ISSN 1814-1196
Научный вестник НГТУ
том 77, № 4, 2019, с. 135–146

http://journals.nstu.ru/vestnik
Science Bulletin of the NSTU
Vol. 77, No. 4, 2019, pp. 135–146

ПРИБОРОСТРОЕНИЕ,
МЕТРОЛОГИЯ
И ИНФОРМАЦИОННО-
ИЗМЕРИТЕЛЬНЫЕ
ПРИБОРЫ И СИСТЕМЫ

INSTRUMENTATION,
METROLOGY AND
INFORMATION-
MEASURING DEVICES
AND SYSTEMS

УДК 519.24 DOI: 10.17212/1814-1196-2019-4-135-146

The analysis of the formant method of speech intelligibility estimation as a method of performing indirect measurements*

V.A. TRUSHIN

Novosibirsk State Technical University, 20 K. Marx Prospekt, Novosibirsk, 630073, Russian Federation

rastr89@mail.ru

Abstract

The paper considers the formant method proposed by N.V. Pokrovsky for determining speech intelligibility. Its essence, algorithmic and technical realization based on the information-measuring approach, as a method of making indirect measurements is analyzed. It is proposed to form a test acoustic effect based on a unified speech-like signal from syllabic or verbal articulation tables. The results of experimental studies of the synthesis of such signals, their energy spectra and probability density distribution are presented. The issues of error estimation in speech intelligibility determination are discussed: the concept of errors for different applications, an error of indirect measurements, and a theoretical error due to the conformity degree of the adopted model of speech intelligibility measurement (estimation) to the real object that is a human hearing aid. Based on the analysis of the results of the peripheral auditory system studies it is concluded that it is necessary to revise the adopted model of speech intelligibility estimation. In conclusion, the critique of the formant approach to the assessment of speech intelligibility and suggestions for their improvement are given. Based on the analysis of the results of the peripheral auditory system studies, an information-measuring model that reflects the functioning of the external ear, eardrum, and basilar membrane more adequately is proposed. The conclusion is drawn that it is necessary to revise the existing model for assessing speech intelligibility, in particular, the criteria for the reliability of its assessment, the model of the test signal and the measurement procedure.

Keywords: speech intelligibility, formant method, articulation tests, method of measurements, intelligibility estimation error, theoretical error, transverse filter, frequency response unevenness, amplitude compression

---

* *Received 16 August 2019.*

INTRODUCTION

The definition of speech intelligibility is necessary in many areas: building acoustics, medicine, psychoacoustics, speech communication systems, problems of protecting speech information, etc. To determine speech intelligibility, a lot of different methods which can be divided into two large groups: subjective (expert) and objective methods are used.

Subjective (expert) methods are based on articulation tests involving teams of speakers and listeners using articulation tables of speech elements: syllables, words, and phrases. There are quite a few varieties of articulation methods for determining speech intelligibility. In Russia, they are regulated by the standards GOST R 50840-95 [1] and GOST 16600-72 [2]. A review of foreign methods is quite fully reflected in [3, 4].

Objective methods for assessing intelligibility are based on measuring the objective parameters of a speech signal, while speakers and listeners are “replaced” by technical devices that to some extent model real speech pathways and human hearing organs (for example, an artificial voice, an artificial ear) [5].

Objective methods are divided into two large groups: formant (additive) and modulation methods. Among the formant methods, we can distinguish methods proposed by N. Pokrovsky [5], Y. Bykova [6], M. Sapozhkova [7], as well as foreign methods such as AI, SII, % ALcons and others [3, 4]. Modulation methods include STI and RASTI [3, 4]. In Russia, formant methods are based on the theory of additive contribution of various frequency lanes to total formant intelligibility. Differences between individual methods are caused by the fact that various characteristics of speech and forms of their interdependencies are selected, while the N.V. Pokrovsky method is widely used, the essence of which is described below:

- The entire frequency range of speech is divided into n-bands (in the general case n-arbitrary bands), for example, equal-articulating, octave, etc.
- For each i-th frequency band, its contribution g_i is determined to the total intelligibility of A_f formants; this contribution is estimated from the so-called formant distribution $A_f(f)$ which, in fact, is the distribution function of the formant occurrence probability in frequency (because human perception of formants has the additivity property $\sum g_i = 1$).

If the listener knowingly accepted everything that the speaker conveys, i.e. if there were an ideal channel (the source of speech information is a receiver), then regardless of the number of bands and their width, formant intelligibility would be 1. However, in real conditions, a part of the formants is not perceived for a number of reasons such as an insufficient volume, distortions in the path, noises, interference, etc., therefore we always have $A_f \leq 1$. This circumstance is taken into account by the formant perception coefficient *P*; in fact, *P* is the probability of the correct reception of formants $0 \leq P \leq 1$.

The perception coefficient is a function of the sensation level of formants, i.e. the number of formants whose intensity is above a certain threshold value.

