INFLUENCE OF INTERFERENCE ON SPEECH INTELLIGIBILITY IN A CAR USING THE BINAURAL METHOD

Marina Milivojević[1]; D. Kostić[2]; Zoran Milivojević[3];
[1]Mikkelsen Electronics, Niš, Srbija, marina.milivojevic@gmail.com
[2]Geonais, Niš, Srbija, koricanac@yahoo.com
[3]College of Applied Technical Science Nis, Niš, Srbija, zoran.milivojevic@akademijanis.edu.rs

Abstract: *The first part of the paper describes the influence of interference on speech intelligibility, which occurs in the environment of a moving car interior, in combination with diffuse noise and early reflections. The second part of the paper describes an experiment in which the intelligibility of the speech signal obtained from the Serbian matrix sentence test base - SMST, was tested by the binaural method, in the presence of early reflections, diffuse noise - DN and car noise - CN. Based on a comparative analysis of the obtained results with the results of similar tests, as well as with the standard IEC 6028-16: 2011, a conclusion was brought on the intelligibility and its classification.*

Keywords: *Intelligibility, diffuse noise, car noise, SMST base.*

1. ITRODUCTION

Intelligibility of speech is very important for good communication beetwen the people.The noise in everyday situation like: babble noise [1], idustrial noise [2], or natural sounds like wind [3], rain [4], production elements like music [5] will have influence to intelligibility of speech. In order to obtain reliable results for evaluation speech intelligibility in the presence of different types of ambient disturbances, there are different types of speech materials for testing. Various authors have dealt with this issue and they used database of word or sentence in their testing. In test with words there are tests which use: a database of known and meaningful words [6], but the results are certainly more reliable using logatom (words without meaning) [7] - [9], because the respondent cannot to sense the meaning of the word. When it comes to sentences, there are also two types of tests: a) testing sentences from everyday speech [10], and b) testing with matrix sentences [11] - [13]. Sentences from everyday conversation are certainly recognizable to the respondent, so it is possible that even if the respondent did not understand well, he/she can sense the meaning of the sentence, because the sentences are semantically correct. When it comes to matrix sentences, the respondent cannot guess their meaning because in the semantic sense they do not have to be correct, while the syntactic structure is important to them. This structure is defined as follows: name, verb, number, adjective, noun. For the purposes of testing in the Serbian language, the authors developed the database of the Serbian Matrix Sentences Test (SMST), described in [11]. The combination of words from the database give the 100000 different sentences, suitable for testing intelligibility. The selection of words from the database is done by random law with the help of a computer.

Intelligibility testing can be performed by subjective and objective methods. Subjective methods include the participation of respondents and evaluation using the MOS test (Mean Opinion Score, MOS). The objective method of testing is computer-assisted testing, using appropriate programs and algorithms. In our case, the STOI algorithm [14] was applied in order to obtain objective test results, and as a result its gives the values of dSTOI coefficient in the range from 0 -1.

In this paper analyze the influence of background noises in ambient inside the car on the move. Intelligibility of speech for passengers in car can be very difficult because interference like: noise from vehicle (engine, tires...), early reflections and other interferences. To evaluate the intelligibility of speech the experiment was performed in next steps: a) from the SMST base a speech signal was created (x), b) then this signal was delay for a predefined delay time Δt = (0, 10, 25, 50) ms (xr), then are c) superimposed the car noise (y) and d) diffuse noise (y1), e) obtained signal y1 and xr gave genereted test signal (z), f) using STOI algorithm and binaural method, for testing signal z, the value of coefficient intelligibility dSTOI, for the left and right ear individually, was got. After that, the results of experiment was analysed, compared with standard IEC 60268-16:11 [15], and with results of similar tests. Based on this the conclusion of intelligibility was brought.

Organization of work as follow. Section 2 describes experiment, experiment's results and their analysis. Section 3 is conclusion.

2. EXPERIMENTAL RESULTS AND ANALYSIS

2.1 Experiment

In fig. 1. the realization of an experiment for testing speech intelligibility, in the presence of background noises, is presented. The experiment was realized in the following steps: a) from the SMST base a test speech signal ***x*** was created, b) this signal was delay for a predefined delay time Δt = (0, 10, 25, 50) ms, signal ***xr***; c) superimpose the car noise, signal **y**, with the speech signal; d) then also superimpose diffuse noise, signal $\mathbf{y_1}$; e) test signal ***z***, was obtained by generating signal y_1 and $\mathbf{x_r}$; f) test signal **z**, passed through the STOI algorithm to obtain results; d) as a result of applying the binary method, the coefficient of intelligibility dSTOI is obtained, for the left and right ear individually.

