

У статті подано розроблений авторами метод виділення вокалізованих ділянок голосового сигналу. Розроблений метод базується на визначенні максимуму короткотривалого спектру сигналу. Виділено ряд особливостей голосового сигналу, які впливають на роботу методу, таких як логічні наголоси, атака та гасіння звуку на початку та в кінці фрази, та здійснено ряд заходів для зменшення їх впливу на результати роботи методу. Експериментальне дослідження розроблених методів на реальних акустичних записах дозволило визначити критерії розмежування вокалізованих ділянок та ділянок шуму, а також двох суміжних вокалізованих ділянок. Результати досліджень показують універсальність розробленого методу, тобто можливість і доцільність його використання незалежно від віку, статі людини, а також мови та тексту аналізованої фрази.

Ключові слова: голосовий сигнал, сегментація, вокалізована ділянка, спектр.

O.YU. MIESHKOV, O.O. NOVIKOV, S.M. ZLEPKO

Kherson National Technical University, Vinnytsia National Technical University

### LOCAL MAXIMUMS METHOD FOR VOCALIZED FRAGMENTS EXTRACTION FROM VOICE SIGNAL

An authors' method for vocalized fragments extraction from voice signal is given in the article. The developed method is based on the determination of the maximum of the short-term spectrum of the signal. A number of peculiarities of the voice signal that are poured into the work of the method, such as logical accents, attack and suppression of sound at the beginning and end of the phrase are highlighted, and a number of measures were taken to reduce their impact on the results of the method. Experimental study of the developed method on real acoustic records allowed determining the criteria for the delimitation of vocalized fragments and noise areas, as well as two adjacent vocalized fragments. The results of research show the universality of the developed method, that is, the possibility and expediency of its use, regardless of age, gender of the person, as well as the language and the text of the analyzed phrase. The main criterion for the delimitation of vocalized sections and noise regions of a signal is the local maxima of the short-range signal spectrum. The authors considered the peculiarities of using the developed method on the basis of real acoustic signals. The initial study was conducted on the pass phrase "Things are going well", which contains all the sounds of the Ukrainian language. The research was conducted on acoustic materials of announcers of different ages, sexes and in rooms with different acoustic characteristics. In all cases, the developed algorithms had the accuracy of the selection of vocalized areas from 3 to 7.5%, which is satisfactory for the tasks of the analysis of the voice signal. At the same time, the study of developed algorithms in the English-language material also yields results that are similar in accuracy. This confirms that the developed algorithms are universal and can be used by different people, regardless of age, gender, anthropometry, recording conditions, human language, and the text of the phrase under discussion.

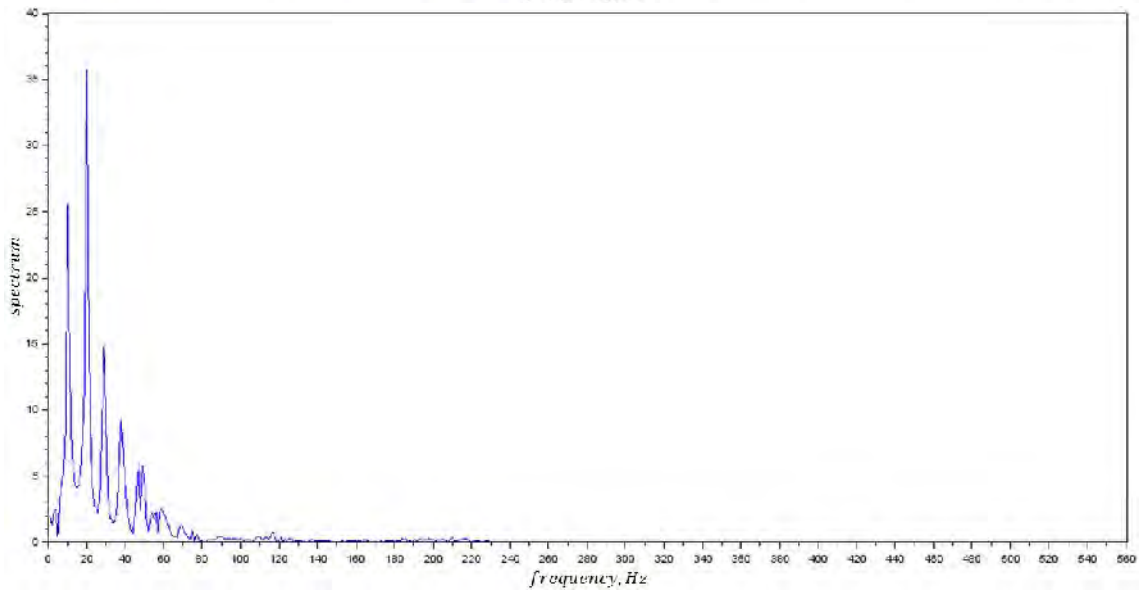
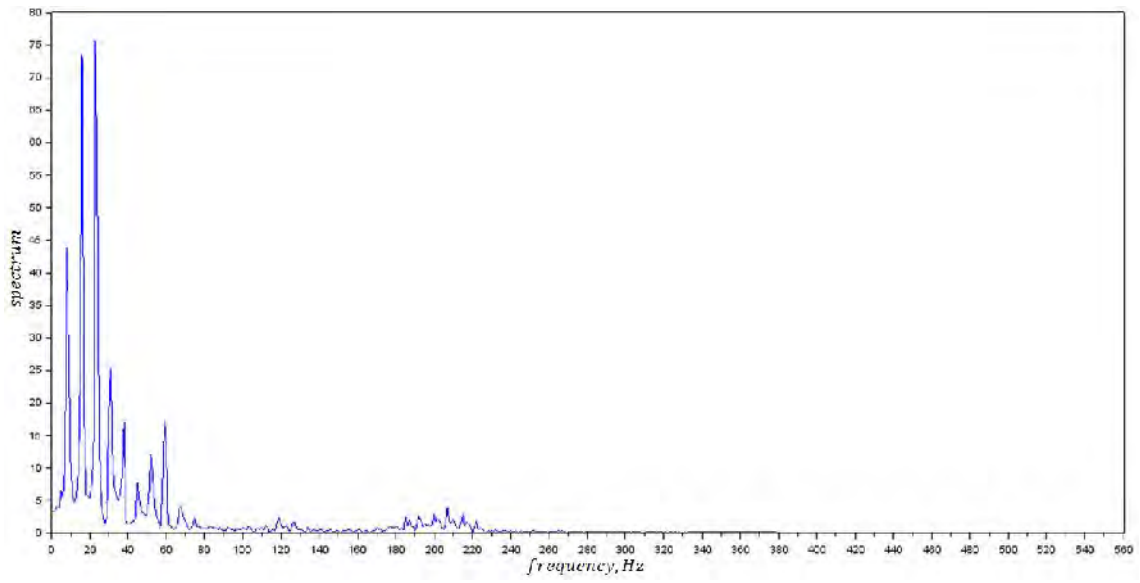
Keywords: voice signal, segmentation, vocalized fragment, spectrum.

[1].

