The Influence of Early Reflections and Babble Noise on the Intelligibility of Speech Signal

Dijana Kostić[1], Zoran Milivojević [2], Zoran Veličković [3]

Abstract – **The first part of the paper describes the effect of the first reflected component on intelligibility. In the second part of the paper, an experiment was described in which the intelligibility of words and sentences spoken on the Serbian language was tested under the conditions of superimposed Babble noise BN8 and in the presence of the first reflection. The analysis was performed for a delay up to 50 ms. Objective (STOI algorithm) and subjective (MOS test) testing were performed. The results are shown tabular and graphically. The analysis of the results determinated the delay, in which obtain the best intelligibility without the influence of the Babble noise, as well as the great destructiveness of the Babble noise to intelligibility.**

Keywords – **Intelligibility, Reflection, Babble noise, SMT.**

I. INTRODUCTION

Sound as a means of communication (speech or music) plays a major role in human life. For successful communication, it is necessary to provide good intelligibility. Intelligibility is one of the essential characteristics of communication systems, but also of spatial acoustics. On its transmission path from the sound source to the listener, the reproduced audio signal may suffer some degradation. Signal degradation refers to the impact of various types of interference: Gaussian Babble, Industrial noise, etc. In speech communication in a spatial acoustic (room, classroom, hall) early reflections also have a significant impact on the speech signal. The effect of the first reflection is particularly significant. In order to assess the intelligibility of speech in conditions of interference, different methods have been developed, and one them is speech audiometry. It involves the use of certain spoken material that is reproduced to the examinee: words (logatoms) or sentences (everyday and matrix), which after reproduction should repeat what they understood.

The words most commonly used in order to obtain objective test results are logatoms. Logatoms represent the meaningless words, composed of consonants (C) and vocals (V), constructing with the logatom type: CCV [1], CVC etc.

The words from everyday speech were created by Plomp and Mimpen in 1979 [2]. They formed a list of 170 sentences spoken in the Dutch language. When forming sentences, it was taken into account, that the words in the sentence do not contain more than 3 syllables.

The matrix sentences test was obtained by forming sentences according to the precisely defined order of the word in the sentence, so-called fixed syntax structure (name, verb, number, adjective, noun). This type of test was first developed for Swedish language by Hagerman [3]. The obtained results this kind of test have proven to be good, so this type of test has been developed for some other languages: Russian [4], Spanish [5], Serbian [6] ...

Respecting the rules set by Hagerman [3], the authors formed the Serbian Matrix Sentence Test - SMST base of 50 words (5 types of words x 10 words of each type) [6]. The authors also took care that the words were phonetically balanced and that they had no more than 3 syllables in their composition. In order to determine whether the words in the SMST base have a good appearance of the phonemes and reflect the spirit of the Serbian language, a comparative analysis was performed with capital literary works written in Serbian language: the novels "Bridge on the Drina" by Ivo Andrić, "Bakonja fra Brne" Sima Matavulj, epics "The Mountain wreath" by Petar Petrović Njegoš and drama "Koštana", author Bora Stanković [6].

In this paper, an experiment was performed to determine the influence of early reflections and the Babble noise on the intelligibility of the speech signal from the SMST base. The experiment is organized in several steps:

a) a speech signal consisting of a words from the SMST base were created, to which the Babble noise is superimposed, with the ratio SNR = {-5, -2, 0} dB;

b) the delay of the speech signal was made for values Δt = {0, 10, 25, 50} ms and

c) The testing intelligibility of speech signal was performing using: MOS and STOI test.

By comparing the results of the experiment with the results shown in [7] for Gaussian and Babble noise, as well as with the International standard [8], a conclusion was made for intelligibility of the speech signal.

The paper is organized in the following way. Section II describes the effect of the first reflexion and the delay time on intelligibility. Section III describes the experiment, experimental results and analysis. Section IV is a conclusion.