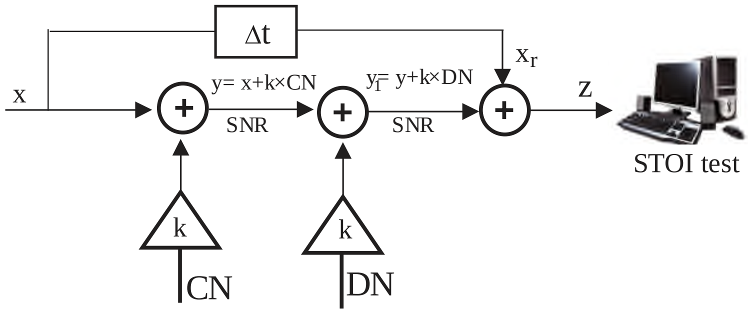
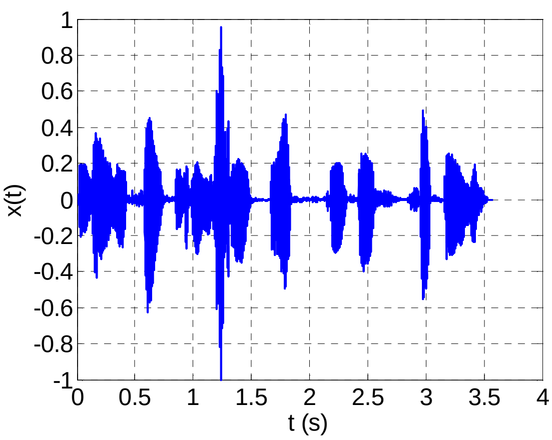


Figure 1: Block diagram of experiment

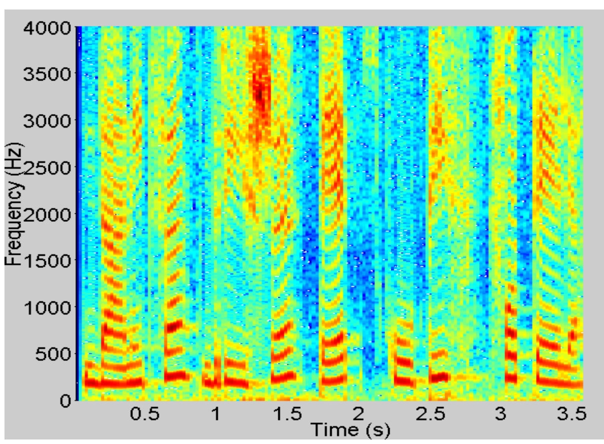
Tests were performed for two cases: a) when the SNR for car noise and diffuse noise are the same and b) when the SNR is different for a car noise and diffuse noise. Parameters used for the realization of the experiment, using the binaural method, are: a) variable signal-to-noise ratio (SNR) is adjusted for the purposes of the experiment by the gain factor k, which values are determined that the signal z is generated with SNR = (0, - 2, -5) dB; b) speech signal angle ϕ_s = 0°; c) angle of reflection ϕ_n = 0°; e) angle of diffuse noise ϕ_{DN} = 0 : 5 : 360°, e) amplitude of reflection Ar = 1; f) delay time between direct and reflected signal Δt = (0, 10, 25, 50) ms. The results of the experiment are presented in tables and graphs. Based on the obtained results, a intelligibility analysis was performed.

2.1.1. The base

Two bases were used in the paper: a) speech signal base and b) noise base. The SMST base described in [11] was used for the speech signal, contains 50 words spoken in Serbian by a professional female speaker. The base was recorded in studio of the Banker Radio in Niš (fs=44.1kHz,16bps). By combining these words, according to a random law, matrix sentences with a precisely defined syntactic structure were obtained: name, verb, number, adjective, noun, for testing. The advantage of these sentences is that they are not repeated, because with a combination of words from the base, it is possible to get 100000 sentences. The noise base used in the paper is available on the website [16]. For the purpose of experiment, the noise of a moving car was used from the base [16]. Figures 2-6 show the time and spectral forms of: a) speech signal (sentences obtained from the SMST base), b) reflected speech signal, b) interference signal (car noise) and d) generated test signal.