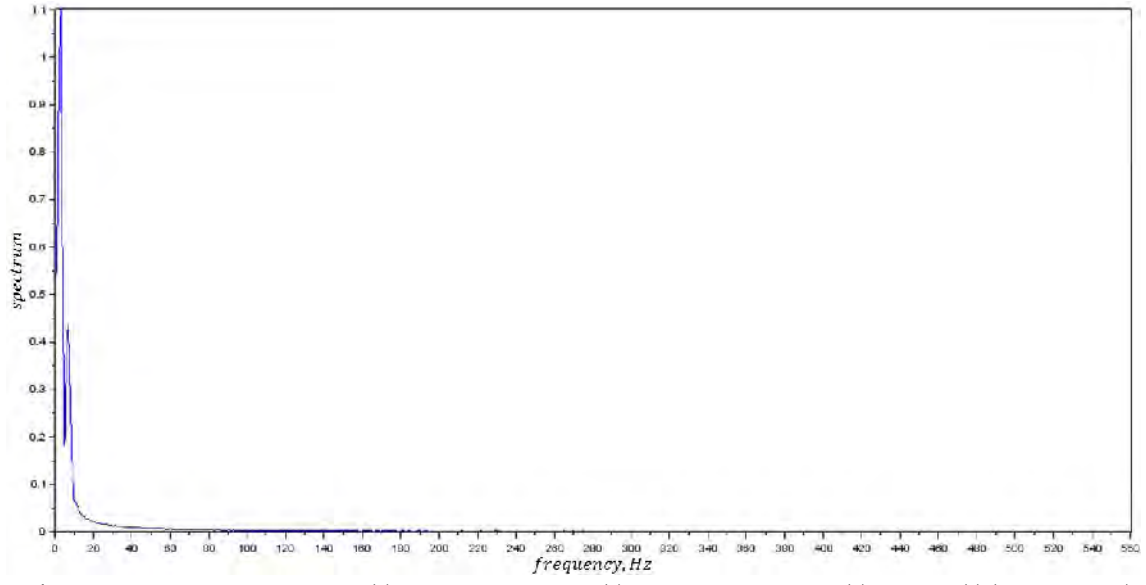
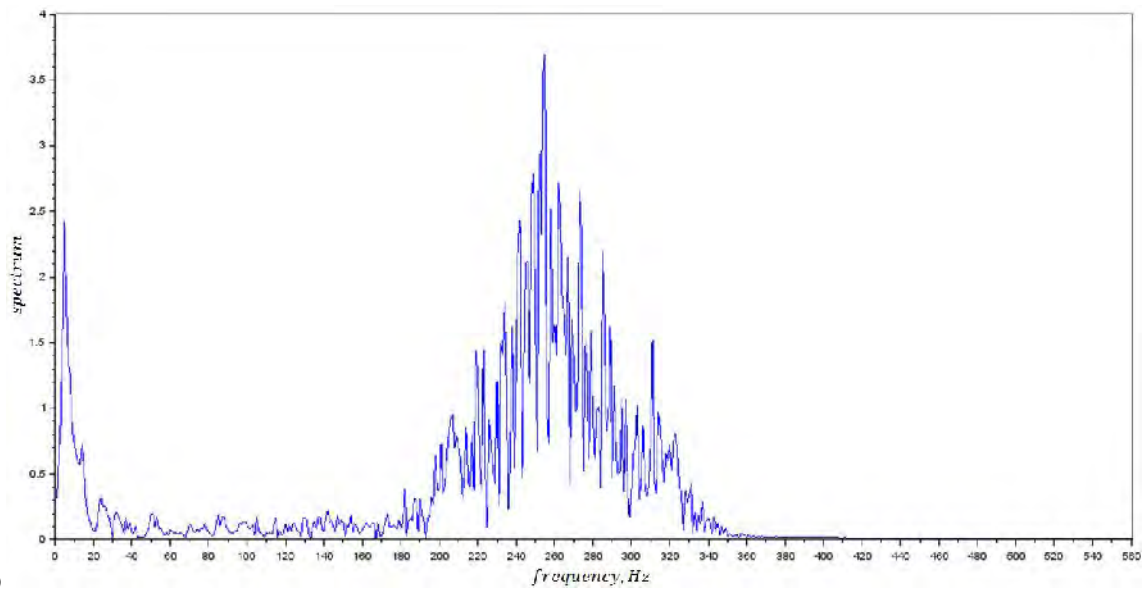
[2].

[1–5].

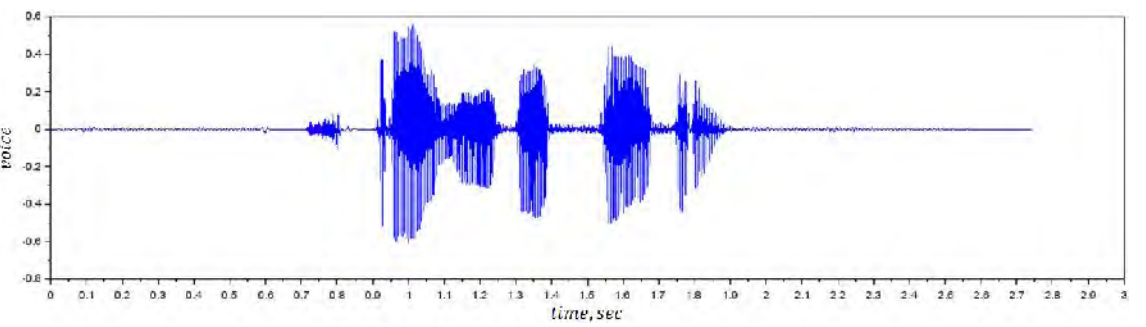
\*.wav 0,05 22050 [2, 3]. 0,05 0,25 1100 550



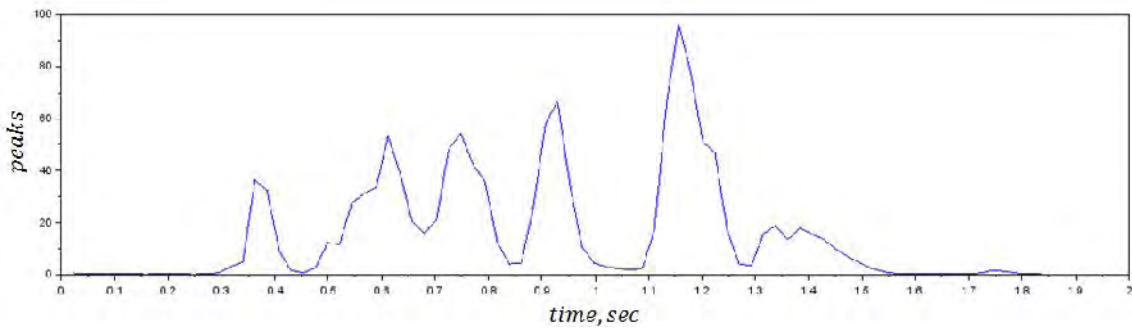
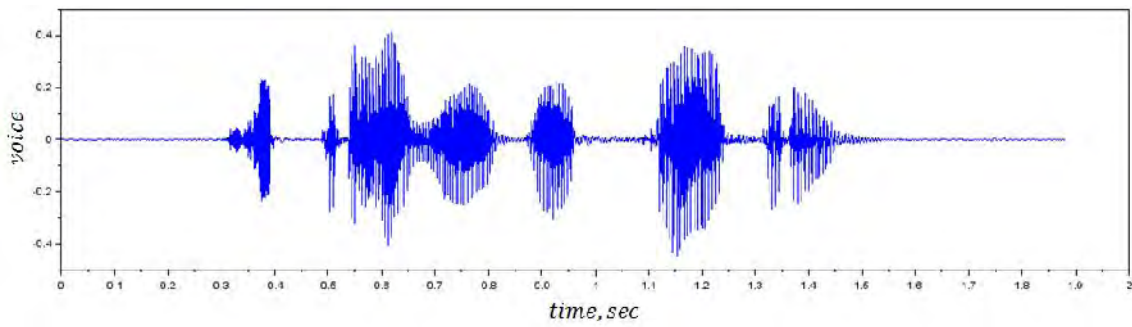
.1. (, (, (, (