II. EFFECT OF THE FIRST REFLECTION

The early reflection is a sound wave reflected from a certain surface (an obstacle encountered in its path): a wall, a ceiling, and a floor and arrives to the listener almost at the same time as a direct sound. The first reflection usually appears very quickly after a direct wave in the range of 0 ÷ 50 ms [9]. In order to ensure good intelligibility, it is considered

[1] Dijana Kostić, College of Applied Technical Sciences of Niš, 20. Aleksandra Medvedeva, St, 18000 Niš, Serbia,
e-mail: koricanac@yahoo.com
[2] Zoran Milivojević, College of Applied Technical Sciences of Niš, 20. Aleksandra Medvedeva, St, 18000 Niš, Serbia,
e-mail: zoran.milivojevic@vtsnis.edu.rs
[3] Zoran Veličković, College of Applied Technical Sciences of Niš, 20. Aleksandra Medvedeva, St, 18000 Niš, Serbia,
e-mail: zoran.velickovic@vtsnis.edu.rs

that the first reflection should appear at an interval of 30 ÷ 80 ms after direct sound [10]. Beside that, in [11] it has been shown that early reflection can improve the intelligibility of speech in the absence of background noise.

It has been shown in [10], [11] that the first reflections that reach to the listener for about 35 ms after a direct wave, contribute to increasing the volume of the sound, and in this way can improve the intelligibility of speech. However, in conditions of ambient background noise there is no improvement in intelligibility [12], [13].

In the next part of the paper, an experiment was performed in which the intelligibility was tested in the presence of a strong first reflection (the amplitude of the direct component and the first reflected components are the same) and the superimposed Babble noise BN8.

III. Experimental results and analysis

The experiment was performed to determine the influence of early reflections and the Babble noise on the intelligibility of the speech signal from the SMST base (words and sentences).

A. Experiment

The experiment was realized in the following steps (Fig.1). *Step 1*: By combining words from the SMST base, sentences with a fixed grammatical structure are formed, i.e., a clear speech signal **x** is formed. *Step 2*: The Bubble noise **BN8** are superimposed to the clear speech signal, after the amplification **k**. In this way, the signal y with the specified SNR is formed. *Step 3*: The reflected speech signal $\mathbf{x_r}$ is generated with the time of delay Δt of the speech signal **x**. *Step 4*: By superimposed of the speech signal **x**, the Babble noise **BN8** and the reflected speech $\mathbf{x_r}$, a test signal **z** is formed. *Step 5*: Testing intelligibility of whole sentences and individual words by subjective MOS test and objective STOI test.

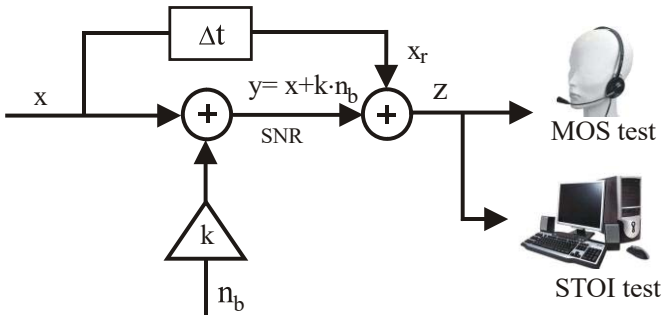

Fig 1. Block diagram of the realization MOS and STOI test.

Amplitude of the reflected signal is $\mathbf{A_r}$=1 [14].

The subjective MOS test (Step 5) was realized as follows. The test signal z was reproduced to the examinee trough the handset. After reproducing the test signal, the examinee repeated out loud what he had heard. After the answer the examiner note the accuracy of certain types of words as well as whole sentences. The statistical processing of the results gives the intelligibility expressed in percentages. The objective test was realized using the STOI algorithm [15]. Test result is intelligibility in percentages. At the end, a comparative analysis of the intelligibility of speech was made in the presence of early reflections and Babble noise with the results of intelligibility of speech in the presence of Gaussian and Babble noise [7].

B. SMST base

For the purposes of experiment, words from the SMST base were used [6]. The words were spoken in Serbian language and they recorded in the studio of "Banker Radio" in Niš (Serbia). The word was read by a professional female speaker. The voice signal is memorized in a **wav** format on a hard disk with Fs = 44.1 kHz, 16 bps.