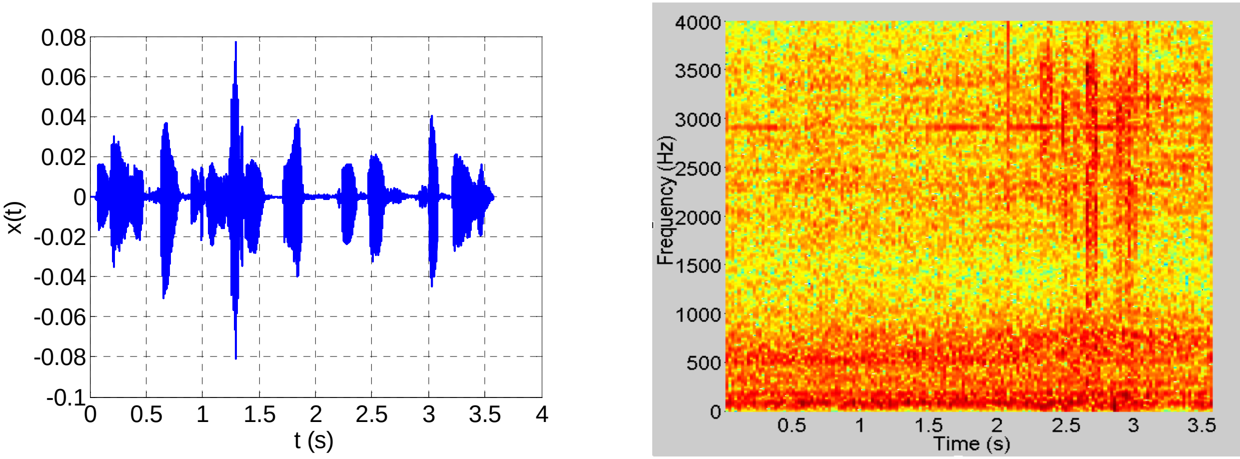


a)



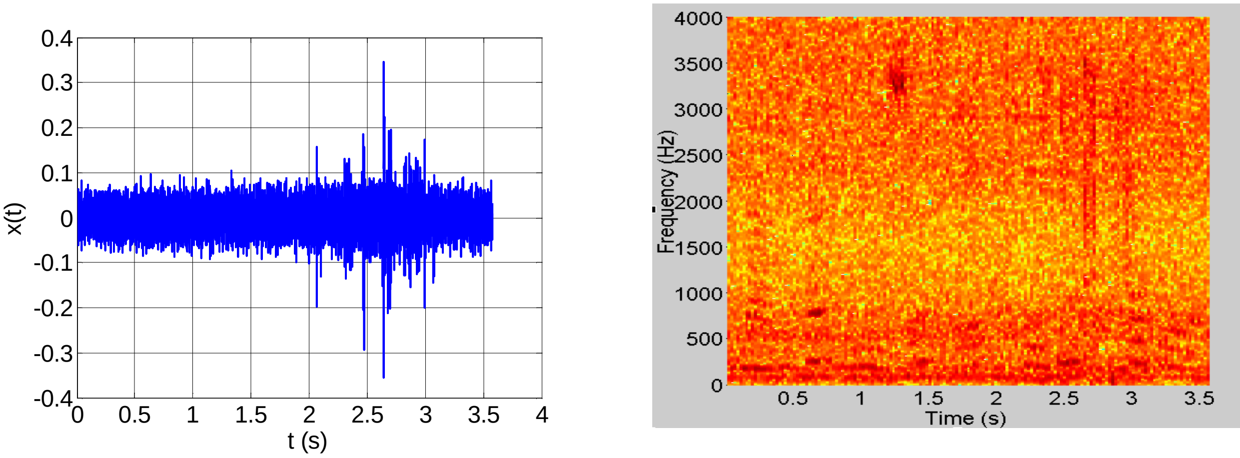
b)

Figure 2: The speech signal of the sentence '*Danica briše pet skupih fotelja*' from the SMST base:
a) time signal and b) spectrogram.