For most practical cases, the level of sensation E_i is determined by the formula:

$$E_i = B_{fi} - B_{ni} = (B_{si} - B_{ni}) - \Delta B_i, \tag{1}$$

where B_{si}, B_{fi}, B_{ni} are, spectral levels of speech, formant, and noise respectively in the *i*-th frequency band; ΔB_i is the difference between the speech spectrum and formants.

Thus, total formant intelligibility is determined by the expression:

$$A_i = \sum_{i=1}^{n} A_{fi} = \sum_{i=1}^{n} g_i P_i. \tag{2}$$

Next, the transition from a formant to any other types of intelligibility (*D*, *S*, *W*, *I*) is carried out according to the dependencies known for a certain language [5, 9].

The use of this method for solving specific practical problems is regulated by appropriate methods; for example, GOST 8031 is used for assessing the quality of communication paths [8]; the methodology proposed in [9] is used for assessing the security of speech information from leakage through technical channels.

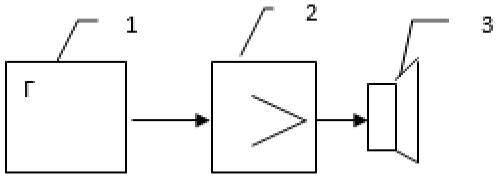
1. THE N.V. POKROVSKY FORMANT METHOD AS A METHOD FOR TAKING MEASUREMENTS

Essentially, these methods are indirect measurement methods in which the final measurement goal is obtained based on direct measurements of certain quantities and the known (accepted) dependencies of these quantities for the purpose of measurement. In this case, we mean direct measurements of noise levels and signal levels + noise" i in the received frequency bands f_i. Then, the levels B_{si}, the ratios (signal / noise) i, the sensation levels of the formants E_i, the perception coefficient p$_i$, the formant intelligibility A_f and the transition according to the known dependencies on the A_{fs} verbal intelligibility W [5, 9] are calculated.

Strictly speaking, it is more correct to speak about evaluation, not about measurement, because the measured value – intelligibility – is dimensionless, but this does not change the essence.

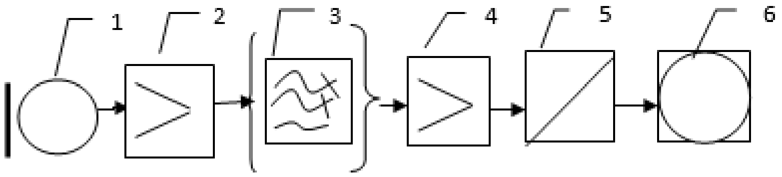
Below is a generalized structural diagram of the implementation of the N.V. Pokrovsky method based on an analog sound level meter [10] (Fig. 1). This diagram is called generalized, because it will be “filled” with various specifics for various applications. Therefore, for the tasks of assessing the security of speech information from leakage through technical channels, measuring microphones and accelerometers are used as primary converters; seven-octave filters with center frequencies of 125, 250, 500, 1000, 2000, 4000.8000 Hz, etc. are used.

As is known, the measurement procedure (method) is a set of specifically described operations, the implementation of which ensures the receipt of measurement results with established accuracy indicators [11].



1 – a tone or white noise generator
2 – an amplifier
3 – an acoustic emitter

a



1 – a primary converter
2–4 – an amplifiers
3 – a set of bandpass filters
5 – an RMS detector
6 – a display device with logarithmic calibration

b

Fig. 1. A generalized block diagram of formant method implementation:
a – a test signal generation device; *b* – measuring part

Рис. 1. Обобщенная структурная схема реализации формантного метода:
а – устройство формирования тестового сигнала; *б* – измерительная часть

According to this formulation and Fig. 1 in the N.V. Pokrovsky formant method the following operations are implemented:

1. Conversion of an acoustic pressure into an equivalent electrical signal (primary transducers).
2. Linear amplification of the electrical signal (amplifiers).
3. Band-pass filtering in octave, third-octave or equal-articulation bands (a set of band-pass filters).
4. Converting the “filtered” signals into average values of their squares (intensity) Ii.
5. A logarithm of the obtained values.

Thus, at the output of the measuring part, a set of values of spectral noise and “signal + noise” levels is formed in each frequency band:

$$B_{ni} = 10\,\lg\frac{I_{ni}}{10} B_{s+n} = 10\,\lg\frac{I_{(s+n)i}}{I_0}, \tag{3}$$

where

$$I_i = \int_o^T p^2(t, fi)\,dt\,. \tag{4}$$

Then, computational operations are carried out to determine the formant intelligibility of A_f and the transition to verbal intelligibility in accordance with the known dependencies [5, 9].

2. THE TEST SIGNAL MODEL

In the adopted technique, either tonal signals corresponding to the middle of the octave bands are used as a test signal (Fig. 1), or a white noise with a normal probability distribution of values is used.

Such an approximation, apparently, is possible when dividing the frequency range into a large number of bands (for example, equally articulatory), which is known as completely inadequate to the real speech signal.

