DOI: 10.20535/2523-4455.2019.24.5.192893

UDC 004.934

Predictive Estimation of Speech Intelligibility Masked by Noise Interference Using Analytical Modeling

A. O. Harasiuk[f], ORCID 0000-0001-7212-4174
M. V. Myronov, ORCID 0000-0001-9331-8699
V. V. Lozinsky, ORCID 0000-0001-7984-3964
Nguyen Thanh Vy, ORCID 0000-0001-5736-3521
A. V. Darchuk, ORCID 0000-0001-8378-9710
A. M. Prodeus[s], Dr.Sc.(Eng.) Prof., ORCID 0000-0001-7640-0850
Department of Acoustic and Acoustoelectronic acoustic.kpi.ua
National technical university of Ukraine "Igor Sikorsky Kyiv polytechnic institute" kpi.ua
Kyiv, Ukraine

***Abstract*—A detailed description of the speech intelligibility prediction algorithm using analytical modeling is presented. The efficiency of the proposed algorithm is tested for four types of noise interference: white, pink, brown and typical for classrooms. The consistency of the results with known similar results indicates the correctness of the proposed components of the analytical algorithm. In addition, we compared the results of evaluating speech intelligibility obtained in accordance with the "classical" approach with the results of evaluating the STI index of speech intelligibility, which allowed us to confirm the thesis of a low camouflage ability of white noise at low signal-to-noise ratios.**

Key words — speech intelligibility; forecast estimate; computer modelling; analytical modeling; noise interference.

I. Introduction

The task of calculating and measuring speech intelligibility is not new, its history currently covers 90 years, if the reference is from pioneering work [1]. Nevertheless, the scope of speech intelligibility assessment applications is constantly expanding, technical means of engineers are changing and improving, the list of factors taken into account when assessing speech intelligibility is increasing. As a result, there is a need for constant updating of the corresponding algorithms and software.

Overseas, the most widely used versions of the Formant Method for assessing speech intelligibility are the Articulation Index AI [2] and the Speech Intelligibility Index (SII) [3]. By the end of the 1950s, several scientific schools had been formed in the USSR headed by N. B. Pokrovskiy, M. A. Sapozhkov and Yu. S. Bykov, where his versions of the Formant Method developed [4]–[6].

In 1973, the Modulation Method appeared, where STI (Speech Transmission Index) is the measure of speech intelligibility [7]. Since the Modulation Method has the ability to take into account the influence of not only noise, but also reverberation on speech intelligibility, some authors even made statements about the "obsolescence" of the Formant Method [8]. At the same time, a careful comparison of the potential capabilities of the Formant and Modulation Methods indicates that the Formant Method is superior to its competitor in accuracy and speed of calculations in conditions where the action of noise prevails over the action of reverb [9].

II. Statement of the Problem

The "classical" computer simulation algorithm for evaluating the intelligibility of noisy speech by the Formant Method is described in [3], [4].

The structure of this algorithm is shown in Fig. 1. At the first stage of calculations, the primary speech signal and noise models are formed in the form of arrays of samples of stationary random processes with specified spectral characteristics. Then, the variance correction of these model processes is performed to provide the required integral signal-to-noise ratio SNR_0. After this correction, the partial signal-to-noise ratios E_k are estimated. At the final stage, speech intelligibility measures are calculated: Formant intelligibility A and verbal intelligibility W.

DOI: 10.20535/2523-4455.2019.24.5.192893

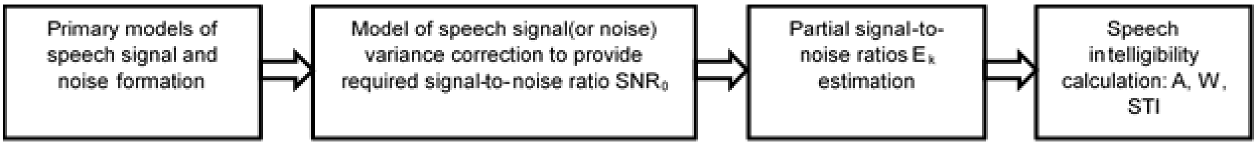

Fig. 1 The structure of the computer simulation algorithm

The essence of the Formant Method for assessing speech intelligibility is as follows. The frequency range of the speech signal is divided into adjacent frequency bands, with center frequencies and boundary frequencies and, within each of which the speech and noise spectra can be considered practically unchanged [4]. Verbal intelligibility is calculated through Formant intelligibility A [10]:

$$W = \begin{cases} 1{,}54 \cdot A^{0{,}25}[1-\exp(-11\cdot A)], & A < 0{,}15; \\ 1-\exp\left(\dfrac{11\cdot A}{1+0{,}7\cdot A}\right), & A \geq 0{,}15; \end{cases} \tag{1}$$

$$A = \sum_{k=1}^{K} p_k \cdot P(E'_k), \tag{2}$$

where p_k – is the probability of staying of Formants in the k-th frequency band:

$$p_k = F(f_{uk}) - F(f_{lk}); \tag{3}$$

$$F(f) = \begin{cases} 2{,}57\cdot 10^{-8} \cdot f^{2{,}4}, f \in (100;400]\,\text{Hz}, \\ 1-1{,}074\cdot e^{-10^{-4}\cdot f^{1{,}18}}, f \in (400;1000]\,\text{Hz}, \end{cases} \tag{4}$$

$P(E'_k)$ - speech perception coefficient.

In accordance with the method of N.B. Pokrovskiy [4],

$$P(E'_k) = \begin{cases} \dfrac{0{,}78+5{,}46\cdot \mathrm{e}^{-4{,}3\cdot 10^{-3}\cdot[27{,}3-|E'_k|]^2}}{1+10^{0{,}1\cdot|E'_k|}}, E'_k \leq 0, \\ 1-\dfrac{0{,}78+5{,}46\cdot \mathrm{e}^{-4{,}3\cdot 10^{-3}\cdot[27{,}3-|E'_k|]^2}}{1+10^{0{,}1\cdot|E'_k|}}, E'_k \leq 0, \end{cases} \tag{5}$$

E'_k – the effective level of sensation of Formants in the k-th frequency band:

$$E'_k = E_k - \Delta B(f_{0k}), \tag{6}$$

E_k – the effective level of sensation of a speech signal in the nth frequency band, equal (at sufficiently high noise levels) to the signal-to-noise ratio in this frequency band:

$$E_k = q_k = 10\lg\frac{D_{sk}}{D_{nk}}, \tag{7}$$

where D_{sk} and D_{nk} are the variances of the signal and noise in the k-th frequency band; $\Delta B(f)$ – the difference between the averaged spectra of speech and Formants:

$$\Delta B(f) = \begin{cases} 200/f^{0{,}43} - 0{,}37, & f \leq 1000\,Гц, \\ 1{,}37 + 1000/f^{0{,}69}, & f > 1000\,Гц. \end{cases} \tag{8}$$

Following the method of M. A. Sapozhkov, the spectrum of Formants is considered to practically coincide with the spectrum of speech, i.e. $\Delta B(f) = 0$, whence follows $E'_k = E_k$ [6]. In addition, as shown in [11], M. A. Sapozhkov differs $P(E'_k)$ from that of N. B. Pokrovskiy.

In [12], the method of M. A. Sapozhkov is clarified by taking into account the dependence of the perception coefficients on the frequency band. In this case, instead of (1), the relation is used:

$$A = \sum_{k=1}^{K} p_k \cdot P(E_k), \tag{9}$$

where the perception coefficients $P(E_k)$ are described by polynomial dependencies (Appendix 1).

In recent years, there has been a tendency toward partial unification of Formant and Modulation [7] Methods for assessing speech intelligibility. So, for example, according to the simplified method for assessing speech intelligibility, presented in GOST R ISO 24504-2015[1], speech intelligibility is evaluated using the STI index:

$$STI = \sum_{k=1}^{7} \alpha_k \cdot T_k - \sum_{k=1}^{6} \beta_k \cdot \sqrt{T_k \cdot T_{k+1}}, \tag{10}$$

$$T_k = \begin{cases} 0, E_k < -15 \\ (E_k + 15)/30, -15 < E_k < 15, \\ 1, E_k > 15 \end{cases} \tag{11}$$

where α_k are weight coefficients, β_k are redundancy coefficients, the values of which for octave bands with center frequencies f_0 are given in Table 1.