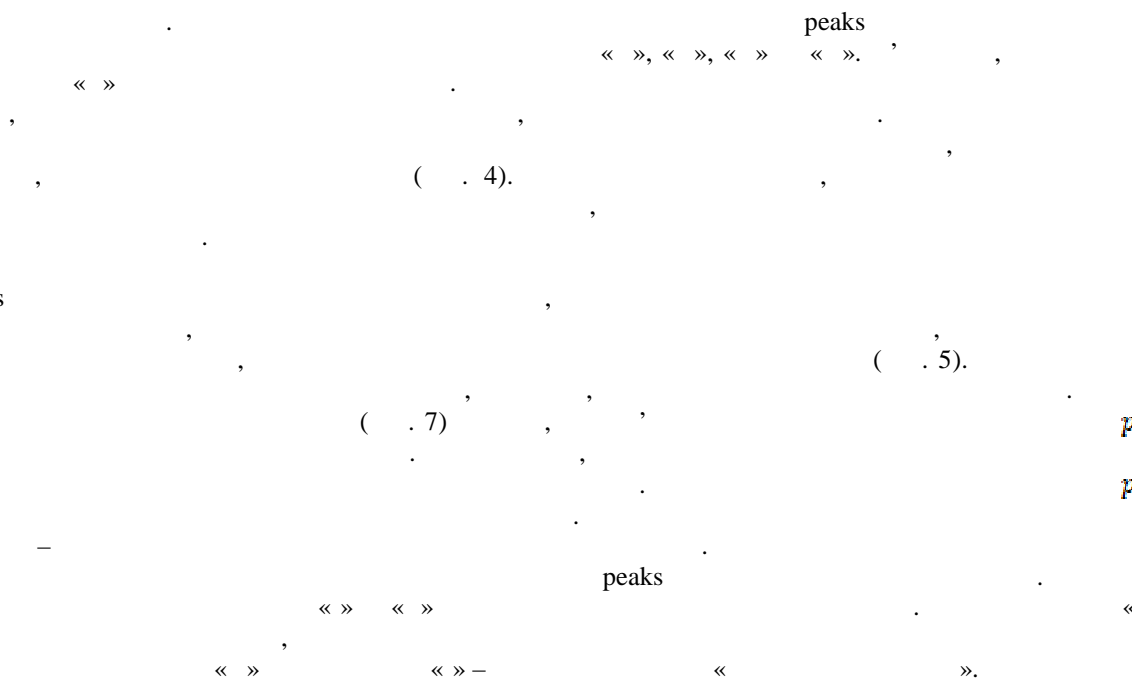
.1. ( ), ( ), ( ) ( ) ( )



.2. « »



.3.



( . 4).

( . 5).

( . 7)