The most natural is the use of pre-recorded real speech signals with a duration of at least the planned time of negotiations; ideally, the voices of the “owners” of the protected premises (for the tasks of protecting information.). However, this approach does not allow certification (if necessary) of such UFTS. In this regard, a unified approach to creating a test signal, in particular formed on the basis of syllabic or verbal tables seems appropriate in accordance with GOST [1, 2]. As an example of such test signals, Fig. 2 and 3 show some results of experimental studies carried out under the guidance and with the participation of the author [12].

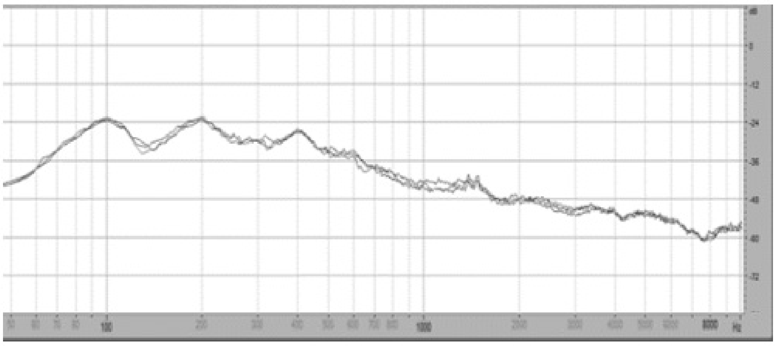
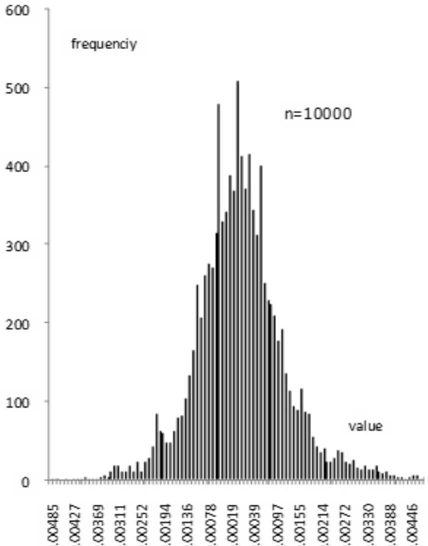


Fig. 2. The energy spectra of a speech choir based on syllables, words, and connected texts (two male voices and one female)

Рис. 2. Энергетические спектры речевого хора на основе слогов, слов и связных текстов (два мужских голоса и один женский)

Fig. 3. The frequency histogram constructed on 10 000 values of a “word” signal

Рис. 3. Частотная гистограмма, построенная по 10 000 значений сигнала «слова»



It can be seen from the figure that the spectra obtained practically coincide and correspond to the average spectrum of Russian speech.

An important statistical characteristic of speech is also the probability distribution density of its values, which is most often approximated by Laplace distributions (double exponential) or a third-order polynomial in a system of exponential functions [13, 14]. Figure 2 shows an experimental probability density of the RP signal (syllables) values, constructed from 10,000 reports with a sampling frequency of 44 kHz. Visual resemblance to the Laplace distribution is obvious.

3. ON ERRORS OF INTELLIGIBILITY ASSESSMENT

From the point of view of information-measuring technologies, the application of the N.V. Pokrovsky formant method and methods that implement it raises a number of questions, and especially the following ones.

1. In accordance with the definition of the term “measurement technique” [17], what will an estimation error W be and what should it be?
2. The degree of conformity of the measurement methodology used to the real object due to the accepted assumptions and simplifications (a theoretical error).

The answers to these questions, of course, require a separate detailed discussion, but in general terms they can be formulated as follows.

As you know, a measurement error is understood to mean the deviation of the measured value of a quantity from its true value. In practice, of course, instead of a true value (which is unknown), an actual value of the measured quantity, i.e. obtained with a higher accuracy is used. Obviously, only the result of properly organized, representative articulation tests can be taken as a real value of W, since the very concept of W has an exclusively expert meaning. Depending on practical applications and goals of articulation tests, they should be organized in different ways. So, to assess speech intelligibility in auditoriums (concert, theater, or cinema halls) according to GOST 25902-83 [8], syllabic speech intelligibility is estimated with gradation in 3 classes:

I. Excellent intelligibility (over 90 %), II. Good intelligibility (80-90 %), III. Satisfactory intelligibility (70…80 %). In this case, spatial curves of equal intelligibility are constructed. To assess speech intelligibility in information protection tasks (assessment of the security of speech information from leakage through technical channels), W is used with the following gradations [9, 15]:

– concealment of negotiation facts in the allocated premises, with W being no more than 10 %;

– hiding the negotiation subject, with $W \leq 20$ %;

– hiding the negotiation content with $W \leq 30$ %;

– an impossibility of compiling a brief summary-annotation of the negotiation content, with $W \leq 40$ %.

– an impossibility of drawing up a detailed certificate of the negotiation content, with $W \leq 60$ %.