[1] GOST R ISO 24504-2015 Ergonomics design. Sound pressure levels of spoken announcements for products and public address systems

DOI: 10.20535/2523-4455.2019.24.5.192893

TABLE 1 WEIGHTING AND REDUNDANCY FACTORS FOR OCTAVE BANDS

f_{0k}, Hz	125	250	500	1000	2000	4000	8000
α_k	0,085	0,127	0,230	0,233	0,309	0,224	0,173
β_k	0,085	0,078	0,065	0,011	0,047	0,095	-

It is easy to see the fundamental similarity of relations (2) and (9), on the one hand, and (10), on the other hand. Moreover, the second term in (10) is a correction that takes into account the correlation of speech signals in adjacent frequency bands, and relation (11) can be interpreted as the result of linearization of the perception coefficient.

Despite the ability of computer simulations to evaluate the performance and effectiveness of software prototypes of real digital measuring systems, a predictive assessment of speech intelligibility is no less urgent. However, it is hardly rational to solve the forecasting problem by computer simulation, given the resource consumption of this method. More economical is the method of analytical modeling, according to which the speech and noise models are described by deterministic functions in the form of power distributions or spectral power densities.

In essence, the stages of analytical and computer modeling are similar. Moreover, relations (1) – (11) used at the final stage are the same for both types of modeling. However, the issue of the analytical description of the initial and intermediate stages in the literature is not adequately covered. One of the goals of this paper is to bridge this gap. Another goal is to compare the results of the assessment using relations (1), (2) and (10). Despite the obvious usefulness of such a comparison, it has not been implemented until recently.

III. NOISY SPEECH PREDICTION ALGORITHM

The structure of the proposed algorithm for predicting the intelligibility of noisy speech is similar to the structure of Fig. 1. Only the method for implementing the individual steps of the algorithm differs, based on the analytical description of the spectral properties of the speech signal and noise. We detail the description of each of the stages.

Stage 1. Formation of input data:

- specification of the analytical spectral model of the speech signal in the form of the dispersion distribution D_{sk}, where k is the number of the frequency band;
- creation of an analytical spectral model of noise in the form of a dispersion distribution D_{nk};
- setting the expected integral signal-to-noise ratio SNR_0 for the signal being listened to.

Stage 2. Correction of the distribution of variances of the speech signal (or noise) to provide a given integral signal-to-noise ratio SNR_0:

- calculation of the "primary" value of the integral signal-to-noise ratio

$$SNR = \frac{D_s}{D_n} = \frac{\sum_{k=1}^{K} D_{sk}}{\sum_{k=1}^{K} D_{nk}}; \tag{12}$$

- correction factor calculation

$$T = SNR_0 / SNR; \tag{13}$$

- adjustment of the dispersion distribution of the speech signal in accordance with the ratio

$$D_{0sk} = T \cdot D_{sk}, \tag{14}$$

- or adjusting the distribution of noise variances in accordance with the ratio

$$D_{0nk} = D_{nk} / T. \tag{15}$$

Stage 3. Calculation of partial signal-to-noise ratios for a given SNR_0:

- calculation of partial signal-to-noise ratios taking into account (14) and (15)

$$\begin{aligned} E_{0k} &= 10\lg\frac{D_{0sk}}{D_{nk}} = 10\lg\frac{D_{sk}}{D_{0nk}} = \\ &= 10\lg\left(D_{sk}/D_{nk}\right) + 10\lg(T) \end{aligned} \tag{16}$$

- correction of the values of E_{0k} in accordance with (6), if the Pokrovskiy technique is used.

Stage 4. Formation of the output:

- calculating articulatory intelligibility and verbal intelligibility in accordance with (1) and (2) or calculating the intelligibility index in accordance with (11).

Let's make some comments about the stage of the input data formation.

As the initial data, we can use the probability density of the signal and the interference, if we take into account the relationship between the dispersion D_k of a random stationary process in the k-th frequency band Δf_k and the average value of the power spectral density P_k within this frequency band, then:

$$P_k = D_k / \Delta f_k. \tag{17}$$

Obviously, when calculating speech intelligibility in communication lines, the value of the expected integral signal-to-noise ratio SNR_0 should be set based on typical operating conditions of the communication line. Slightly more difficult is to set SNR_0, if the speech is audible indoors.

