EVALUATION OF THE INFLUENCE OF PINK NOISE (PN) ON THE INTELLIGIBILITY OF SPEECH IN THE SERBIAN LANGUAGE, USING SERBIAN MATRIX SENTENCE TEST (SMST) BASE AND STOI ALGORITHM

Dijana Kostić[1], Marina Milivojević[2], Zoran Milivojević[3]
[1]*Geonais Niš, Serbia*
[2]*Mikkelsen Electronics, Niš, Serbia,*
[3]*College of Applied Technical Sciences of Niš, Serbia*

Abstract
The first part of the paper describes the influence of Pink Noise (PN) and early reflection on intelligibility of sentences spoken on Serbian language, obtained from Serbian Matrix Sentence Test (SMST) base. Experiment was performed by objective method, using STOI test (STOI algorithm). In the second part of the paper the results of experiments, are shown in tabular and graphical form. By comparing the obtained results with the results of similar tests, and International Standard IEC 60268-16 the conclusion of intelligibility was brought.

Keywords: Intelligibility, SMST, Pink Noise, STOI.

1. INTRODUCTION

In the speech communication there are: the source that sends the message and the recipients who receives the message. In the transmitting a voice signal, we have a diffrent type of interference. The type and level of noise have, a diffrent influence on intelligibility. Intelligibility represented the measure intelligibility of speech in different conditions: a) background noise (industrial [1], babble [2], wind [3]), b) spatial effect (reverberation, early reflection [4]), c) distortion (equipment problems). Whatever form the noise takes, it will have influence to intelligibility when noise, and consonants of speech falls in the same frequency range. Different autors have a diffrent aprouch to the problem of the intelligibility. To determine the intelligibility, the researchers using test list with: a) words (meaning [5], without meaning-logatom [6]), or sentence (sematically corect [7] or matrix - syntactically defined [8]). In the paper [4] speech intelligibility was measured in diffuse noise by varyying the signal-to-noise ratio by either increasing the direct sound energy, or the energy of the early reflections. Arweiler and Buchholz, concluded that temporal integration of early reflections was facilitated when they arrived from the same direction as the direct sound. In addition, they found that speech intelligibility was better in binaural than in monaural listening conditions by 2 to 3 dB, which could be explained by spatial unmasking in the presence of the diffuse masker [4].

This paper analyzed the influence of early reflections, pink and diffuse noise on speech intelligibility in the Serbian language, using matrix sentences, obtained from the SMST base [9]. The experiment was performed in the College of Applied Technical Sciences of Niš, Serbia, using an objective test method, STOI test. The result intelligibility of sentence was determined using the dSTOI coefficient (using the STOI algorithm)[10]. The coefficient dSTOI goes in range from $0 \div 1$, where 0 is absolute inintelligibility and 1 is absolute intelligibility. By comparing results with the results of similar tests and the standard IEC 60268-16: 11 the conclusion of intelligibility was brought.

The organization of work is as follows. Section 2 describes early reflections and pink

noise. Section 3 describes the experiment, results and analysis of results of experiment. Section 4 is the conclusion.

2. EARLY REFLECTION AND PINK NOISE

The early reflection

The early reflection is defined as the sound observed within 50ms-80ms time delay after the arrival of the direct sound [11], where 50ms is is limit for speech, and 80ms for music. In paper [12] was shown that the reflection and increase of the delay time can increase intelligibility. In this paper was analyzing the influence of early reflection, with a time of delay Δt, from 0 to 50 ms.

Pink noise

In pink noise (PN), there is equal energy in all octaves of frequency, and power at a constant bandwidth of pink noise falls off at 3 dB per octave. The noise is subjectively hissy, unpleasant quality if presented at a volume that can provide significant acoustical privacy. Pink noise have same frequency balance as speech that so it provides the better masking, then white noise, but less then babble noise. The Fig.1 represents the spectral characteristic of pink noise.