Obviously, the inaccuracy of the intelligibility assessment in the first case can lead to a change in class and, consequently, a comfortable perception by listeners. In the second case, errors in the estimation of W can lead to a speech information leakage with a wide variety of consequences. However it can be seen from the foregoing that a required error in the assessment of intelligibility is of the order of percentage units in the first case for syllabic intelligibility with its high values (0.7…0.9), and in the second case for verbal intelligibility with its rather small values (0.1…0.5).

The N.V. Pokrovsky concept of equally accurate measurements, i.e. actually "tied" to the accuracy of formant intelligibility is introduced.

"If measurements of any kind are equally accurate, independent and free from systematic errors and gross errors, then the number of measurements n for a given absolute error ε is determined by the dependence that is valid for each type" [5, p. 99]

$$n = f\left(\frac{\varepsilon}{r}\right), \tag{5}$$

where "r" is a standard deviation.

In this case, a single measurement is a table, i.e. n is a number of tables to be transmitted along the communication path.

Now let us consider the "measured", i.e. obtained using the above methodology, value of W and the diagram that implements it (Fig. 1).

The instrumental component of the estimation error W is determined by the metrological characteristics of the technical equipment used, as shown in Fig. 1, the requirements for which are obvious from the general provisions of information-measuring technologies. A methodological error of the estimation of W is determined according to the recommendation of MI 2083-90 [18]. A methodological error in determining W was estimated in [17, 18]. It is shown that an absolute error of indirect measurements of W can be 0.07, and a relative error is 0.35.

4. ON THE ORGANIZATION OF ARTICULATORY TESTS

Regarding the organization of articulation tests, it is important to note the following circumstance. Research conducted by N.B. Pokrovsky was primarily associated with assessing the quality of means and communication channels to obtain their objective characteristics. For this, it is extremely important that the actuation tables are not correlated (sounds, syllables, words), while in a number of practical applications (for example, in information protection tasks) negotiations and conversations i.e. connected, meaningful "texts" take place. You can restore a missed word by context, and if it is possible to record conversations, listen to them repeatedly and filter, in this case intelligibility is significantly increased (the use of modern noise reduction methods increases the signal-to-noise ratio to 8 dB depending on the frequency band) [15]. Therefore, articulation tests with coherent texts [18] should be conducted for these purposes.

5. ON CONFORMITY OF THE ACCEPTED MEASUREMENT MODEL TO A REAL OBJECT

Obviously, a human hearing system, a degree of compliance with which should be evaluated is a real object. Studies of human hearing organs have been conducted for more than a century, and many articles, monographs, patents, and dissertations have been devoted to them. Nevertheless, a number of conclusions can be singled out that are directly related to the issue under consideration. As you know, the peripheral part of the human auditory system consists of the outer, middle and inner ear.

The outer ear organs (the auricle and the external auditory canal) perform the functions of perception and amplification of sound waves, but also have their own resonant frequencies in the region of 3…4 kHz with a gain of 10–12 dB [20, 21].

The eardrum is located at an angle to the ear canal and has asymmetric elasticity, which leads, according to the Helmholtz theory, to the formation of combination frequencies. In addition, "the presence of a special mechanism controlling the sensitivity of the ear in the form of a muscle stretching the eardrum" [20–22] weakens a high level of sound pressure by 10…30 dB depending on the frequency, i.e. protects the ear from overload. At the same time, it is generally accepted that transformations in the outer and middle ear are linear up to sound pressure levels of the order of 80 dB [4, 22, 23], and nonlinearities appear only in the inner ear.

At the same time, in a number of papers opposite opinions are expressed about the place of occurrence of nonlinearities. Therefore, their occurrence in structures preceding the filtration mechanism is described, with the form of nonlinearity being a logarithmic or cubic root [22].

Modern methods of non-contact measurements of displacements of the nanometer range have revealed a logarithmic dependence of the movements of the eardrum (in particular in the Umbo region) on the sound pressure amplitude on its outer side in the range of ultrasound of 30…80 dB. In this case, the displacement reaches 60…250 nanometers depending on the frequency [24, 25].

Filtration of speech signals in the cochlea of the inner ear is carried out by not octave or equiarticulatory etc. bands, but by critical bands, the number of which is 24 [4]. In addition, band-pass filtering is carried out in a tonotopic mode (frequencies are "tied" to a specific location on the basilar membrane) from high-frequency to low-frequency; this explains a greater masking effect of the high-frequency part of the low-frequency spectrum. Apparently, the closest analogue to such a filter is transversal [26], which agrees well with G. Bekesy's traveling wave theory. In addition, the presence of nonlinear compression on the basilar membrane carried out by external hair cells is considered a fact [4, 21, and 23].

An analysis of papers devoted to studies of the peripheral auditory system, presented in a concise form, allows us to conclude that the accepted measurement model does not correspond to real processes in the human auditory system and to propose a more adequate model (Fig. 4).

