of stochastic signals. Lviv: The center of strategic research of eco-bio-technical systems; 1997.