DOI: 10.20535/2523-4455.2019.24.5.192893

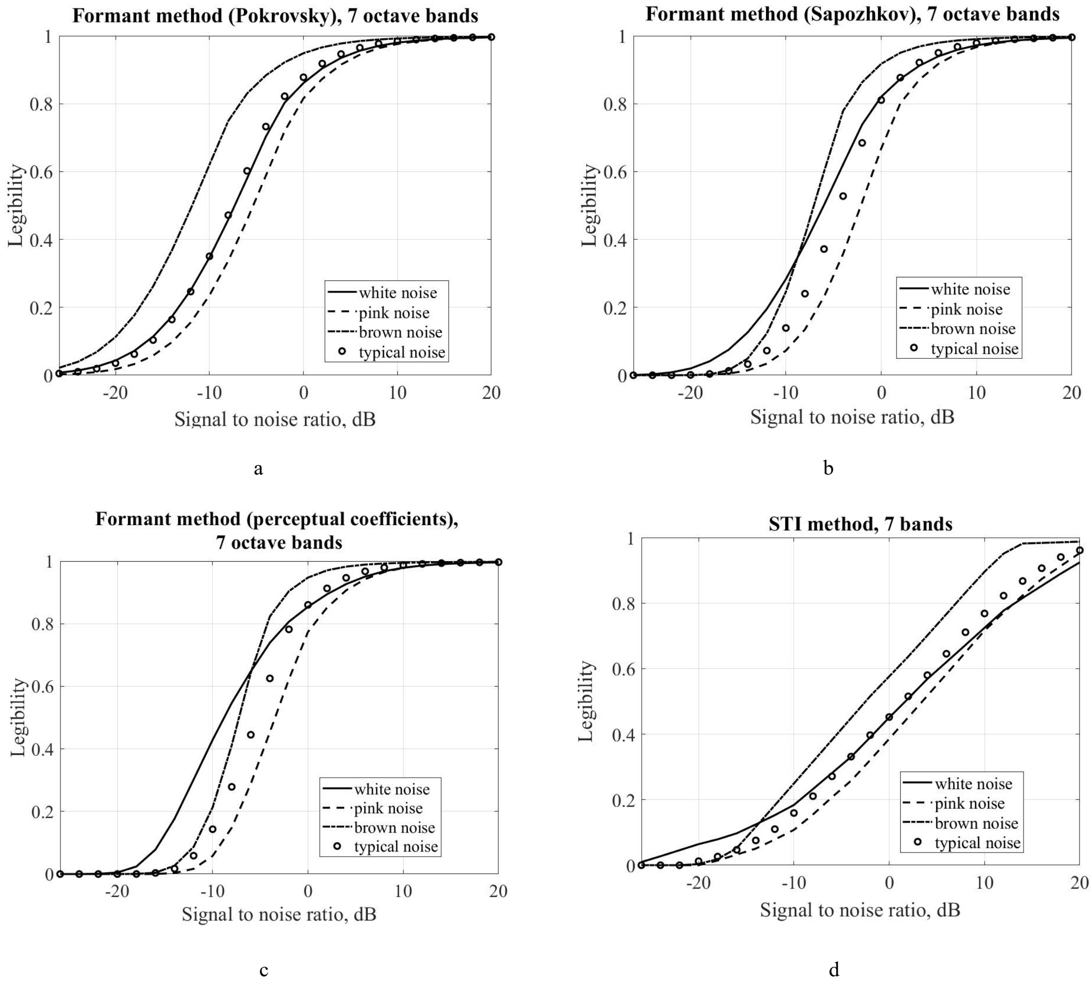

Fig. 2 Assessment of verbal (a, b, c) legibility and legibility index (d)

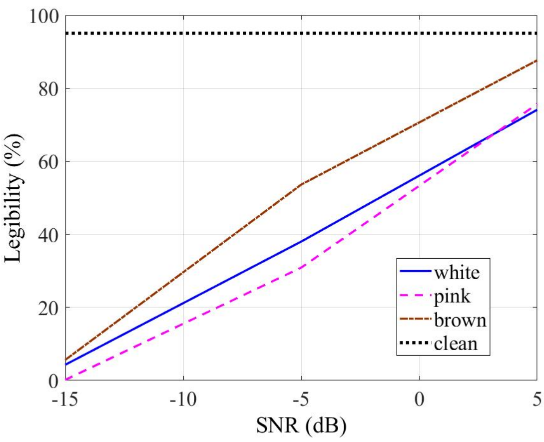

Fig. 3 Subjective assessments of syllabic intelligibility [13]

When calculating speech intelligibility in open-plan rooms (ISO 3382-3 : 2013[2] standard), which include offices, library halls, classrooms, for small (2-4 m) R distances between the speaker and the listener, we can assume that the speech signal attenuates in the same way as in free space, i.e. 6 dB with doubling the distance. With an increase in R, however, this pattern is violated, and then the calculation of the signal level at the listening point is somewhat more complicated.