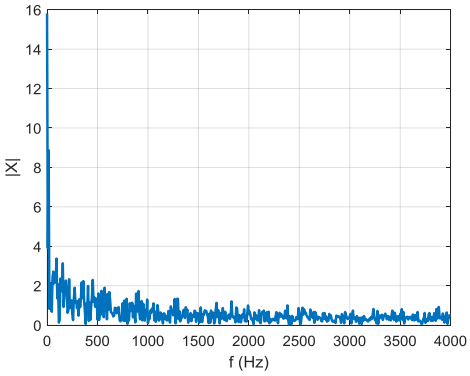

Fig. 1. *Spectral characteristic of Pink Noise (PN)*

3. EXPERIMENTAL RESULTS AND ANALYSIS

This section describes an experiment, performed to determine the intelligibility of sentences from the SMST base in the presence of pink and diffuse noise and early reflections.

3.1. EXPERIMENT

The experiment was realized in the following steps, shown in the block diagram in Fig.2:

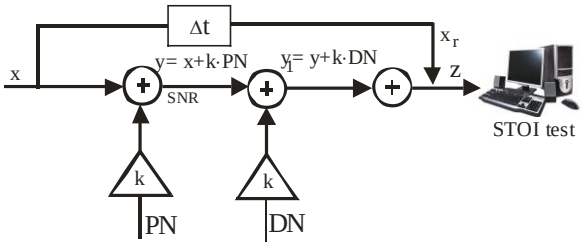

Fig. 2. *Block diagram of generated signal* ***z,*** *used for testing intelligibility with STOI tests*

By combining words from the SMST base, matrix sentences were obtained that represent a pure speech signal x. The speech signal is superimposed with pink noise, and using the parameter k we define the desired signal-to-noise ratio SNR = {-5, -2, 0} dB to get the signal y. The signal y is superimposed with the diffuse noise, y_1, also with a predefined signal-to-noise ratio SNR = {-5, -2, 0} dB. The signal used for binaural testing is obtained by sum the signal with noise y_1 and the reflected speech signal of signal x_r (generated in the circuit Δt). The parameters used during the performance of the experiment are: a) the angle of the speech signal ϕs = 0°, b) angle of reflection ϕn = 0°, c) angle of diffuse noise ϕDN = 0 : 5 : 360°, d) amplitude of reflection Ar=1, e) delay time between direct sound and reflection Δt = {0, 10, 25, 50} ms. Intelligibility was tested by STOI test [9] for both ears separately, for the case when SNR_{PN} and SNR_{DN}: a) equal and b) different. The results are presented in table and graphs form. After that, a comprehensibility analysis was performed based on the obtained values of the dSTOI coefficient.

SMST base

Sentences used for intelligibility testing were obtained from the SMST database described in [2]. The words were spoken in Serbian language, and they were recorded in studio of "Banker Radio" in Niš readed by a professional female speaker. Sentences are obtained by combining words by a computer, according to a random law, with a precisely defined sentence structure: name-verb-number-adjective-noun. The sampling frequency is Fs = 8kHz

3.2. THE RESULTS

The intelligibility results obtained using the STOI test (dSTOI coefficient) are shown in Tables 1-2 and the Fig 14-18. Table 1 shows the intelligibility of sentences in the presence of pink and diffuse noise for SNR_{PN}= SNR_{DN}. Table 2 shows the intelligibility of sentences in the presence of pink and diffuse noise for SNR_{PN}= SNR_{DN}, while Table 3 shows the mean values of intelligibility when the delay time is disregard Δt. An example of the time form of one of the tested sentences (”Danica pravi pet lepih brodova”) is given in Fig.3, for SNR_{PN}= SNR_{DN}=0dB and Δt=10ms.