C. Test group

The test group were formed from students of the College of Applied Technical Science in Niš, with a gender structure: 13 men and 7 women (aged 19 to 45), μ = 22.9 years. Before the test, the examinees said, they are not aware that they had hearing problems.

D. The results

The results of the experiment are shown in the Tables I ÷ III and graphically on Figs.2 ÷ 6. The experiment was realized for SNR = {-5, -2, 0} dB and Δt = {0, 10, 25, 50} ms.

TABLE I INTELLIGIBILITY OF SPEECH WITH SUPERIMPOSED BN8 (SNR = 0 dB), IN THE PRESENCE OF EARLY REFLECTION (Δt)

	Δt (ms)	0	10	25	50	μ
Inteligibility (%)	Name	30	30	40	30	32,50
	Verb	55	35	55	40	46,25
	Number	55	60	65	55	58,75
	Adjective	65	50	30	35	45
	Object	60	35	45	45	46,25
	Sentence	5	0	0	0	1,25
	dSTOI	59,44	58,66	61,32	59,40	59,70

TABLE II INTELLIGIBILITY OF SPEECH WITH SUPERIMPOSED BN8 (SNR = -2 dB), IN THE PRESENCE OF EARLY REFLECTION (Δt)

	Δt (ms)	0	10	25	50	μ
Inteligibility (%)	Name	65	40	50	25	45
	Verb	80	55	45	55	58,75
	Number	90	55	50	30	56,25
	Adjective	70	50	45	35	50
	Object	40	45	25	25	33,75
	Sentence	15	10	10	0	8,75
	dSTOI	54,13	54,05	56,28	55,66	55,01

TABLE III INTELLIGIBILITY OF SPEECH WITH SUPERIMPOSED BN8 (SNR=-5dB), IN THE PRESENCE OF EARLY REFLECTION (Δt)

	Δt (ms)	0	10	25	50	μ
Inteligibility (%)	Name	50	30	10	20	27,50
	Verb	70	35	25	20	37,50
	Number	75	55	45	35	52,50
	Adjective	40	10	15	30	23,75
	Object	40	20	5	10	18,75
	Sentence	10	0	0	0	2,50
	dSTOI	44,87	45,46	45,70	45,45	45,37

TABLE IV INTELLIGIBILITY OF SPEECH IN PRESENCE OF GAUSSIAN AND BABBLE NOISE [7]

		Inteligibility (%)		
SNR (dB)		-5	-2	0
Name	Gaussian	46.67	66.67	63.33
	Babble	26.67	40	53.33
Verb	Gaussian	53.33	63.33	70
	Babble	6.67	10	36.67
Number	Gaussian	53.33	63.33	70
	Babble	10	36.67	53.33
Adjectiv	Gaussian	56.67	60	63.33
	Babble	3.33	20	46.67
Object	Gaussian	33.33	53.33	53.33
	Babble	6.67	16.67	30
Sentence	Gaussian	0	20	20
	Babble	0	6.67	6.67

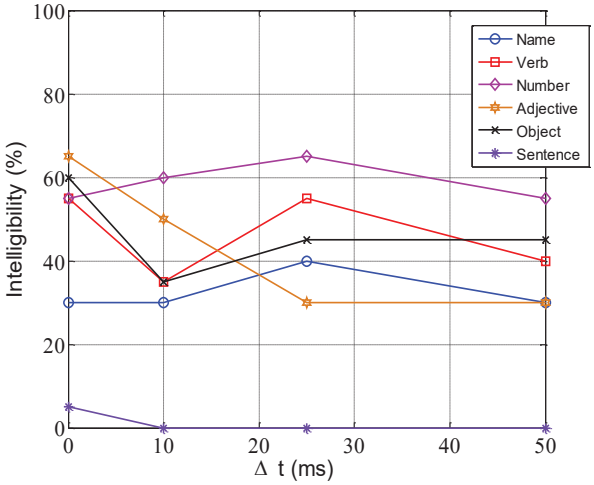

Fig. 2. Intelligibility of speech for: a) sentences and b) word for the SNR=0dB.