IV. CHECKING THE PERFORMANCE OF THE ALGORITHM

The performance of the proposed algorithm was tested for 4 types of noise: white, pink, brown and typical for classrooms (Table 2). Note that the typical distribution of noise levels D_{nk}, dB over the frequency channels for classrooms is borrowed from the regulatory document SSN 3.3.6.037-99 [3], and the parameters of the long-term speech spectrum are borrowed from [4].

For maximum ease of subsequent calculations values, are given in Table 2, the dispersion distribution values are normalized to the dispersion value in the fourth frequency channel (f_{04} = 1000 kHz).

[2] ISO 3382-3 : 2013 ACOUSTICS - MEASUREMENT OF ROOM ACOUSTIC PARAMETERS - PART 3: OPEN PLAN OFFICES
[3] SSN 3.3.6.037-99 State Sanitary Norms. Sanitary norms of virobiotic noise, ultrasound and infrasound, Kyiv, 1999.

DOI: 10.20535/2523-4455.2019.24.5.192893

TABLE 2 PRIMARY DISTRIBUTION OF VARIANCES OF SPEECH AND NOISE

Frequency channels parameters							
f_{0k}	125	250	500	1000	2000	4000	8000
Δf_k	90	175	355	690	1400	2800	5270
White noise parameters							
D_{nk}	0,130	0,254	0,514	1	2,029	4,058	7,638
Pink noise parameters							
D_{nk}	1,043	1,014	1,029	1	1,014	1,014	0,955
Brown noise parameters							
D_{nk}	8,348	4,058	2,058	1	0,507	0,254	0,119
Typical noise hindrance parameters							
$D_{nk,dB}$	61	54	49	45	42	40	38
D_{nk}	5,193	2,017	1,284	1	1,023	1,273	1,5471
Long-term speech spectrum parameters							
$P_{sk,dB}$	60	70	68	59,5	52	46	40,5
D_{sk}	0,146	2,846	3,642	1	0,361	0,181	0,096

Fig. 2 presents the results of predictive calculations in the form of dependencies of speech intelligibility estimates on the expected signal-to-noise ratio SNR_0 in the range of values SNR_0 = -26 to +20 dB. Graphs Fig. 2a, Fig. 2b and Fig. 2c, constructed using perceptual coefficients N. B. Pokrovskiy, M. A. Sapozhkov and perceptual coefficients from Appendix 1, are in good agreement with the results presented in [11], [12]. Fig. 2d shows the estimates of STI (Fig. 2).

Comparing the graphs with each other, we see that the graphs obtained for the perception coefficients N. B. Pokrovskiy (Fig. 2a), fundamentally differ from the rest of the graphs in that in the entire range of considered values SNR_0 the masking properties of white noise turn out to be better than that of brown noise. Meanwhile, in [11], [12], the incorrectness of perception coefficients was first noted and corrected by N. B. Pokrovskiy.

Graphs Fig. 2b and Fig. 2c, constructed using adjusted perceptual coefficients, indicate that for small SNR_0 ($SNR_0 <$ -8 dB) the masking properties of white noise are inferior to those for brown noise. It is noteworthy that this result is consistent with the behavior of the graphs in

Fig. 2d, although in this case the threshold value of SNR_0 is $SNR_0 = -13$ dB.

Presented in fig. 3, the results of subjective articulation tests [13] also indicate a pronounced tendency to deteriorate the masking properties of white noise at low signal-to-noise ratios SNR_0. The absence of a clear loss of white noise in this case can be explained by the insufficient correctness of the organization of articulation tests.

As expected, the computational time for analytical modeling turned out to be an order of magnitude shorter than the time required for computer simulation and did not exceed 1 s for an FDA with a clock frequency of 2.66 GHz, 4 GB of RAM, and 32-bit OS.

CONCLUSION

A detailed description of the initial and intermediate stages of the speech intelligibility prediction algorithm by analytical modeling is presented. The efficiency of the proposed algorithm is tested for 4 types of noise interference: white, pink, brown and typical for classrooms. The consistency of the results with known similar results indicates the correctness of the proposed components of the analytical algorithm.