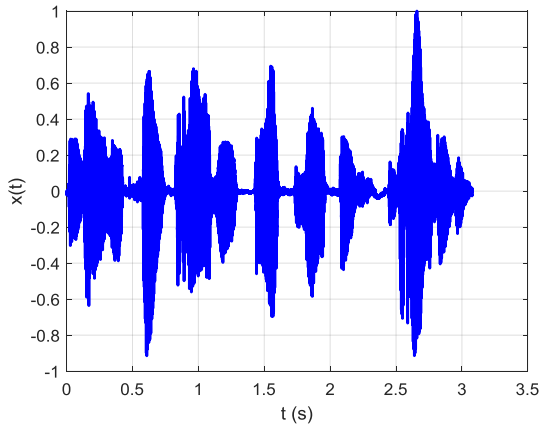

Fig. 3. *Time characteristic of speech signal of sentence from SMST base*

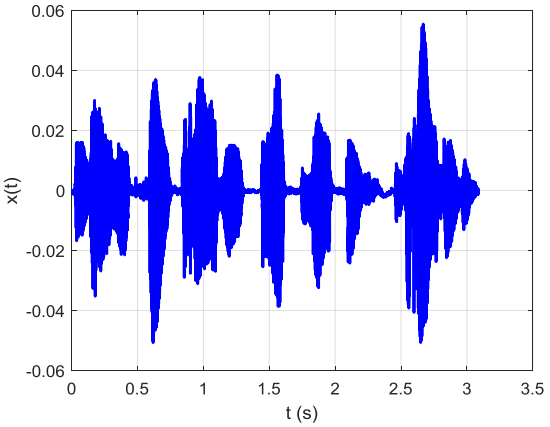

Fig.4. *Time characteristic of reflected speech signal*

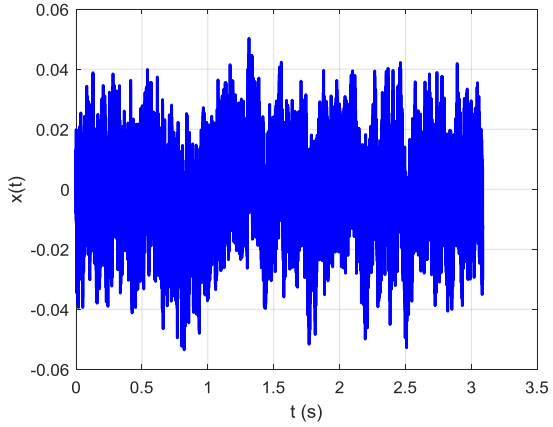

Fig.5. *Time characteristic of noise*

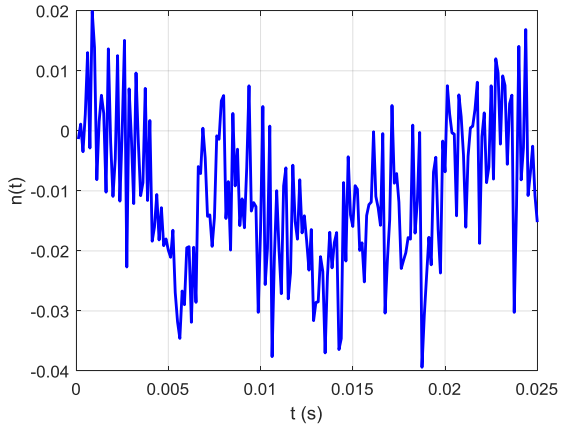

Fig.6. *Time characteristic of noise (detailed)*

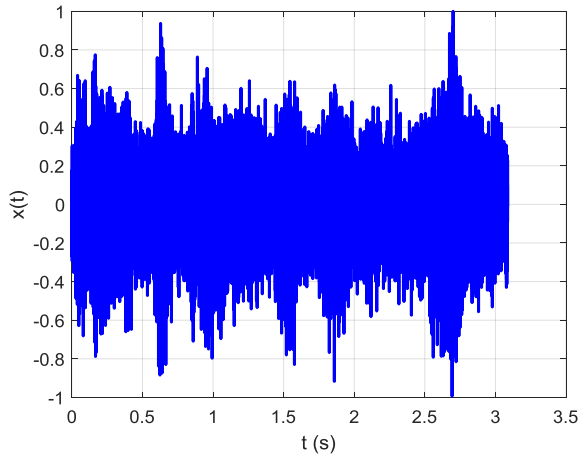

Fig.7. *Time characteristic of generated test signal*

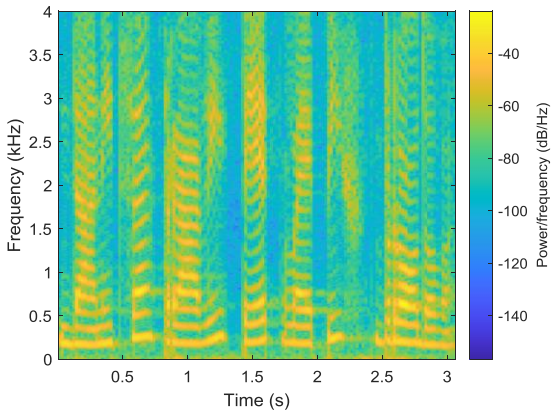

Fig.8. *Spectral characteristic of speech signal*

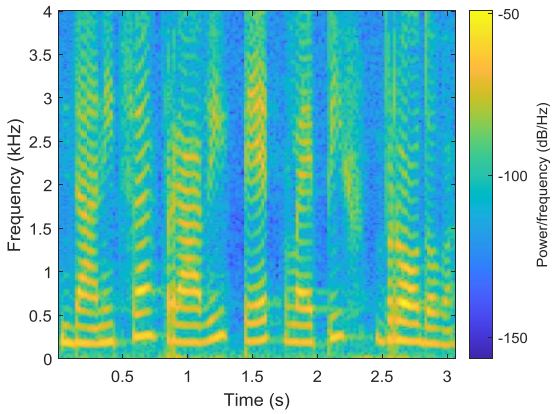

Fig.9. *Spectral characteristic of reflected speech signal*