peaks

peaks

peaks

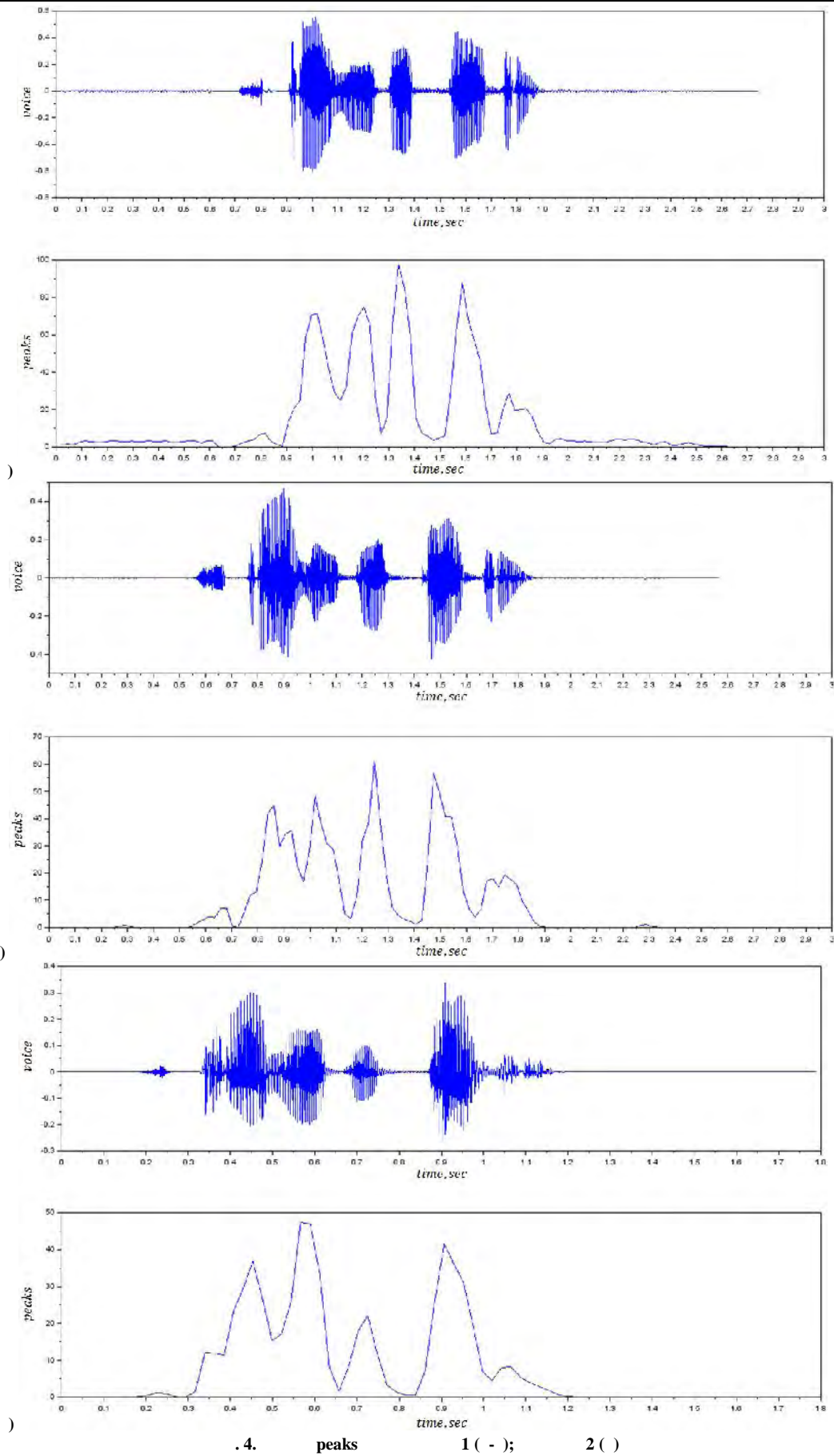
<>

<>

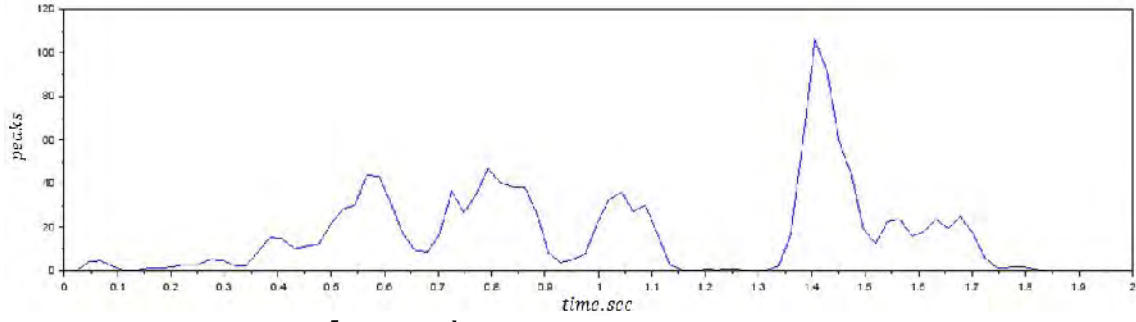
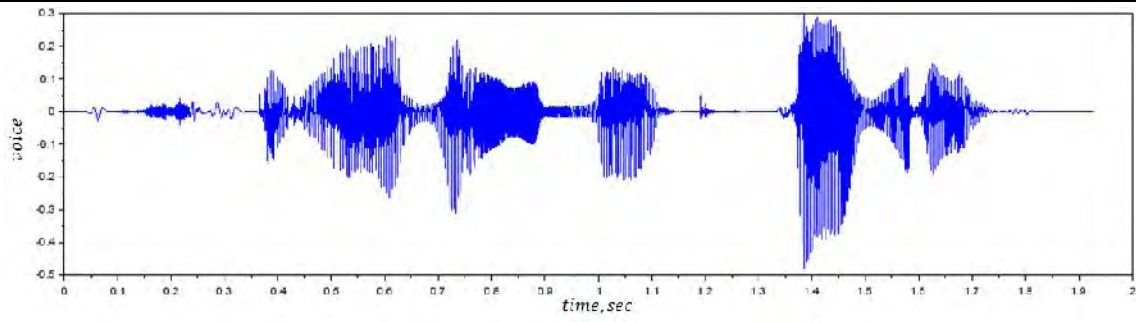
<>-

<

>.



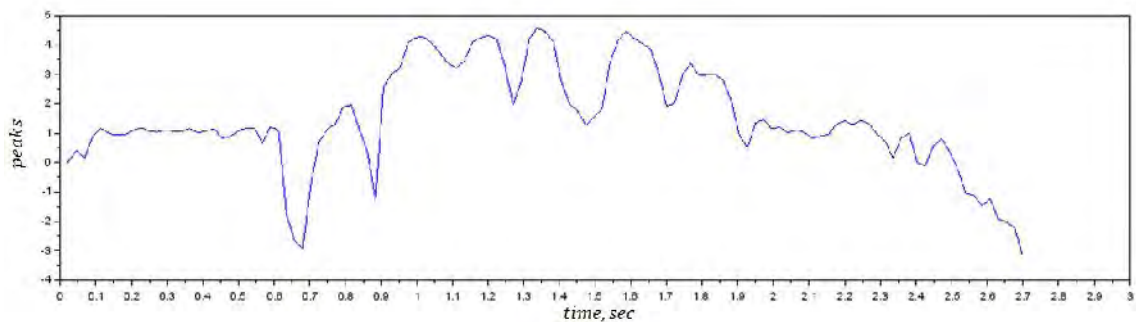
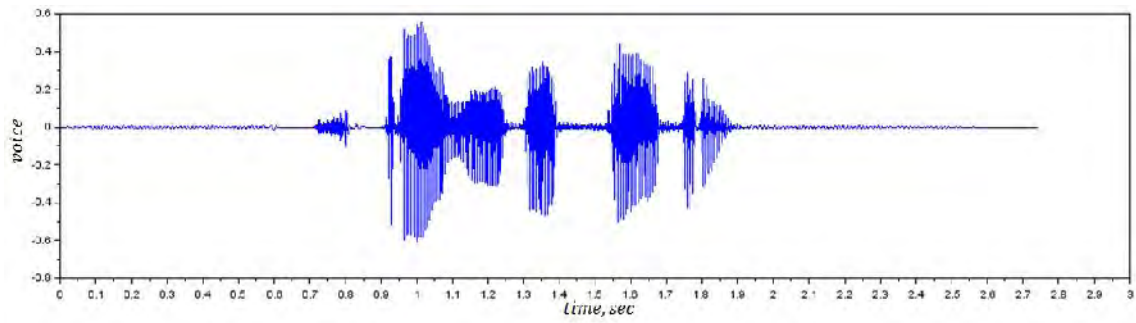
.4. peaks 1 (-); 2 ( )



.5. peaks

, — peaks.

.6-7.