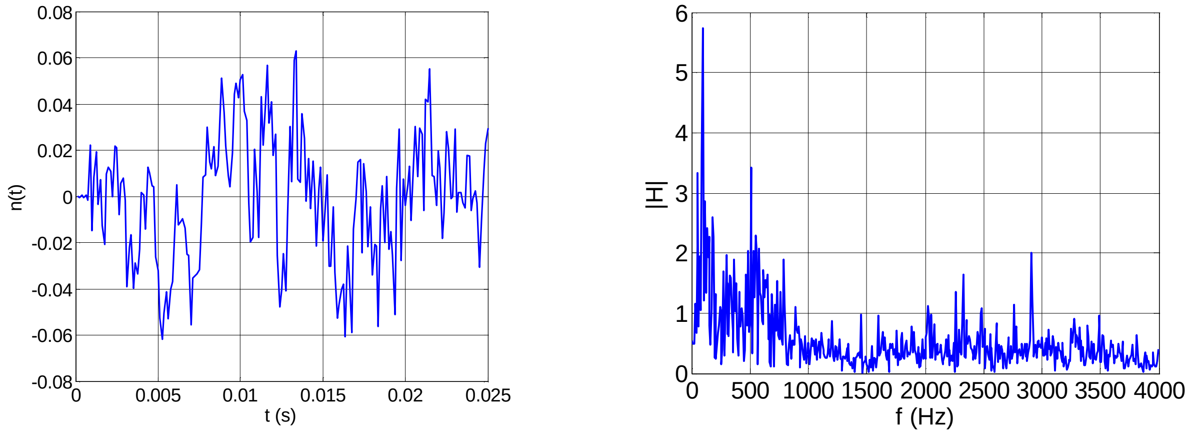
a) b)

Figure 3: Reflected speech signal of the sentence *'Danica briše pet skupih fotelja'*
a) time signal and b) spectrogram.



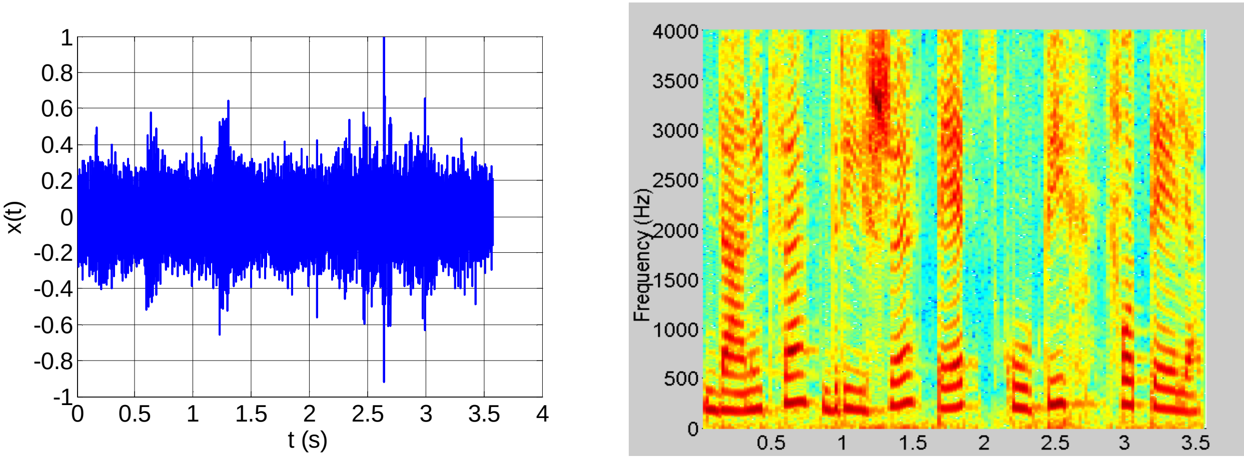
a) b)

Figure 4: Signal of car noise: a) time signal and b) spectrogram.



a) b)

Figure 5: Signal of car noise in details: a) time signal and b) spectrogram.