A comparison of the results of evaluating speech intelligibility obtained in accordance with the “classical” approach with the results of evaluating the STI index of speech intelligibility made it possible to confirm the thesis of a low camouflage ability of white noise at low signal-to-noise ratios. In the future, it is advisable to carry out an additional verification of this thesis by the method of articulation tests.

ACKNOWLEDGEMENTS

The authors thank the students of the Department of Acoustics and Acoustoelectronics of the National Technical University of Ukraine "Igor Sikorsky Kyiv Polytechnic Institute" for participating in the development and testing of the automated system of articulation tests.

REFERENCES

[1] J. Collard, “A Theoretical Study of the Articulation and Intelligibility of a Telephone Circuit,” *Electr. Commun.*, vol. 7, pp. 168–186, 1929.

[2] K. D. Kryter and J. H. Ball, “SCIM - A meter for measuring the performance of speech communicationsystems,” ESD-TDR-64-674, 1964.

[3] K. S. Rhebergen and N. J. Versfeld, “A Speech Intelligibility Index-based approach to predict the speech reception threshold for sentences in fluctuating noise for normal-hearing listeners,” *J. Acoust. Soc. Am.*, vol. 117, no. 4, pp. 2181–2192, Apr. 2005, **DOI:** 10.1121/1.1861713.

[4] N. B. Pokrovskiy, *Raschet i izmerenie razborchivosti rechi [Prediction and measurement of speech intelligibility]*. Moskow, USSR: Svyazizdat, 1962.

[5] M. A. Sapozhkov, *Rechevoy signal v kibernetike i svyazi [Speech signal in cybernetics and communication]*. Moscow: Svyazizdat, 1963.

[6] Y. S. Bykov, *eoriya razborchivosti rechi i povysheniye effektivnosti radiotelefonnoy svyazi [The theory of speech intelligibility and improving the effectiveness of radiotelephone communications]*. Moskow-Leningrad: Gosenergoizdat, 1959.

[7] H. J. M. Steeneken and T. Houtgast, “A physical method for measuring speech-transmission quality,” *J. Acoust. Soc. Am.*, vol. 67, no. 1, pp. 318–326, 1980, **DOI:** 10.1121/1.384464.

[8] I. Aldoshina and R. Pritts, *Muzykal’naya akustika [Musical acoustics]*. St. Petersburg: Composer, 2006, **ISBN:** 5-7379-0298-6.

[9] A. N. Prodeus, L. B. Dronzhevskaya, V. A. Klimkov, and D. A. Shagitova, “Formantnyy i formantno-modulyatsionnyy metody otsenki razborchivosti rechi. Chast’ 2. Tochnost’ i skorost’ izmereniy [Formant and formant-modulation methods for assessing speech intelligibility. Part 2. Accuracy and speed of measurements.],” *Electron. Commun.*, vol. 16, no. 6, pp. 16–24, 2011.

[10] V. K. Zheleznyak, Y. K. Makarov, and A. A. Horev, “Nekotoryye metodicheskiye podkhody k otsenke effektivnosti zashchity rechevoy informatsii [Some methodological approaches to assessing the effectiveness of protection of speech information],” *Spetsial’naya tekhnika*, no. 4, p. 39, 2000.

[11] A. N. Prodeus, A. V. Gavrilenko, and V. S. Didkovskiy, "Sopostavleniye versiy formantnogo metoda otsenki razborchivosti rechi [Comparison of versions of the formant method for assessing speech intelligibility]," *Electron. Commun.*, pp. 227–231, 2008.

[12] V. S. Didkovskiy, M. V. Didkovskaya, and A. N. Prodeus, *Akusticheskaya ekspertiza kanalov rechevoy kommunikatsii. Monografiya [Acoustic examination of the speech communication channels. Monograph]*. Kyiv, Ukraine: Imeks-Ltd, 2008, **ISBN:** 978-966-8861-85-7.

[13] A. M. Prodeus, K. V. Bukhta, P. V. Morozko, O. V. Serhiienko, I. V. Kotvytskyi, and O. O. Dvornyk, "Automated Subjective Assessment of Speech Intelligibility in Various Listening Modes," *Microsystems, Electron. Acoust.*, vol. 23, no. 3, pp. 49–57, Jun. 2018, **DOI:** 10.20535/2523-4455.2018.23.3.130367.