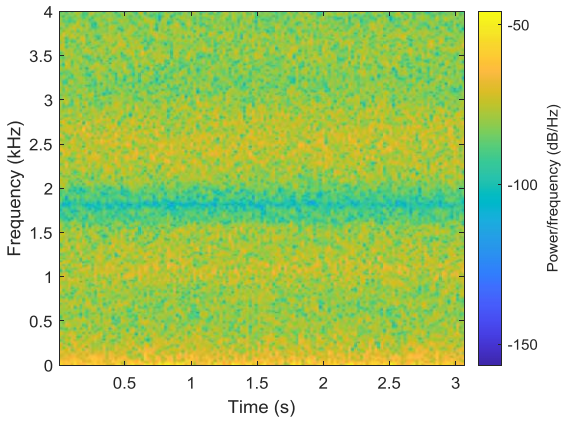

Fig.10. *Spectral characteristic of noise*

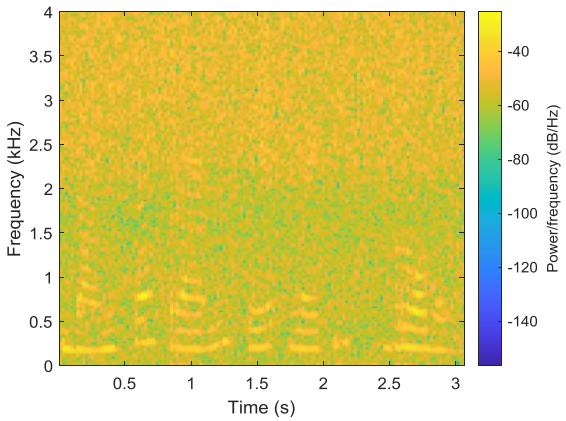

Fig.11. *Spectral characteristic of generated test signal*

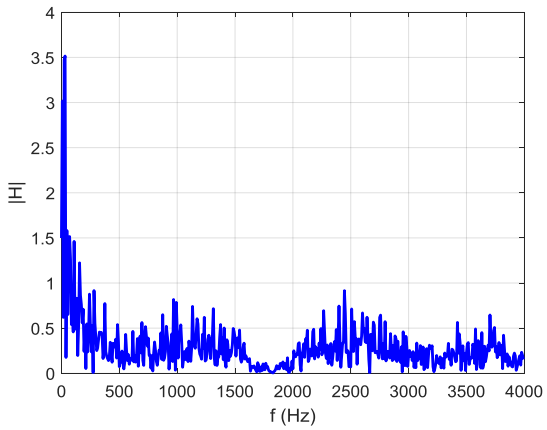

Fig.12. *Spectral characteristic of noise (detailed)*

Table 1. Intelligibility of sentence in the presence of PN and DN, $SNR_{PN} \neq SNR_{DN}$

Intelligibility						
SNR (dB)	**dSTOI**	**Δt (ms)**				
		0	**10**	**25**	**50**	**μ**
-2	$dSTOI_L$	0,7442	0,6199	0,5954	0,5768	0,6341
	$dSTOI_R$	0,7469	0,6226	0,6048	0,5764	0,6377
-5	$dSTOI_L$	0,7285	0,6230	0,6034	0,5585	0,6284
	$dSTOI_R$	0,7321	0,6250	0,6048	0,5650	0,6317