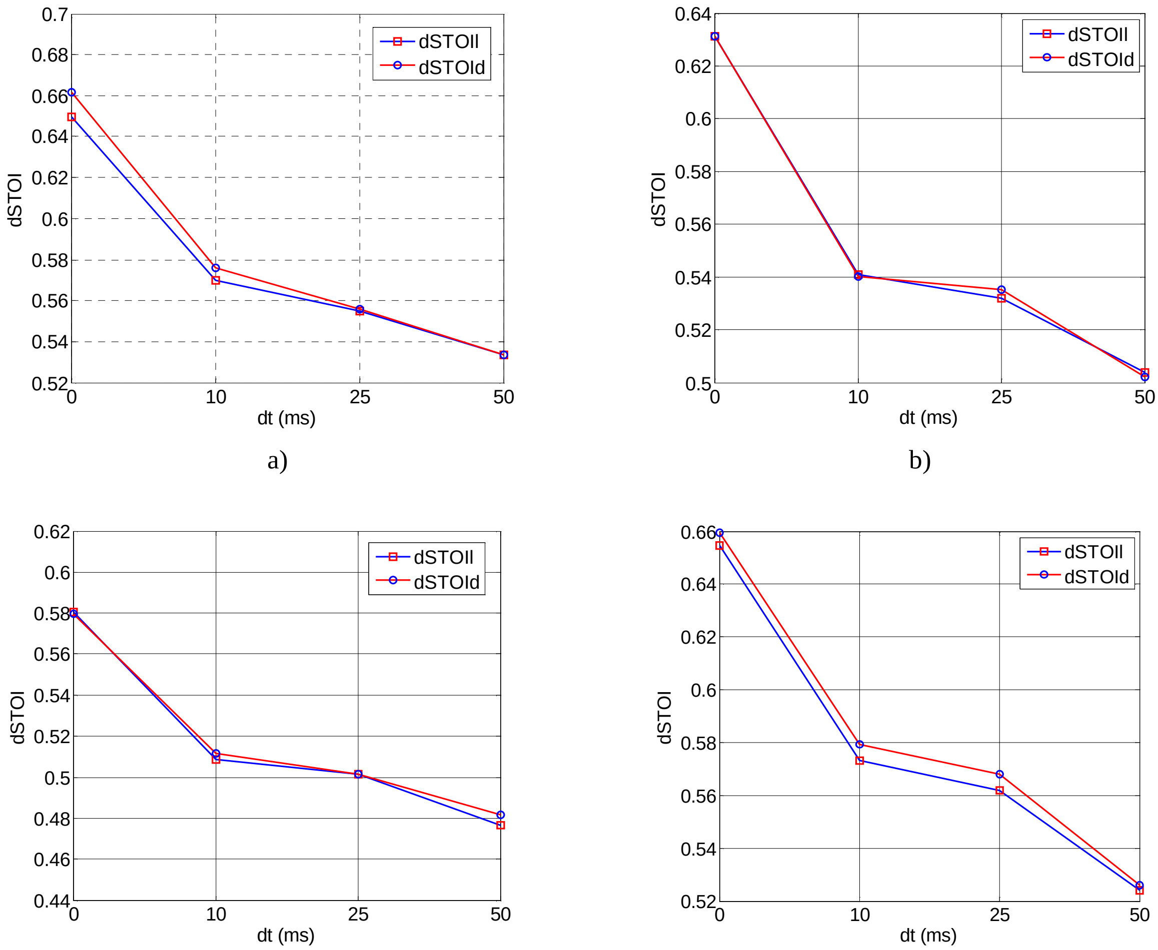


a) b)

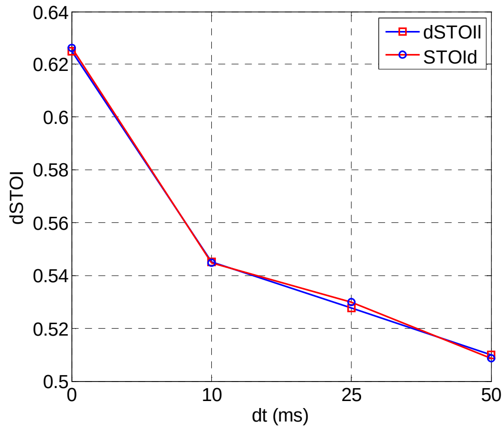
Figure 6: Speech test signal **z** (SNR=-5 dB): a) time signal and b) spectrogram.