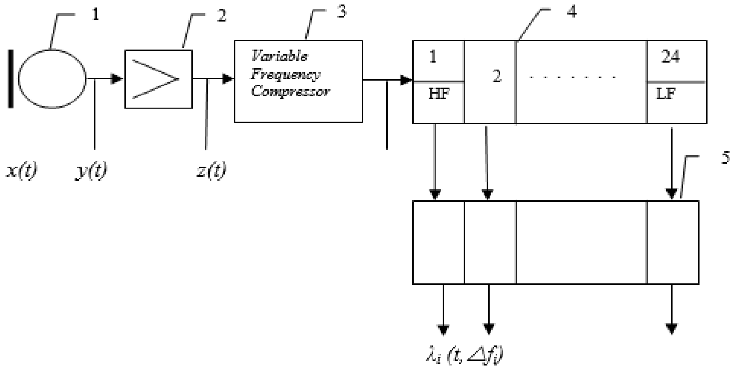


Fig. 4. A model of peripheral auditory information system for intelligibility measurements:

1 – a primary converter; *2* – an amplifier with an uneven frequency response (rise by 12…15 dB in the region of 3…4 kHz); *3* – a frequency-dependent amplitude compressor (a quasi-logarithmic or cubic root); *4* – a transverse filter with bands corresponding to critical hearing bands; *5* – an integrator

Рис. 4. Информационно-измерительная модель периферической слуховой системы для определения разборчивости:

1 – первичный преобразователь; *2* – усилитель с неравномерной АЧХ (подъем на 12…15 дБ в районе 3…4 кГц); *3* – частотозависимый амплитудный компрессор (квазилогарифмический или корень кубический); *4* – трансверсальный фильтр с полосами, соответствующими критическим полосам слуха; *5* – интегратор

In Fig. 4 $y(t) = k_1 f\, x(t)$; $z(t) = k_2 f\, F[y(t)]$, $\lambda_i = \int z(t,\ \Delta f_i)\, dt$, where F is the compression function.

The proposed model, of course, requires experimental validation and determination of its numerical parameters.

CONCLUSION

Quite a number of publications have been devoted to criticizing formant methods for assessing speech intelligibility, the results of which are reflected in the most generalized form in [27, 28, and 29]. The above analysis allows us to draw the following conclusions:

– results of correctly organized representative articulation tests (with coherent texts taking into account the effect of speech forcing, etc.) should be used as criteria for the reliability of speech intelligibility assessments;

– test signals should take into account real processes of speech formation;

– measurement methods should be based on an adequate model of auditory perception.

REFERENCES

1. *GOST R 50840–95. Peredacha rechi po traktam svyazi. Metody otsenki kachestva, razborchivosti i uznavaemosti* [State Standard R 50840–95. Speech transmission over varies communication channels. Techniques for measurements of speech quality, intelligibility and voice identification]. Moscow, Gosstandart Rossii Publ., 1996. 229 p.

2. *GOST 16600–72. Peredacha rechi po traktam radiotelefonnoi linii. Trebovaniya k razborchivosti rechi i metody artikulyatsionnykh ispytanii* [State Standard 16600–72. Speech transmission over the radiocommunication. Requirements for speech legibility and methods of articulation measurements]. Moscow, Standards Publ., 1972. 90 p.

3. Rashevskii Ya.I., Kargashin V.L. Obzor zarubezhnykh metodov opredeleniya razborchivosti rechi [Review of foreign methods for determining speech intelligibility]. *Spetsial'naya tekhnika – Special Equipment*, 2002, no. 4, pp. 37–46.

4. Aldoshina I.A., Pritts R. *Muzykal'naya akustika* [Musical acoustics]. St. Petersburg, Kompozitor Publ., 2006. 760 p. (In Russian).

5. Pokrovsky N.B. *Raschet i izmerenie razborchivosti rechi* [Calculation and measurement of speech intelligibility]. Moscow, Svyaz'izdat Publ., 1962. 392 p.

6. Bykov Yu.S. *Teoriya razborchivosti rechi i povysheniya effektivnosti radiotelefonnoi svyazi* [Theory of legibility and increasing the effectiveness of radiotelephone communications]. Moscow, Gosenergoizdat Publ., 1959. 351 p.

7. Sapozhkov M.A. *Rechevoi signal v kibernetike i svyazi* [The speech signal in cybernetics and communications]. Moscow, Svyaz'izdat Publ., 1963. 452 p.

8. *GOST 25902–83. Zritel'nye zaly. Metod opredeleniya razborchivosti rechi* [State Standard 25902-83. Auditoria. Speech intelligibility determination method]. Moscow, Standards Publ., 1984. 10 p.

9. Zheleznyak V.K., Makarov Yu.K., Khorev A.A. Nekotorye metodicheskie podkhody k otsenke effektivnosti zashchity rechevoi informatsii [Some methodological approaches to assessing the effectiveness of voice information protection]. *Spetsial'naya tekhnika – Special Equipment*, 2000, no. 4, pp. 39–45.

10. *GOST 17187–2010. Shumomery*. Ch. 1. *Tekhnicheskie trebovaniya* [State Standard 17187–2010. Sound level meters. Part 1. Technical requirements]. Moscow, Standartinform Publ., 2010. 31 p.