Table 2. Intelligibility of sentence in the presence of PN and DN, $SNR_{PN} = SNR_{DN}$

Intelligibility						
SNR (dB)	**dSTOI**	**Δt (ms)**				
		0	**10**	**25**	**50**	**μ**
0	$dSTOI_L$	0,7172	0,6311	0,6074	0,5704	0,6315
	$dSTOI_R$	0,7195	0,6335	0,6097	0,5681	0,6331
-2	$dSTOI_L$	0,6865	0,5861	0,5767	0,5020	0,5878
	$dSTOI_R$	0,6876	0,5921	0,5735	0,5218	0,5937
-5	$dSTOI_L$	0,6259	0,5738	0,5302	0,4971	0,5568
	$dSTOI_R$	0,6269	0,5537	0,5362	0,5042	0,5552

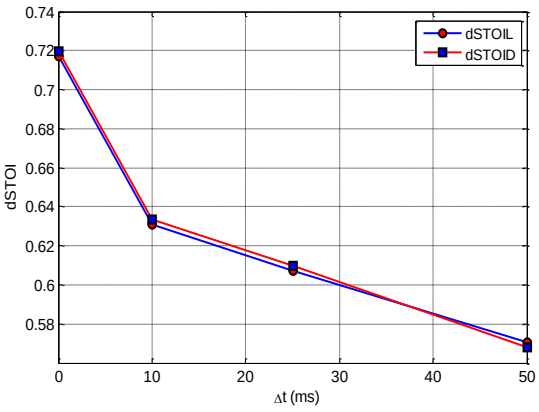

Fig. 14. *Intelligibility of sentence, STOI test for $SNR_{PN}=SNR_{DN}=0dB$*

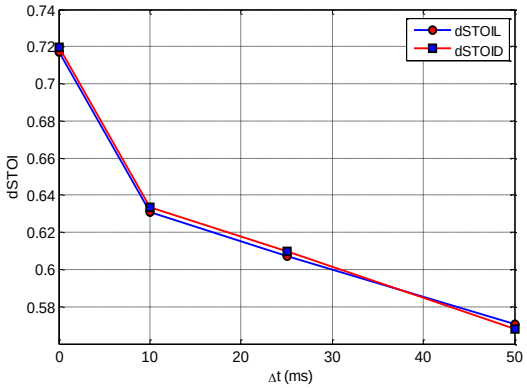

Fig. 15. *Intelligibility of sentence, STOI test for SNR_{PN}=-2dB, SNR_{DN}=0dB*