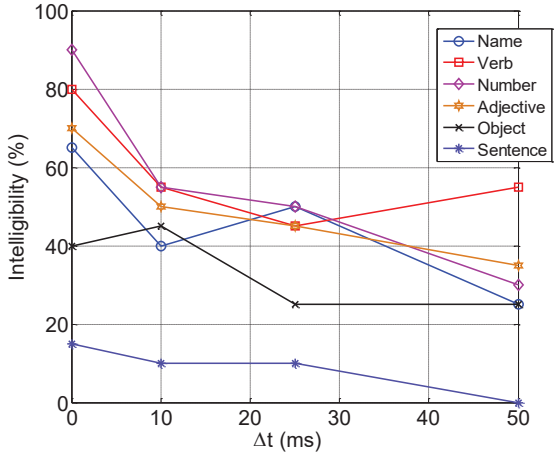

Fig 3. Intelligibility of speech for: a) sentences and b) word word for the SNR=-2dB.

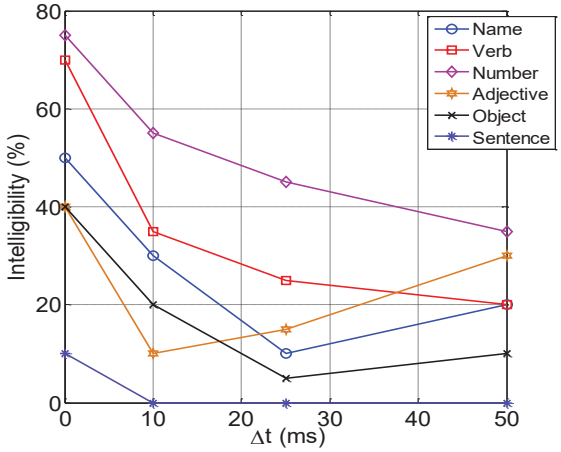

Fig 4. Intelligibility of speech for: a) sentences and b) word word for the SNR= -5dB.

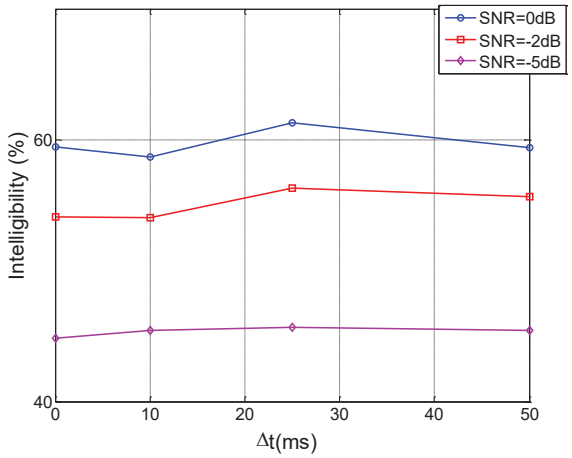

Fig 5. Intelligibility of speech dSTOI coefficient.

A. Analysis the results

Based on the results shown in Table I ÷ III and Fig. 2 ÷ 4 it can be concluded that intelligibility of type of word go in range:

a) for the 0 dB from 30 ÷ 65%,
b) for the -2 dB from 25 ÷ 90%,
c) for the -5 dB from 5 ÷ 75%,

Based on the results shown in Table I ÷ III and Fig. 2 ÷ 4 it can be concluded that intelligibility of word: *'Number'* is the best 90% for -2 dB and Δt = 0 ms, and the worst is *'Object'* for -5 dB and Δt = 25 ms.

Based on the results shown in Table I ÷ III and Fig. 2 ÷ 4 it can be concluded that intelligibility of sentence goes in range from 0 ÷ 15%, and it is the best for -2 dB (15%) and Δt = 0 ms, and the worst is for 0 dB (0%) and Δt = (10, 25, 50) ms, for 2 dB (0%) and Δt = 50 ms, for -5 dB (0%) and Δt = (10, 25, 50) ms.