Надійшла до редакції 20 вересня 2019 р.

DOI: 10.20535/2523-4455.2019.24.5.192893

Appendix 1

Analytical Description of Perceptual Coefficients for Seven Octave Frequency Bands

Perception coefficients are described by the expression:

$$P_i'(x) = \begin{cases} 0, & x \le x'_{\min}; \\ \sum_{n=0}^{N'} a'_n x^n, & x'_{\min} < x \le 0 \\ \sum_{m=0}^{M'} b'_m x^m, & 0 < x \le x'_{\max}; \\ 1, & x > x'_{\max}. \end{cases} \quad \text{(A1.1)}$$

where the values of the coefficients a'_n and b'_m are presented in table. A1.1 and A1.2. These perceptual coefficients differ from the coefficients given in [12] by the presence of not five, but seven octave frequency bands.

TABLE A1.1. ODDS a'_n

N / F_{AVE}, Hz	0	1	2	3	4	5
125	0,324508	0,086098	0,004887	-0,00053	-6,56E-05	-1,87E-06
250	0,358141	0,070497	-0,00767	-0,00311	-0,00029	-8,76E-06
500	0,309301	0,054827	-0,00207	-0,00104	-7,77E-05	-1,82E-06
1000	0,19779	0,028848	0,000603	-0,00011	-7,30E-06	-1,45E-07
2000	0,238407	0,042094	0,000906	-0,00028	-2,29E-05	-5,43E-07
4000	0,231306	0,040168	0,001327	-0,00014	-1,07E-05	-2,07E-07
8000	0,178871	0,01705	-0,00064	-6,50E-05	2,91E-06	1,91E-07

TABLE A1.2. ODDS b'_m

M / F_{AVE}, Hz	0	1	2	3	4	5
125	0,315151	0,057481	-0,00116	-4,45E-05	2,10E-06	-2,22E-08
250	0,35763	0,088259	-0,00646	0,000232	-3,89E-06	2,46E-08
500	0,301964	0,07124	-0,00382	9,63E-05	-8,28E-07	-9,67E-10
1000	0,195479	0,032135	0,000349	-2,04E-05	-1,22E-07	5,59E-09
2000	0,238336	0,042175	0,002185	-0,00025	7,53E-06	-7,75E-08
4000	0,233574	0,037324	0,003033	-0,0003	8,87E-06	-9,05E-08
8000	0,185163	0,01171	0,002126	-2,81E-05	-2,24E-06	4,55E-08

УДК 004.934

DOI: 10.20535/2523-4455.2019.24.5.192893

Прогнозоване оцінювання розбірливості мови, замаскованої шумовою завадою, використовуючи аналітичне моделювання

Гарасюк[f] А. О., ORCID 0000-0001-7212-4174
Миронов М. В, ORCID 0000-0001-9331-8699
Лозінский В.В., ORCID 0000-0001-7984-3964
Тхань Ві Нгуєн, ORCID 0000-0001-5736-3521
Дарчук А. В., ORCID 0000-0001-8378-9710
Продеус[s] А. М., д.т.н. проф., ORCID 0000-0001-7640-0850
Кафедра Акустики та Акустоелектроніки acoustic.kpi.ua
Національний технічний університет України
"Київський політехнічний інститут імені Ігоря Сікорського" kpi.ua
Київ, Україна

***Анотація*—Результати акустичної експертизи приміщень та засобів зв'язку, що полягає в оцінюванні розбірливості мовних сигналів, є необхідними для сертифікації приміщень та комунікаційних систем. Оскільки технічні засоби звукоінженерів постійно змінюються й удосконалюються, а також з огляду на зростання переліку факторів, що можуть бути врахованими при оцінюванні розбірливості мови, зростає й кількість апаратно-програмних додатків для такого оцінювання. Таким чином, розробка математичного та програмного забезпечення для прогнозування та вимірювання розбірливості мови є актуальним завданням.**