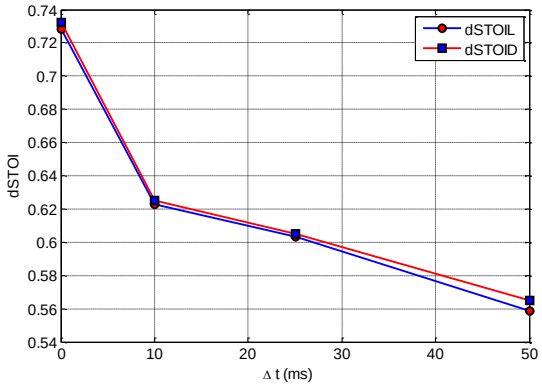

Fig. 16. *Intelligibility of sentence, STOI test SNR_{PN}=-5dB, SNR_{DN}=0dB*

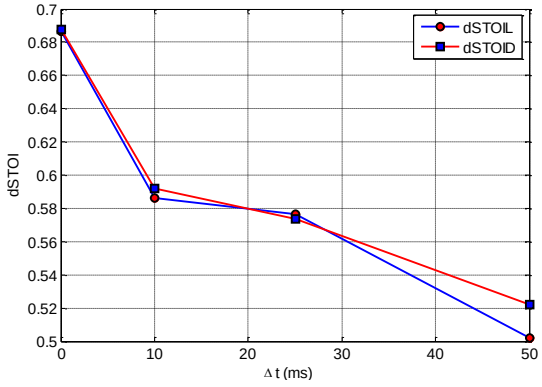

Fig. 17. *Intelligibility of sentence, STOI test for SNR_{PN}=SNR_{DN}=-2dB*

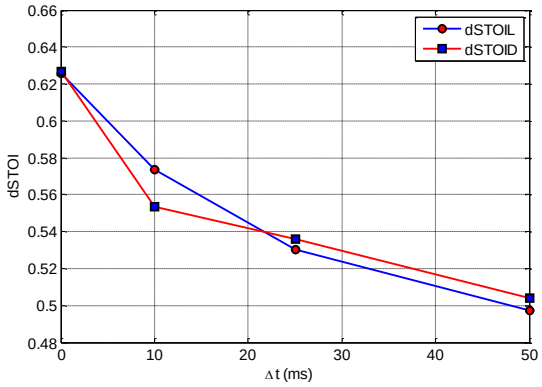

Fig. 18. *Intelligibility of sentence, STOI test for SNR_{PN}=SNR_{DN}=-5dB*

3.3. ANALYSIS OF RESULTS

Analyzing the results shown in Tables 1-2 and the Fig. 14-18 can be concluded that intelligibility of sentences for Δt={0,10,25,50} (ms) goes in range:

for SNR_{PN}=0dB, SNR_{DN}=-2dB:
a) from 0,5768 to 0,7442 for left ear,
b) from 0,5764 to 0,7469 for right ear,

for SNR_{PN}=0dB, SNR_{DN}=-5dB:
a) from 0,5585 to 0,7285 for left ear,
b) from 0,5650 to 0,7321 for right ear,

for SNR_{PN}=SNR_{DN}=0dB:
a) from 0,5704 to 0,7172 for left ear,
b) from 0,5681 to 0,7195 for right ear,

for SNR_{PN}=SNR_{DN}=-2dB
a) from 0,5020 to 0,6865 for left ear,
b) from 0,5218 to 0,6876 for right ear,

for SNR_{PN}=SNR_{DN}=-5dB
a) from 0,4971 to 0,6259 for left ear,
b) from 0,5042 to 0,6269 for right ear,