11. *GOST 8.563–2009. Metodiki (metody) izmerenii* [State Standard 8.563-2009. Procedures of measurements]. Moscow, Standartinform Publ., 2010. 15 p.

12. Trushin V.A., Khitsenko V.E. About the methods of forming a test signal in the instrumental evaluation of speaker clearance. *2019 Dynamics of Sistems, Mechanisms and Machines (Dinamics)*, Omsk, 2019.

13. Vykhodets V., ed. *Spravochnik po radioveshchaniyu* [Handbook to broadcasting]. Kiev, Tekhnika Publ., 1981.

14. Kropotov Yu.A. Model' odnomernoi plotnosti veroyatnosti rechevykh signalov [Model-dimensional probability density function of the speech signal]. *Sistemy upravleniya, svyazi i bezopasnosti – Systems of Control, Communication and Security*, 2015, no. 4, pp. 158–170.

15. Dvoryankin S.V., Makarov Yu.K., Khorev A.A. Obosnovanie kriteriev effektivnosti zashchity rechevoi informatsii ot utechki po tekhnicheskim kanalam [Substantiation of criteria for the effectiveness of protection of speech information from leakage through technical channels]. *Zashchita informatsii. Insaid*, 2007, no. 2, pp. 18–25. (In Russian).

16. Kozlachkov S.B., Dvoryankin S.V., Bonch-Bruevich A.M. Printsipy formirovaniya testovykh rechevykh signalov pri otsenkakh effektivnosti tekhnologii shumoochistki [Principles of forming test speech signals in assessing the efficiency of noise reduction technologies]. *Voprosy kiberbezopasnosti – Cybersecurity issues*, 2018, no. 3 (27), pp. 9–16.

17. *MI 2083–90. Izmereniya kosvennye. Opredelenie rezul'tatov izmerenii i otsenka ikh pogreshnostei* [Recommendations about metrology 2083–90. Indirect measurements. Determination of measurement results and assessment of their errors]. Moscow, Standartinform Publ., 1991. 9 p.

18. Trushin V.A., Ivanov A.V., Reva I.L. O metodicheskikh pogreshnostyakh otsenki slovesnoi razborchivosti rechi v zadachakh zashchity informatsii [On the methodological errors in evaluating verbal speech intelligibility in information protection problems]. *Doklady Tomskogo gosudarstvennogo universiteta sistem upravleniya i radioelektroniki – Proceedings of Tomsk State University of Control Systems and Radioelectronics*, 2012, no. 1 (25), pt. 2, pp. 180–185.

19. Trushin V.A., Ivanov A.V., Khitsenko V.E. O vybore modeli testovogo signala pri otsenke zashchishchennosti rechevoi informatsii ot utechki po tekhnicheskim kanalam [On the selection of a test signal model for assessing the protection of speech information from leakage through technical channels]. *Trudy SPIIRAN – SPIIRAS Proceedings*, 2015, no. 3, pp. 122–133.

20. Schiffman H.R. *Sensation and perception: an integrated approach*. New York, Wiley, 2001 (Russ. ed.: Shiffman Kh.R. *Oshchushchenie i vospriyatie*. 5th ed. St. Petersburg, Piter Publ., 2003. 928 p.).

21. Altman Ya.A., Tavartkiladze G.A. *Rukovodstvo po audiologii* [Handbook on audiology]. Moscow, DMK-Press Publ., 2003. 360 p.

22. Labutin V.K., Molchanov A.P. *Modeli mekhanizmov slukha* [Models of hearing mechanisms]. Moscow, Energiya Publ., 1973. 200 p.

23. Chistovich L.A., Ventsov A.V., Granstrem M.P. et al. *Fiziologiya rechi. Vospriyatie rechi chelovekom* [Physiology of Speech. The perception of speech by a person]. Leningrad, Nauka Publ., 1976. 388 p.

24. Mareev G.O. *Vozmozhnosti lazernogo avtodinnogo izmeritelya nanosmeshchenii v diagnostike zabolevanii ukha*. Avtoref. diss. dokt. med. nauk [Possibilities of a laser autodyne measuring device of nanodisplacements in the diagnosis of ear diseases. Author's abstract of Dr. medical sci. diss.]. Saratov, 2013. 22 p.

25. Usanov D.A., Mareev O.V., Skripal' A.V., Mareev G.O. Lazernye avtodinnye izmereniya parametrov dvizheniya barabannoi pereponki [Laser autodyne measurements of the parameters of the movement of the eardrum]. *Rossiiskii zhurnal biomekhaniki – Russian Journal of Biomechanics*, 2012, vol. 16, no. 1, pp. 8–21. (In Russian).

26. Vargin P.S. *Matematicheskaya model' ulitki vnutrennego ukha* [Mathematical model of the cochlea of the inner ear]. Available at: http://www.telenavt.ru/downloads/Ulitka%20vnutrennego%20uha%20cheloveka.pdf (accessed 16.12.2019).