Analysis the results shown in Table I÷III and Fig.5 it can be notice, that is dSTOI coefficient of intelligibility goes in range:

a) for the 0 dB from 58,66 ÷ 61,32%
b) for the -2 dB from 54,05 ÷ 56,28%
c) for the -5 dB from 44,87 ÷ 45,70%.

It can be notice that the greater time of delay have influence on the intelligibility. The great time of delay in combination with parameter SNR, when is strength of noise signal increased relative to strength of speech signal, have as results lower intelligibility of speech.

Observing the results from [7] in Table IV, comparing intelligibility of the speech signal in the presence of Gaussian and in the presence of Babble noise, it can be concluded that the results for Gaussian noise is better, while the results in the presence of the Babble noise, with and without the presence of the early reflections, do not differ significantly.

IV. Conclusion

In this paper the experimental results of speech intelligibility testing in the presence of Babble noise and early reflections are presents. The intelligibility test was performed by a subjective (MOS test) and an objective (STOI algorithm) test. Subjective MOS results show that increasing of value SNR = (-5 ÷ 0) dB increases the intelligibility of the word (18,75 ÷ 58,75) %, and sentences (2,5 ÷ 8,75) %. Increasing the time of delay of first reflection Δt = (0 ÷ 50) ms leads to a decrease intelligibility of word (90 ÷ 5) %, and sentences (15 ÷ 0) %. Objective STOI results refer to the intelligibility of whole sentences. These results show that with the increasing of the SNR = (-5 ÷ 0) dB and Δt = (0 ÷ 50) ms have influence on intelligibility of sentences. Increasing the SNR value the intelligibility of the sentences have been increases (45,37 ÷ 59,70) %, while when Δt increase the intelligibility of sentence decrease (61,32 ÷ 45,45) %.

The test results confirm the conclusion that there is a constant intelligibility of the word [10] up to 20 ms, and that the delay of about 30 ms can raise speech Intelligibility [12].

From the Figs. 2 it can be noticed that after "constant intelligibility" the intelligibility of words and sentences have been improved at Δt = 25 ms.

According to the standard IEC 60268-16, the obtained results for intelligibility are classified as "bad intelligibility" words (0 ÷ 67) % and sentences (0 ÷ 89) %.

The analysis shows that there is a greater degradation of intelligibility in the presence of Babble noise because of the identical distribution of energy in the spectral domain.

References

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

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

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

[4] M. Boboshko, A. Warzybok, M. Zokoll, N. Maltseva, RUMatrix test: construction, evaluation and clinical validation. Otorhinolaryngologia Hungarica. Vol. 59, N 2., 2013.

[5] S. Hochmuth, T. Brand, M. Zokoll, F. Zenker, N. Wardenga, B. Kollmeier, A Spanish matrix sentence test for assessing speech reception thresholds in noise. Int. J. Audiol. 51(7) pp. 536-544, 2012.

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

[7] Z. Milivojević, D. Kostić, D. Brodić,"Performanse razumljivosti Srpskog MST-a u prisustvu akustičkog Gausovog šuma", INFOTEH Jahorina, 2017.

[8] 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.

[9] I. Arweiler, J. M. Buchholz, "The influence of spectral characteristics of early reflections on speech intelligibility", J.Acoust.Soc.Am.130 (2), pp. 996-1005, 2011.

[10] L. L. Beranek, T. J. Mellow, "Acoustics: Sounde Fields and Transducers", 2012.

[11] J. Lochner, F. Burger, "The influence of reflections on auditorium acoustics", J Sound Vib. 1 (1964).

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

[13] A. K. Nabelek, J. M. Pickett "Monaural and binaural speech perception through hearing aids under noise and reverberation with normal and hearing-impaired listeners", J Speech Hear. Res. 7 (1974).

[14] A.Warzybok, J. Rennie, T.Brand, S.Doclo, B. Kollmeier, "Effect of spatial and temporal integration of a single early reflection on speech intelligibility", J.Acoust.Soc.Am.133(1), pp. 269-282, 2013

[15] Z. Milivojević, D. Kostić, Z. Veličković, "The optimization of the STOI algorithm parameters in presence of the WGN", ICEST 2018, Sozopol, Bulgaria.