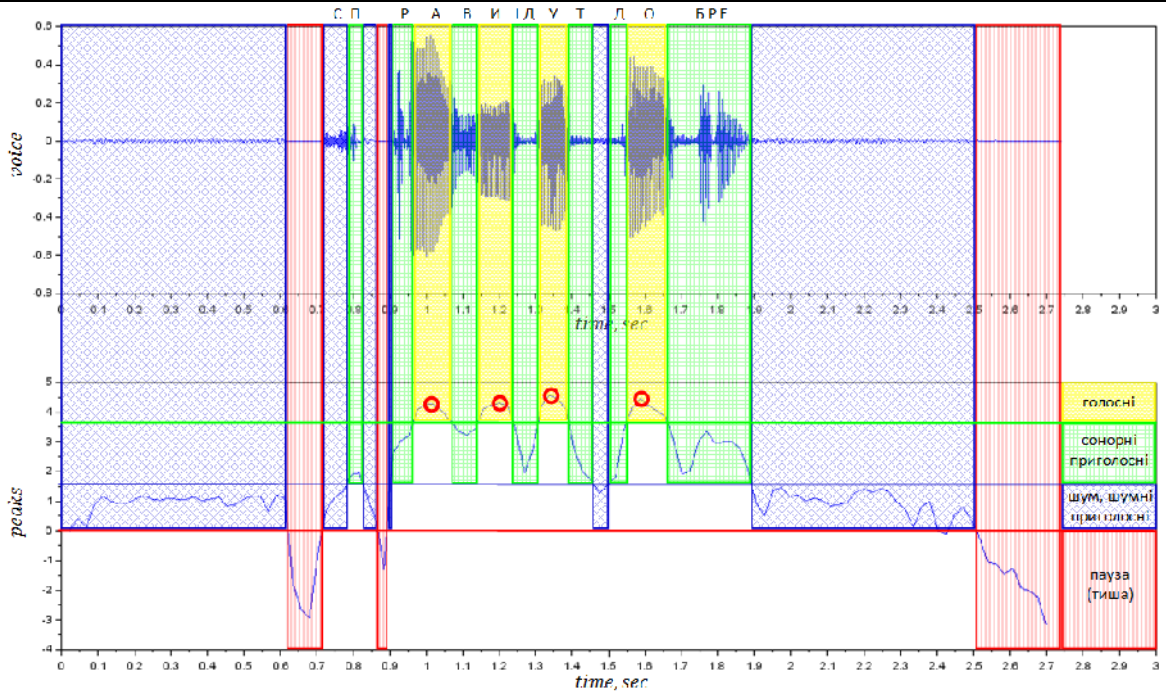


.6. peaks

— « » « » ,

.7.,



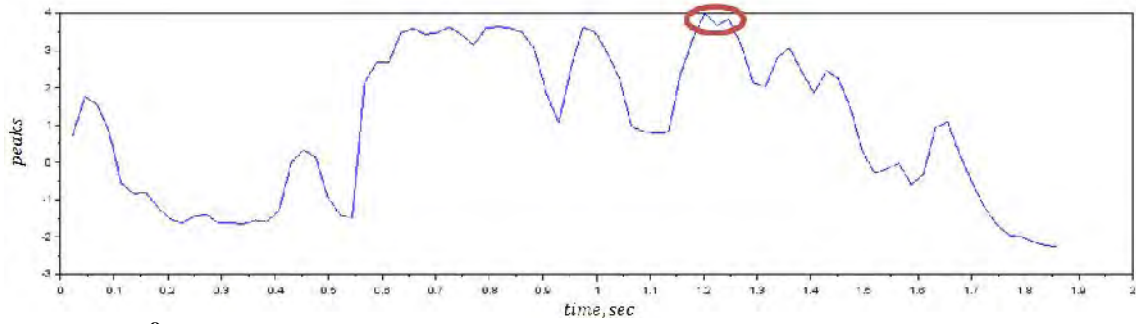
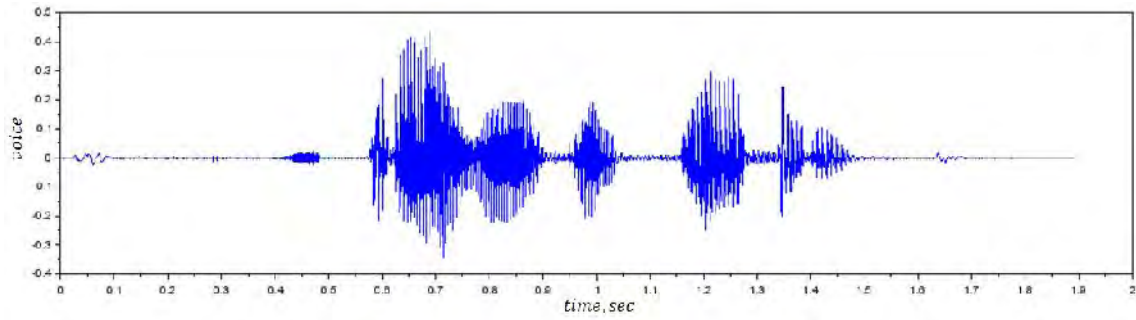


. 7.

peaks

peaks

( . 8.).



. 8.

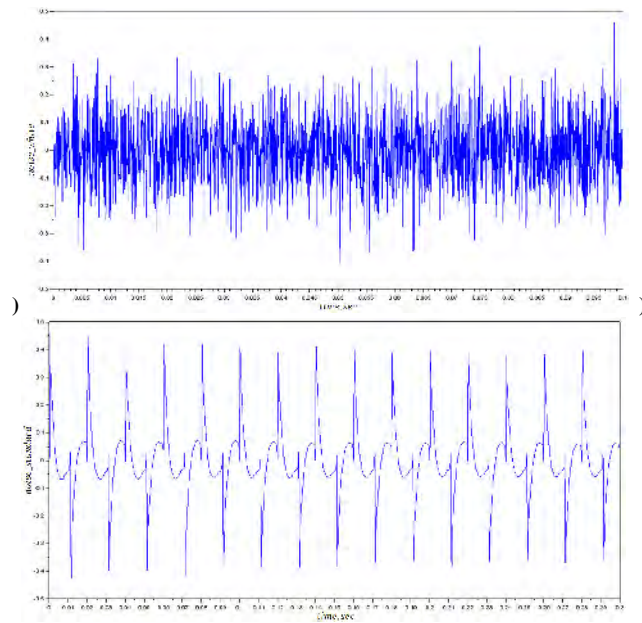
peaks.

peaks.

$$\frac{\sqrt{3}}{2}$$

0,05-0,25

5 1 5. peaks,



. 9.

« »

« »



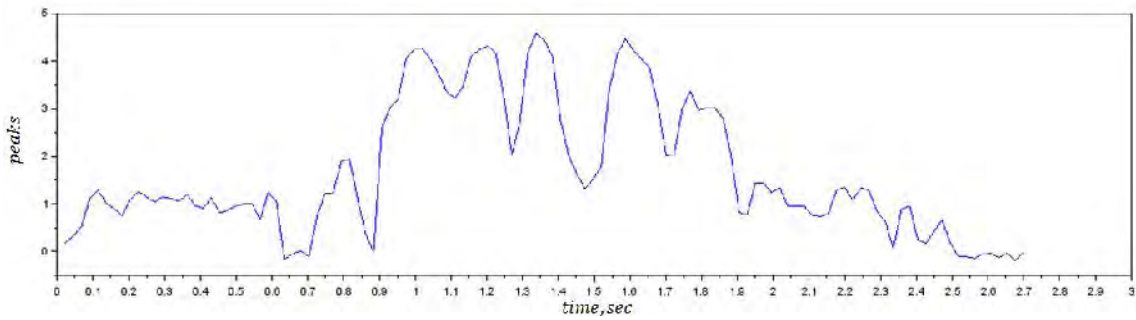
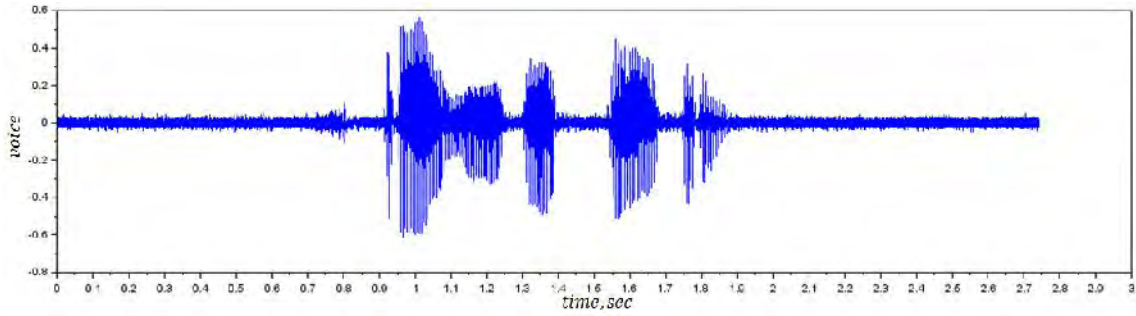
« »

50 , ( . 9).