Найбільш поширеними на сьогодні є формантний та модуляційний методи оцінювання розбірливості мови. Формантний метод є дещо обмеженим, оскільки не дозволяє враховувати дію реверберації. Модуляційний метод, в якому мірою оцінювання розбірливості мови є індекс передачі мови (Speech Transmission Index), є вільним від цього недоліку. Тому в деяких роботах можна зустріти висловлювання про «застарілість» формантного методу. Проте, ретельне зіставлення потенційних можливостей формантного та модуляційного методів свідчить, що формантний метод перевершує свого конкурента за точністю і швидкістю обчислень в умовах, коли дія шуму переважає над дією реверберації. Найбільшого поширення набули такі версії формантного методу оцінювання розбірливості мови як індекс артикуляції (Articulation Index) й індекс розбірливості мови (Speech Intelligibility Index). На території колишнього СРСР найбільш поширеними були версії формантного методу, розвинуті в наукових школах, очолюваних Н. Б. Покровським, М. А. Сапожковим і Ю. С. Биковим. Згідно із формантним методом, область частот мовного сигналу розбивають на суміжні частотні смуги, в межах кожної з яких спектри мови та шуму можна вважати практично незмінними, й формантну розбірливість обчислюють як певну функцію парціальних відношень сигнал-шум, а словесну розбірливість обчислюють через формантну розбірливість.

У даній статті представлено детальний опис алгоритму прогнозування розбірливості мови шляхом аналітичного моделювання. У загальному вигляді алгоритм складається з наступних кроків: на першому етапі обчислень здійснюється формування первинних моделей мовного сигналу і шуму у вигляді масивів вибірок стаціонарних випадкових процесів із заданими спектральними характеристиками. Потім виконується корекція дисперсій цих модельних процесів, щоб забезпечити необхідне інтегральне відношення сигнал-шум. Після такої корекції оцінюються парціальні відношення сигнал-шум. На заключному етапі обчислюються показники розбірливості мови, такі як формантна розбірливість, словесна розбірливість, індекс передачі мови.

Працездатність запропонованого алгоритму перевірена для 4-х видів шумової завади: білої, рожевої, коричневої та типової для навчальних приміщень. Узгодженість отриманих результатів з відомими аналогічними результатами свідчить про коректність запропонованих компонентів аналітичного алгоритму. Окрім того, виконано співставлення результатів оцінювання розбірливості мови, отриманих у відповідності до «класичного» підходу, з результатами оцінювання індексу передачі мови, що дозволило підтвердити тезу про низьку маскувальну здатність білого шуму при малих відношеннях сигнал-шум. Запропонований в даній роботі алгоритм прогнозування розбірливості мови буде корисним студентам в галузі звукоінженерії, фахівцям в галузях розробки та налагодження систем зв'язку, архітектурного проектування, озвучення залів різного призначення, а також керівникам установ, де розбірливість мови є важливим показником.

Ключові слова — розбірливість мови; прогнозована оцінка; комп'ютерне моделювання; аналітичне моделювання; шумова завада.

DOI: 10.20535/2523-4455.2019.24.5.192893

УДК 004.934

Прогнозированное оценивание разборчивости речи, маскируемой шумовой помехой, с использованием аналитического моделирования

Гарасюк[f] А. О., ORCID 0000-0001-7212-4174
Миронов М. В, ORCID 0000-0001-9331-8699
Лозинский В.В., ORCID 0000-0001-7984-3964
Тхань Ви Нгуен, ORCID 0000-0001-5736-3521
Дарчук А. В., ORCID 0000-0001-8378-9710
Продеус[s] А. Н., д.т.н. проф., ORCID 0000-0001-7640-0850
Кафедра Акустики та Акустоэлектроники acoustic.kpi.ua
Национальный технический университет Украины
"Киевский политехнический институт имени Игоря Сикорского" kpi.ua
Киев, Украина

***Аннотация*—Представлено детальное описание алгоритма прогнозирования разборчивости речи методом аналитического моделирования. Работоспособность предложенного алгоритма проверена для 4-х видов шумовой помехи: белой, розовой, коричневой и типовой для учебных помещений. Согласованность полученных результатов с известными аналогичными результатами свидетельствует о корректности предложенных компонентов аналитического алгоритма. Кроме того, произведено сопоставление результатов оценивания разборчивости речи, полученных в соответствии с «классическим» подходом, с результатами оценивания индекса разборчивости речи STI, что позволило подтвердить тезис о низкой маскировочной способности белого шума при малых отношениях сигнал-шум.**

Ключевые слова: разборчивость речи; прогнозная оценка; компьютерное моделирование; аналитическое моделирование; шумовая помеха.