27. Avdeev V.B. O nekotorykh napravleniyakh sovershenstvovaniya metodicheskikh podkhodov, primenyaemykh pri otsenke effektivnosti tekhnicheskoi zashchity informatsii [About some areas of improving methodological approaches used in assessing the effectiveness of technical information protection]. *Spetsial'naya tekhnika – Special Equipment*, 2013, no. 62, pp. 1–10.

28. Kozlachkov S.B., Dvoryankin S.V., Bonch-Bruevich A.M. Ogranicheniya formantnoi teorii razborchivosti rechi v prilozheniyakh zashchity rechevoi informatsii [Limitations of the formant theory of speech intelligibility in voice information protection applications]. *Voprosy kiberbezopasnosti – Cybersecurity Issues*, 2016, no. 5 (18), pp. 28–35.

29. Reva I.L., Trushin V.A., Ivanov A.V., O korrektirovke metodiki otsenki zashchishchennosti rechevoi informatsii ot utechki po tekhnicheskim kanalam [On correction of estimation technique for voice information protection against leakage through TCIL]. *Spetsial'naya tekhnika – Special Equipment*, 2016, no. 6, pp. 29–35.

Trushin Viktor Aleksandrovich, PhD, senior researcher, associate professor at the Department of Information Security, Faculty of Automation and Computer Engineering, Novosibirsk State Technical University. The main field of his research is information-measuring systems and information security. He is the author of more than 100 publications including 1 monograph and 22 patents. E-mail: rastr89@mail.ru

Трушин Виктор Александрович, кандидат технических наук, старший научный сотрудник, доцент кафедры защиты информации факультета автоматики и вычислительной техники Новосибирского государственного технического университета. Основное направление научных исследований – информационно-измерительные системы, защита информации. Имеет более 100 публикаций, в том числе одну монографию, 22 патента. E-mail: rastr89@mail.ru

DOI: 10.17212/1814-1196-2019-4-135-146

Анализ формантного метода оценки разборчивости речи как метода выполнения косвенных измерений[*]

В.А. ТРУШИН

630073, РФ, г. Новосибирск, пр. Карла Маркса, 20, Новосибирский государственный технический университет

rastr89@mail.ru

В работе рассматривается формантный метод определения разборчивости речи, предложенный Н.Б. Покровским. Анализируется его суть, алгоритмическая и техническая реализация с точки зрения информационно-измерительного подхода, т. е. как метод выполнения косвенных измерений. Предлагается формирование тестового акустического воздействия на основе унифицированного речеподобного сигнала из слоговых или словесных артикуляционных таблиц. Приводятся результаты экспериментальных исследований по синтезу таких сигналов, их энергетические спектры и плотности распределения вероятностей. Обсуждаются вопросы оценки погрешности в определении разборчивости речи: самого понятия погрешности для разных случаев применения, погрешности косвенных измерений, теоретической погрешности, обусловленной степенью соответствия принятой модели измерения (оценки) разборчивости речи реальному объекту, т. е. слуховому аппарату человека. На основе анализа результатов исследований периферийной слуховой системы делается вывод о необходимости пересмотра принятой модели оценки разборчивости речи.

Ключевые слова: разборчивость речи, формантный метод, артикуляционные испытания, методика выполнения измерений, погрешность оценки разборчивости

СПИСОК ЛИТЕРАТУРЫ

1. ГОСТ Р 50840-95. Передача речи по трактам связи. Методы оценки качества, разборчивости и узнаваемости. – Введ. 1997–01–01. – М.: Госстандарт России, 1996. – 229 с.
2. ГОСТ 16600–72. Передача речи по трактам радиотелефонной линии. Требования к разборчивости речи и методы артикуляционных испытаний. – Взамен ГОСТ 16600-71: введ. 01.01.1974. – М.: Изд-во стандартов, 1973. – 90 с.
3. *Рашевский Я.И., Каргашин В.Л.* Обзор зарубежных методов определения разборчивости речи // Специальная техника. – 2002. – № 4. – С. 37–46.
4. *Алдошина И.А., Приттс Р.* Музыкальная акустика. – СПб.: Композитор, 2006. – 760 с.
5. *Покровский Н.Б.* Расчет и измерение разборчивости речи. – М.: Связьиздат, 1962. – 392 с.
6. *Быков Ю.С.* Теория разборчивости речи и повышения эффективности радиотелефонной связи. – М.: Госэнергоиздат, 1959. – 351 с.
7. *Сапожков М.А.* Речевой сигнал в кибернетике и связи. – М.: Связьиздат, 1963. – 452 с.
8. ГОСТ 25902–83. Зрительные залы. Метод определения разборчивости речи. – Введ. 01.01.84. – М.: Изд-во стандартов, 1984. – 10 с.
9. *Железняк В.К., Макаров Ю.К., Хорев А.А.* Некоторые методические подходы к оценке эффективности защиты речевой информации // Специальная техника. – 2000. – № 4. – С. 39–45.