. 10.,

« », « », « » « »

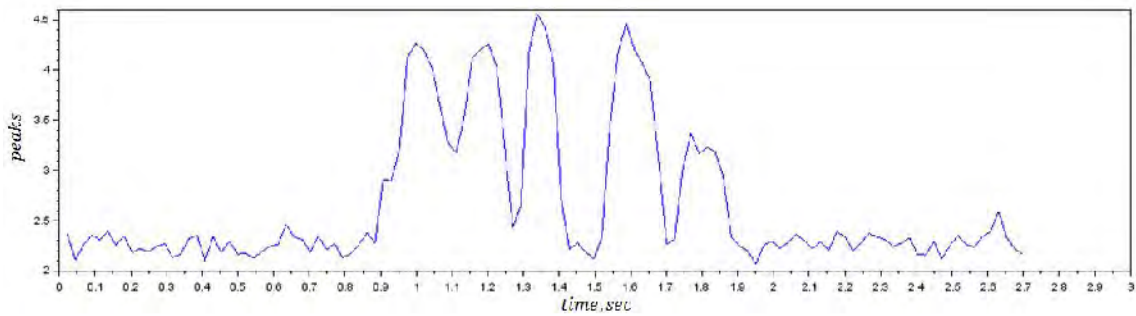
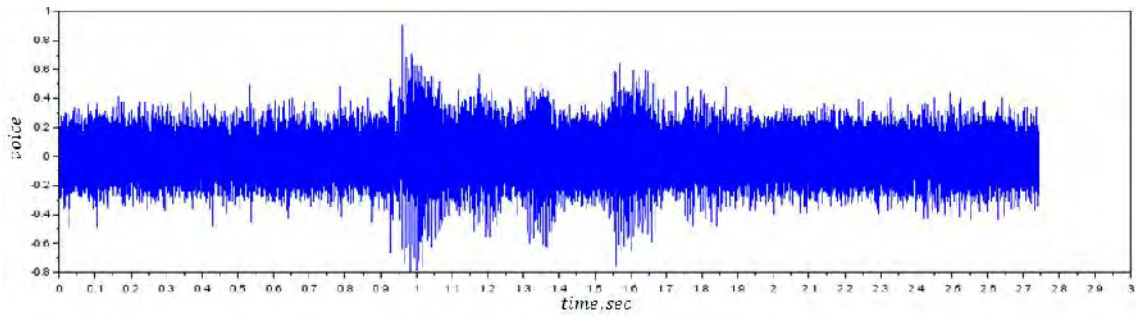
« »



)

. 10.

) 10 ; ) 20

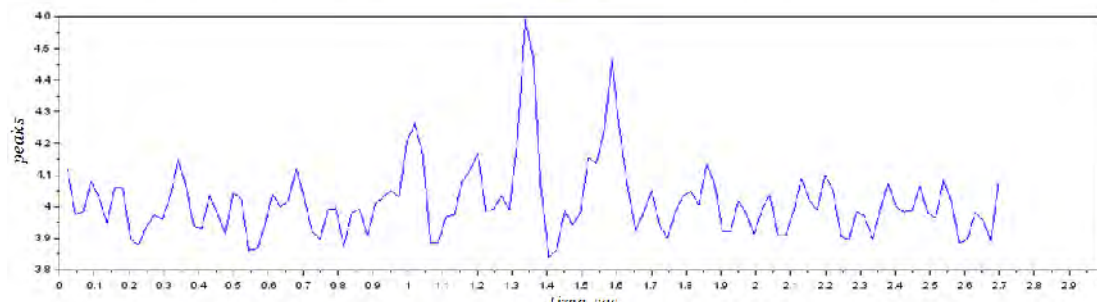
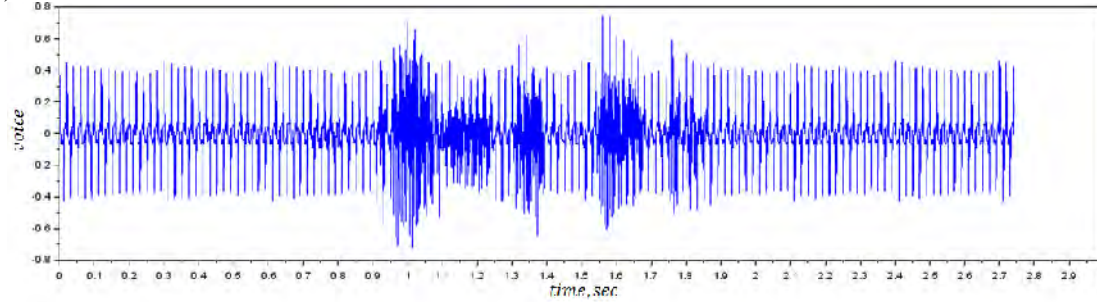
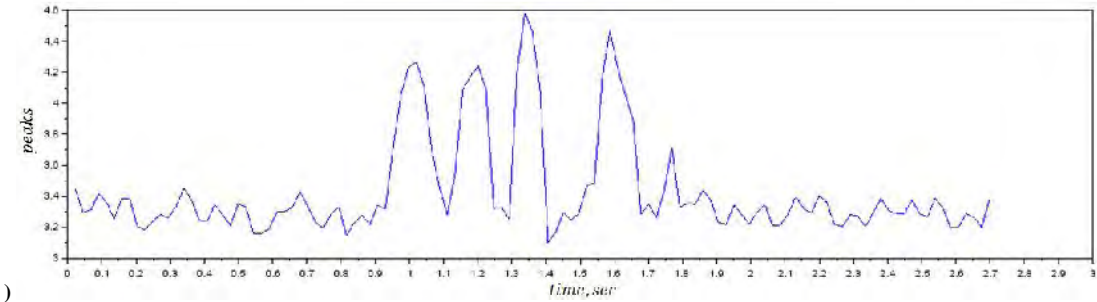
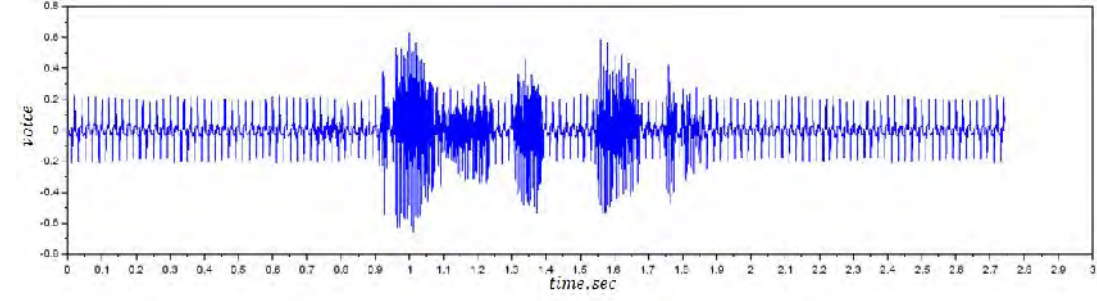
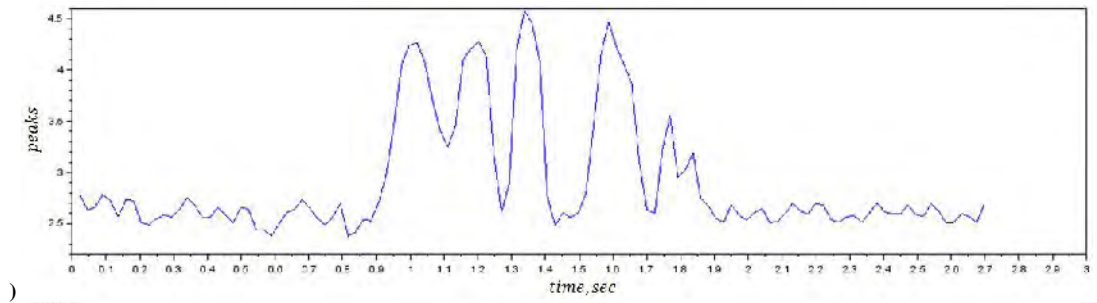
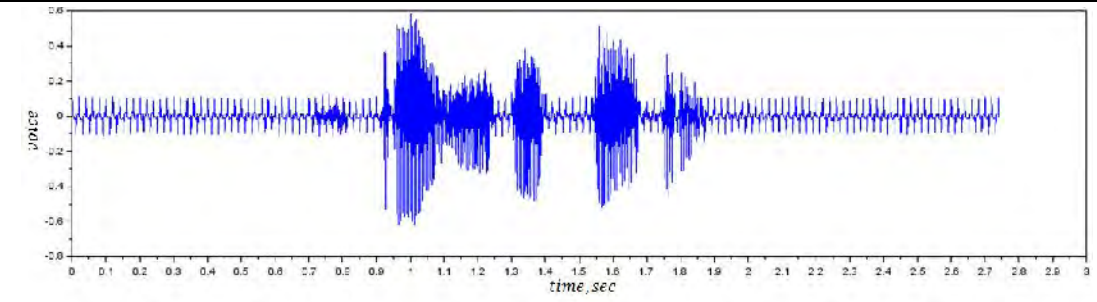