2.2. The results

The results of the experiment are presented in tables and graphs. In tables 1-2, as well as in fig. 10-14 graphically shows the intelligibility of sentences for: a) $SNR_{CN} = SNR_{DN}$ = (0, -2, -5) dB; b) $SNR_{CN} \neq SNR_{DN}$ = (-2, -5) dB. Table 3 presents a comparative analysis of the results with the results for noise: a) pink; b) Gaussian and c) applause.



a) b) c) d)



e)

Figure 7: Intelligibility of speech for: a) SNR=0 dB, b) SNR=-2 dB, c) SNR=-5 dB za $SNR_{CN} = SNR_{DN}$ i d) SNR=-2 dB,SNR=-5 dB za $SNR_{CN} \neq SNR_{DN}$

Table 1: Speech intelligibility for $SNR_{CN} = SNR_{DN}$

Intelligibility						
SNR (dB)	**dSTOI**	**Δt (ms)**				
		0	**10**	**25**	**50**	**μ**
0	$dSTOI_L$	0.6487	0.5699	0.5548	0.5335	0.5770
	$dSTOI_R$	0.6616	0.5761	0.5557	0.5335	0.5817
-2	$dSTOI_L$	0.6311	0.5408	0.5319	0.5039	0.5519
	$dSTOI_R$	0.6313	0.5401	0.5350	0.5021	0.5521
-5	$dSTOI_L$	0.5804	0.5085	0.5014	0.4764	0.5167
	$dSTOI_R$	0.5798	0.5115	0.5013	0.4817	0.5186

Table 2: Speech intelligibility for $SNR_{CN} \neq SNR_{DN}$

Intelligibility						
SNR (dB)	**dSTOI**	**Δt (ms)**				
		0	**10**	**25**	**50**	**μ**
-2	$dSTOI_L$	0.6548	0.5733	0.5619	0.5240	0.5785
	$dSTOI_R$	0.6596	0.5794	0.5681	0.5259	0.5832
-5	$dSTOI_L$	0.6250	0.5450	0.5278	0.5100	0.5519
	$dSTOI_R$	0.6262	0.5447	0.5299	0.5086	0.5523

Table 3: Limit values of speech intelligibility in the presence of noise type: pink, Gaussian, applause

dSTOI	**middle value -μ**		
	Pink[17]	**Gaussian [17]**	**Applause[17]**
$dSTOI_L$	0.5568-0.6341	0.4713-0.6167	0.5615-0.6185
$dSTOI_R$	0.5552-0.6377	0.5593-0.6169	0.4408-0.6144

2.3. Analysis of results

Analyzing the obtained results of intelligibility of sentences spoken in the Serbian language in the presence of early reflection, diffusion noise, and car noise, for time of delay Δt = (0, 10, 25, 50) ms, and SNR = (0, -2, -5) dB, shown in Tables 1-2 and in Fig. 7, it is concluded that the coefficient dSTOI, goes in range from 0.5086 to 0.6612.

Considering the individual results of intelligibility, expressed through the STOI coefficient it can be concluded that the intelligibility is:

a) the best for the right ear 0.6616 (Δt = 0ms), and the worst for the left ear 0.5334 (Δt = 50ms), ($SNR_{CN} = SNR_{DN}$ = 0dB),
b) the best for the right ear 0.6313 (Δt = 0ms), and the worst for the left ear 0.5021 Δt = 50ms), ($SNR_{CN} = SNR_{DN}$ = -2dB),
c) the best for the right ear 0.5804 (Δt = 0ms), and the worst for the left ear 0.4764 (Δt = 50ms), ($SNR_{CN} = SNR_{DN}$ = -5dB),
d) the best for the right ear 0.6560 (Δt = 0ms), and the worst for the left ear 0.5240 Δt = 50ms), (SNR_{CN} = -2, SNR_{DN} = 0dB),
e) the best for the right ear 0.6262 (Δt = 0ms), and the worst for the left ear 0.5086 (Δt = 50ms), (SNR_{CN} = -5, SNR_{DN} = 0dB),

Analyzing the results shown for the mean value in Tables 1-2 for SNR = (0, -2, -5) dB, not observing the delay time, it is concluded that intelligibility is the best right ear 0.5167 for SNR_{CN} = -2, SNR_{DN} = 0dB, and the worst for the left ear 0.5832, for $SNR_{CN} = SNR_{DN}$ = -5dB.

Looking at the obtained results and comparing with the standard IEC 60268-16: 2011, it is concluded that intelligibility belongs to the classification of poor intelligibility, if we look at the results in percentage form (0 ÷ 89%). Comparing the obtained results with the results of similar tests, comparative analysis with the results for superimposed Gaussian noise [17], Pink noise [18] and applause noise [19], it is concluded that better intelligibility is present everywhere in the right ear, except for noise of applause where is the better intelligibility on the left ear. In paper [20], the results for several different algorithms used to assess intelligibility are presented for the different type of noises, from which it can be seen that the percentage of intelligibility (for car noise) goes in the range of 40-65%, depending on the applied algorithm, for SNR = 0dB.