---

[*] *Статья получена 16 августа 2019 г.*

10. ГОСТ 17187–2010. Шумомеры. Ч. 1. Технические требования. – Взамен ГОСТ 17187–81 в части технических требований: введ. 2012–11–01. – М.: Стандартинформ, 2012. – 31 с.

11. ГОСТ 8.563–2009. Методики (методы) измерений. – Взамен ГОСТ Р 8.563–96: введ. 2010–04–15 / Федеральное агентство по техническому регулированию и метрологии. – М.: Стандартинформ, 2010. – 15 с.

12. *Trushin V.A., Khitsenko V.E.* About the methods of forming a test signal in the instrumental evaluation of speaker clearance // 2019 Dynamics of Sistems, Mechanisms and Machines (Dinamics). – Omsk, 2019.

13. Справочник по радиовещанию / под общ. ред. В. Выходца. – Киев: Техніка, 1981.

14. *Кропотов Ю.А.* Модель одномерной плотности вероятности речевых сигналов // Системы управления, связи и безопасности. – 2015. – № 4. – С. 158–170.

15. *Дворянкин С.В., Макаров Ю.К., Хорев А.А.* Обоснование критериев эффективности защиты речевой информации от утечки по техническим каналам // Защита информации. Инсайд. – 2007. – № 2. – С. 18–25.

16. *Козлачков С.Б., Дворянкин С.В., Бонч-Бруевич А.М.* Принципы формирования тестовых речевых сигналов при оценках эффективности технологий шумоочистки // Вопросы кибербезопасности. – 2018. – № 3 (27). – С. 9–16.

17. МИ 2083–90. Измерения косвенные. Определение результатов измерений и оценка их погрешностей. – М.: Изд-во стандартов, 1991. – 9 с. – (Рекомендация).

18. *Трушин В.А., Иванов А.В., Рева И.Л.* О методических погрешностях оценки словесной разборчивости речи в задачах защиты информации // Доклады ТУСУР. – 2012. – № 1 (25), ч. 2. – С. 180–185.

19. *Трушин В.А., Иванов А.В., Хиценко В.Е.* О выборе модели тестового сигнала при оценке защищенности речевой информации от утечки по техническим каналам // Труды СПИИРАН. – 2015. – Вып. 3. – С. 122–133.

20. *Шиффман Х.Р.* Ощущение и восприятие. – 5-е изд. – СПб.: Питер, 2003. – 928 с. – (Мастера психологии).

21. *Альтман Я.А., Таварткиладзе Г.А.* Руководство по аудиологии. – М.: ДМК-Пресс, 2003. – 360 с.

22. *Лабутин В.К., Молчанов А.П.* Модели механизмов слуха. – М.: Энергия, 1973. – 200 с.

23. Физиология речи. Восприятие речи человеком / Л.А. Чистович, А.В. Венцов, М.П. Гранстрем и др. – Л.: Наука, 1976. – 388 с.

24. *Мареев Г.О.* Возможности лазерного автодинного измерителя наносмещений в диагностике заболеваний уха: автореф. дис. … д-ра мед. наук: 14.01.03. – Саратов, 2013. – 39 с.

25. Лазерные автодинные измерения параметров движения барабанной перепонки / Д.А. Усанов, О.В. Мареев, А.В. Скрипаль, Г.О. Мареев // Российский журнал биомеханики. – 2012. – Т. 16, № 1. – С. 8–21.

26. *Варгин П.С.* Математическая модель улитки внутреннего уха [Электронный ресурс]. – URL: http://www.telenavt.ru/downloads/Ulitka%20vnutrennego%20uha%20cheloveka.pdf (дата обращения: 16.12.2019).

27. *Авдеев В.Б.* О некоторых направлениях совершенствования методических подходов, применяемых при оценке эффективности технической защиты информации // Специальная техника. – 2013. – № 62. – С. 1–10.

28. *Козлачков С.Б., Дворянкин С.В., Бонч-Бруевич А.М.* Ограничения формантной теории разборчивости речи в приложениях защиты речевой информации // Вопросы кибербезопасности. – 2016. – № 5 (18). – С. 28–35.

29. *Рева И.Л., Трушин В.А., Иванов А.В.,* О корректировке методики оценки защищенности речевой информации от утечки по техническим каналам // Специальная техника. – 2016. – № 6. – С. 29–35.

Для цитирования:

Трушин В.А.. Анализ формантного метода оценки разборчивости речи как метода выполнения косвенных измерений // Научный вестник НГТУ. – 2019. – № 4 (77). – С. 135–146. – DOI: 10.17212/1814-1196-2019-4-135-146. – Текст англ.

For citation:

Trushin V.A. The analysis of the formant method of speech intelligibility estimation as a method of performing indirect measurements. *Nauchnyi vestnik Novosibirskogo gosudarstvennogo tekhnicheskogo universiteta – Science bulletin of the Novosibirsk state technical university*, 2019, no. 4 (77), pp. 135–146. DOI: 10.17212/1814-1196-2019-4-135-146.