)

. 10.

) 10 ; ) 20 ( )

( . 11.)



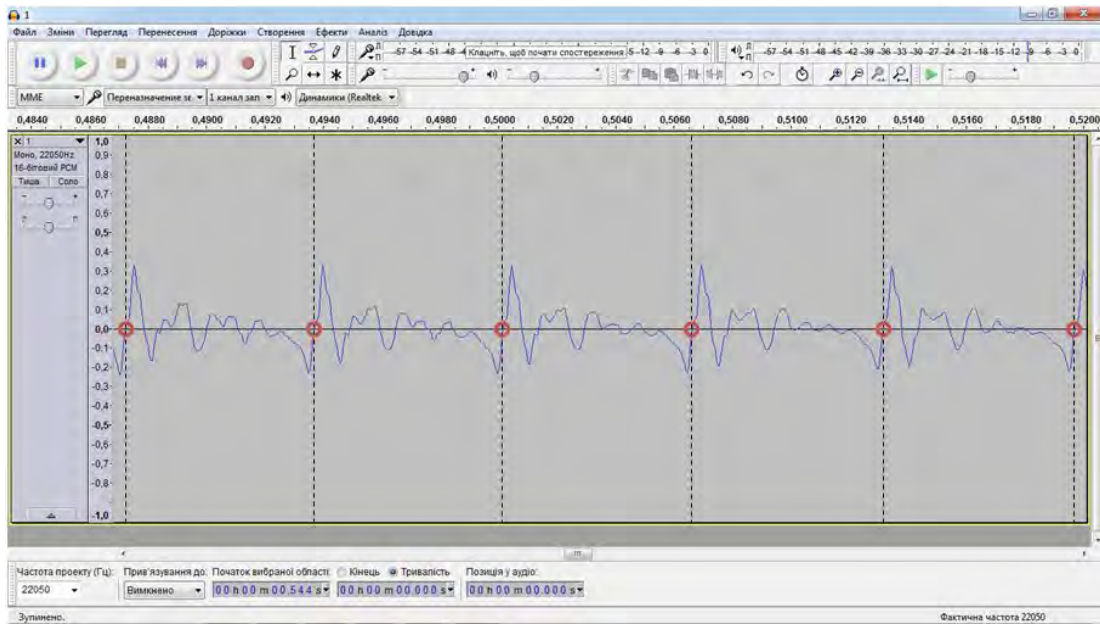
. 11.

) 10 ; ) 20 ; ) 40

[8].

( )

( . 12).



. 12.

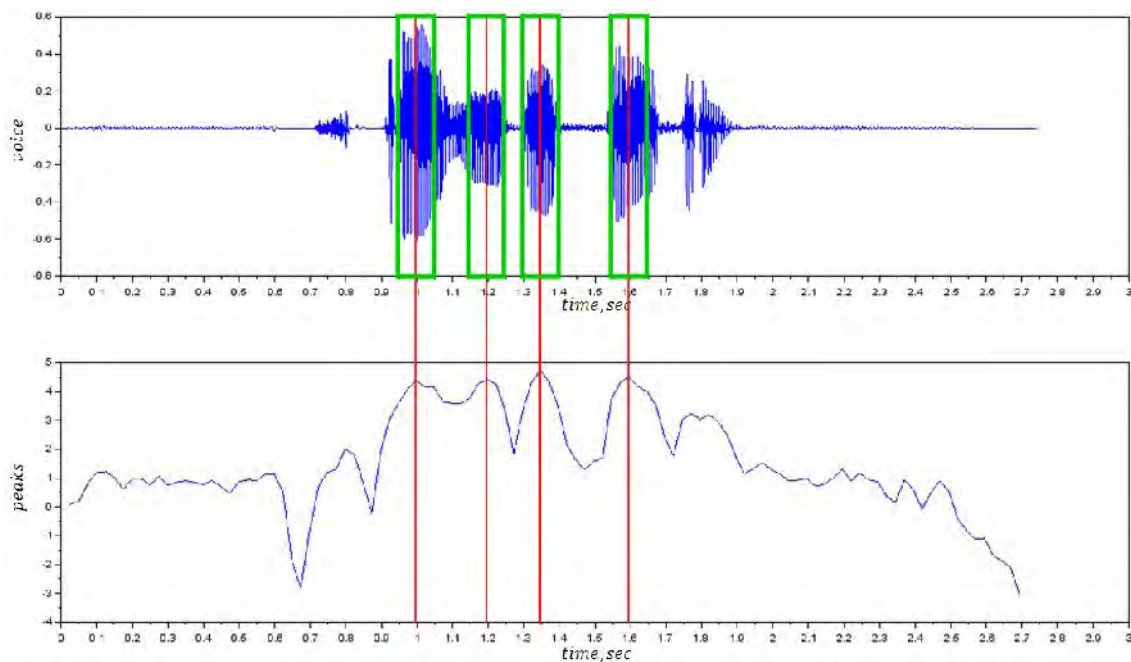
( « » )