3. CONCLUSION

The paper presents the results of an experiment in which speech intelligibility was evaluated in the presence of noise type car. Evaluation of intelligibility was performed using STOI test, for SNR = (0, -2, -5) dB and Δt = (0, 10, 25, 50) ms. The speech signal from the SMST base [11] and the signal of car noise from the noise base [16] were used. Analysis of the results showed that the reproduced speech signal have better intelligibility on the right ear. It can also be noticed that, as the delay time Δt t (ms) increases, the intelligibility of the speech signal decreases. Comparative analysis shows that the intelligibility for car noise is also better in the right ear. The answer to this better intelligibility can be found in the medical perception of man, because it is known that the left side of the brain processes the received information, which is usually accepted by the right ear [21].

REFERECES

[1] Kostić D., Milivojević Z., Veličković Z, The influence of early reflections and Babble noise on the intelligibility of speech signal, ICEST 2019, Ohrid, Macedonia
[2] 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.
[3] Milivojević Z., Kostić D., Veličković Z., Brodić Z, Efekat dejstva akustičke smetnje tipa vetar na razumljivost govora, INFOTEH Jahorina, pp. , 2018.
[4] Kostić D., Milivojević Z., Veličković Z.,Influence of the rain noise to intelligibility in Serbian language, UNITEH Gabrovo, 2018
[5] Milivojević Z., Kostić D., Veličković Z., Brodić D., Influence of the musical noise, type major accord, on the comprehensibility of speech in the Serbian language, SED 2017, Užice, Srbija
[6] Clark J.E., Four PB word lists for Australian English, Aust. J. Audiology 3(1), 21-31, 1981.
[7] 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, pp 153-156, Ohrid, Macedonia, 2016.
[8] Stojanović V., Kostić D., Milivojević Z., Veličković Z., Subjective evaluation of speech intelligibility in orthodox church based of the test intelligibility Nasals, Laterals and Affricates, UNITEH Gabrovo, 2016
[9] Tillman W. W., Carhart R., An expanded test for speech discrimination utilizing CNC monosyllabic word, Brooks Air Force Base, TX: Northwestern University AuditoryTest No. 6. USAFSchool of Aerospace Medicine TechnicalReport, 1966.
[10] Plomp R., Mimpen A. M., Improving the reliability of testing the speech reception threshold for sentences, Audiology, 18, pp. 43–52, 1979.
[11] Milivojević Z., Kostić D., Veličković Z., Brodić D., Serbian sentence matrix test for speech intelligibility measurement in different reverberation conditions, UNITEH Gabrovo, 2016
[12] Hagerman, B., Sentences for testing speech intelligibility in noise, Scand Audio, Vol. 11, pp. 79-87, 1982
[13] Wagener K., Lignel Josvassen J., Ardenkjær R., Design, optimization and evaluation of a Danish sentence test in noise, International Journal of Audiology, Volume 42, 2003 - Issue 1.
[14] 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.
[15] INTERNATIONAL Electrotechnical Commission IEC 60268-16 - International Standard: Sound system equipment – Part 16: Objective rating of speech intelligibility by speech transmission index, Switzerland: IEC, 2011.
[16] https://bigsoundbank.com/search?CatID=VEHCar
[17] Milivojević Z., Veličković Z., Kostić D., Razumljivost govora u ambijentu ranih refleksija, Gausovog i difuznog šuma, INFOTEH Jahorina, 2020
[18] Kostić D., Milivojević M., Milivojević Z.,Evaluation of the influence of Pink Noise (PN) on the intelligibility of speech in the Serbian language, using Serbian Matrix Sentence Test (SMST) base on STOI algorithm, Uniteh Gabrovo, pp- 2020
[19] Milivojević Z., Veličković Z., Kostić D., Milivojević M., Speech intelligibility in the Serbian language in the presence of ambient noise of applause, INFOTEH Jahorina, 2021
[20] Hu Y., Loizou P.C., A comparative intelligibility study of speech enhancement algorithms, 2007 IEEE International Conference on Acoustics, Speech and Signal Processing - ICASSP '07, 4, IV-561-IV-564.
[21] Acoustical Society of America. (2017, December 6). Want to listen better? Lend a right ear. ScienceDaily. Retrieved August 30, 2021 from www.sciencedaily.com/releases/2017/12/171206090611.htm