Analyzing the results shown in Tables 1 and 2, if we disregard the influence of the delay time Δt (ms), it can be concluded that the mean value intelligibility of sentence μ:

for SNR_{PN}=0dB, SNR_{DN}=-2dB:
a) 0,6341- left ear,
b) 0,6377 - right ear,

for SNR_{PN}=0dB, SNR_{DN}=-5dB:
a) 0,6284 - left ear,
b) 0,6317 - right ear,

for SNR_{PN}=SNR_{DN}=0dB:
a) 0,6315 - left ear,
b) 0,6331 - right ear,

for SNR_{PN}=SNR_{DN}=-2dB
a) 0,5878 - left ear,
b) 0,5937 - right ear,

for SNR_{PN}=SNR_{DN}=-5dB
a) 0,5568 - left ear,
b) 0,5552 - right ear,

Using comparative analysis and comparing with the standard IEC 60268-16: 2011, it is concluded that intelligibility belongs to the classification of poor intelligibility (0÷89%).

4. CONCLUSION

In this paper, speech intelligibility was tested in the presence of pink (PN) and diffuse (DN) noise under the influence of early reflections, for SNR values = {-5, -2, 0} dB, using sentences from the SMST database. The test results show that there is a difference in intelligibility between the left and right ear, in favor of the right, where the results of a little bit better intelligibility are present. Comparing the results of speech signal intelligibility in the presence of Gaussian and diffusion noise [13], with the influence of early reflections, it is noticed that there is no significant deviation in the results, but there is a difference in signal intelligibility in relation to the ear.

The hypothesis in [12] that the reflection and increase of the delay time can increase intelligibility, it is not confirmed here, because with the increase of the time of delay we also have a decrease of intelligibility.

REFERENCE

[1] Milivojević Z., Kostić D., Brodić D., "The Influence of Industrial Noise on the Performance of Speech Intelligibility Serbian Sentence Matrix Test", ICEST Niš, pp., 2017.

[2] Kostić D., Milivojević Z., Veličković Z., "The influence of early reflections and Babble noise on the intelligibility of speech signal", ICEST 2019, Ohrid, Macedonia.

[3] Milivojević Z., Kostić D., Veličković Z., Brodić D., "Efekat dejstva akustičke smetnje tipa vetar na razumljivost govora", INFOTEH Jahorina, pp., 2018.

[4] Arweiler, I., Buchholz, J. M. "The influence of spectral characteristics of early reflections on speech intelligibility", Journal of the Acoustical Society of America 130 (2): 996–1005, 2011.

[5] Egan J. P., "Articulation testing methods", Laryngoscope 58, 955-981, septembar, 1948.

[6] Kostić D., Milivojević Z., Stojanović V., "The Evaluation of Speech Intelligibility in the Orthodox Church on the Basis of MOS Test Intelligibility Logatom Type CCV", ICEST 2016, Ohrid, Macedonia, 2016.

[7] Plomp R., Mimpen A. M., "Improving the reliability of testing the speech reception threshold for sentences", Audiology, 18, pp. 43–52, 1979.

[8] Hagerman B., "Sentences for testing speech intelligibility in noise", Scand Audio, Vol. 11, pp. 79-87, 1982.

[9] Milivojević Z., Kostić D., Veličković Z., Brodić D., "Serbian sentence matrix test for speech intelligibility measurement in different reverberation conditions", UNITEH Gabrovo, pp. 173-178, 2016.

[10] Milivojević Z., Kostić D., Veličković Z., "The Optimization of the STOI Algorithm Parameters in Presence of the White Gaussian Noise (WGN)", ICEST 2018, pp-, 2018.

[11] Bradley J. S., Sato H., Picard M., "On the importance of early reflections for speech in rooms", J. Acoust. Soc. Am., vol. 113, no. 6, pp. 3233–3244, 2003.

[12] Parizet E., Polack J.D., "The influence of an early reflection upon speech intelligibility in the presence of a background-noise", Acustica 77, 21-30, 1992.

[13] Milivojević Z.,Kostić D., Veličković Z., "Speech intelligibility in the ambience of early reflections, Gaussian and diffuse noise", INFOTEH Jahorina, pp., 2020.