0,1

22050

2200

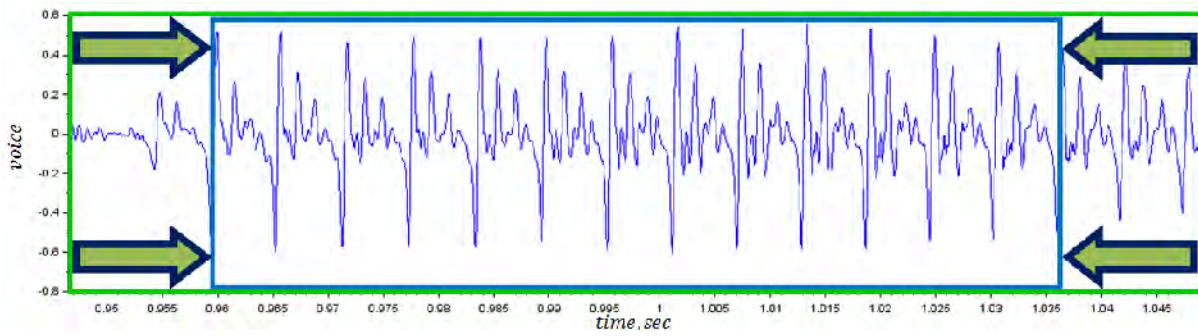
( . 13).



. 13.



( .14).



.14.

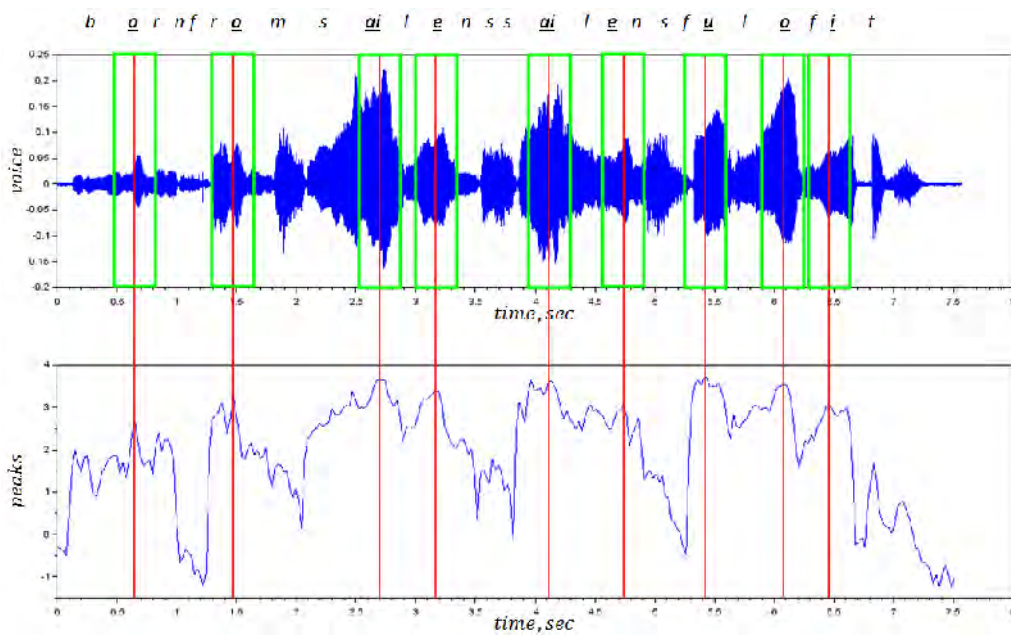
0,9

[8, 9].

« ».

( Nightwish, Metallica Avenged Sevenfold,

.15.



.15.

«Born from silence, silence full of it»

SciLab 6.0.0.

3 7,5%,  
15%,

« »,

3 7,5%,

1. // . – 2011. – 17. – . 44–50.
2. . – 2011. – 4. – . 146–152. //
3. // . – 2011. – 4 (20). – . 85–94.
4. . – 2004. – 2. 4. – . 202–220. //
5. . – 2006. – 3. 6. – . 177–207. //
6. // . – 2011. – 1-1. – . 192–196.
7. // . – 1981. – 496 .
8. Mathematical model of human voice for the task of personal identification and human condition analysis / O.Yu. Mieshkov, . . Novikov // . – 2016. – 1. – . 163–169.

9. Oleksandr Yu. Mieshkov, Oleksandr O. Novikov, Vsevolod O. Novikov, Leonid S. Fainzilberg, Andrzej Kotyra, Saule Smailova, Ainur Kozbekova, Baglan Imanbek, "Identification and human condition analysis based on the human voice analysis", Proc. SPIE 10445, Photonics Applications in Astronomy, Communications, Industry, and High Energy Physics Experiments 2017, 104453T (7 August 2017); doi: 10.1117/12.2281003

## References

1. Zhylyakov E. G., Proxorenko E. Y., Boldshev A. V., Fyrsova A. A., Fatova M. V. Segmentatsiya rechevix signalov na osnove analiza osobennostej raspredeleniya dolej energii po chastotnim yntervalam // Vestnyk NTU XPY. – 2011. – 17. – S.44-50.
2. Algoritm fonemnoj segmentatsiy na osnove analiza skorosty yzmeneniya energii diskretnogo veivlet-perobrazovaniya / O.A. Vyshnyakova, D.N. Lavrov // Vestnyk Omskogo unyversyteta. – 2011. – 4. – S. 146-152.
3. Metodyka videleniya v zvukovom signale uchastkov rechevoj aktivnosti / V.E. Gaj, V.A. Utrobyn // Yzvestiya visshyx uchebnix zavedenij. Povolzhskij region. Tekhnicheskiye nauky. – 2011. – 4 (20). – S. 85-94.
4. Segmentatsiya y raspoznaniye glasnix / V.N. Sorokyn, A.Y. Cziplyxyn // Ynformacyonnie processy, Tom 4. – 2004. – 2. – S. 202-220.
5. Segmentatsiya rechy na kardynalne element // A.Y. Cziplyxyn, V.N. Sorokyn // Ynformacyonnie processy, Tom 6. – 2006. – 3. – S. 177-207.
6. Viyavlenye poroga chuvstvitelnosti spektralnogo metoda k poysku peryodycheskogo signala v normalnom shume / G.A.Semenov // Yzvestiya AltGU. – 2011. – 1-1. – S.192-196.
7. Rabyner L.R., Shafer R.V. Cyfrovaya obrabotka rechevix signalov: Per. s angl. / Pod red. M.V.Nazarova y Yu.N.Proxorova. – M.: Radyo y svyaz, 1981. – 496 s.
8. Mathematical model of human voice for the task of personal identification and human condition analysis / O.Yu. Mieshkov, O.O. Novikov // Visnyk Xersonskogo nacionalnogo texnichnogo unyversytetu. – 2016. – 1. – S.163-169.
9. Oleksandr Yu. Mieshkov, Oleksandr O. Novikov, Vsevolod O. Novikov, Leonid S. Fainzilberg, Andrzej Kotyra, Saule Smailova, Ainur Kozbekova, Baglan Imanbek, "Identification and human condition analysis based on the human voice analysis", Proc. SPIE 10445, Photonics Applications in Astronomy, Communications, Industry, and High Energy Physics Experiments 2017, 104453T (7 August 2017); doi: 10.1117/12.2281003

/Peer review : 11.12.2017 .

/Printed :27.03.2018 .

: . . . . .