The Evaluation of Speech Intelligibility in the Orthodox Church on the Basis of MOS Test Intelligibility Logatom Type CCV

Dijana Kostić[1], Zoran Milivojević[2] and Violeta Stojanović [3]

Abstract **– This paper presents the evaluation of the intelligibility of speech in the Orthodox Church on the basis of MOS test intelligibility logatom type CCV in the presence of Babble noise, for some values of the SNR. First part of the paper described the experiment and the results of the experiment are shown in graphical and tabular form. The second part is an analysis of the results and a comparison with the International Standard IEC 60268 – 16. On the basis of comparative analysis was shown that the intelligibility of speech in the Orthodox Church is poor.**

Keywords – **intelligibility of speech, the room impulse response (RIR), Babble noise, MOS, intelligibility of logatom.**

I. Introduction

From the aspect of information content the speech signal is the most important acoustic signal. It holds a huge amount of information and is therefore very important to ensure transmission quality and understanding of the messages that carries [1].

French, Steinberg and Beranek 1947 were the first who highlight the problem of transmission and speech intelligibility [2]. Kryten in 1962 introduced the articulation index, *AI*. Peutz in 1971 has developed an algorithm to predict the intelligibility of speech in auditoriums and suggested parameter which refers to the loss of articulation of consonants, Al_{cons} [3]. Houtgast and Steeneken in 1980 confirmed an objective method to measure the quality of voice transmission in the room [2], and suggested the acoustic parameters of the speech transmission index, *STI* [4]. Subjective systems evaluation for transmitting voice using open and closed tests intelligibility, were first performed by Fletcher and Steinberg in 1929 [2], then Egan in 1944, Miller and Nicely in 1955, House in 1965 and Voiers in 1977 etc. Tests intelligibility was represented by words test (usually with logatoms (monosyllabic words without meaning) type CVC, VCV, CV, VC, CCVC, CVCC, (C - consonant, V- vocal)), test sentence or test of syllables. Review of tests for evaluating intelligibility of speech was given by Pols in 1991 and Steeneken in 1992.

The following factors have an influence on the speech intelligibility [1]: the width of the band, distortion in the figurative sense, the strength of speech, ambient noise and Babble noise, reverberation and influence direct and reflected sound. Babble noise is a consequence at the same time speech *N* persons (BNN) and its presence leads to the degradation of the useful acoustic signal which is a measure of the Signal-to-Noise Ratio SNR [5]. In the given physical circumstances intelligibility of speech depends on: the content of speech, speech production and precision of diction speaker, listener's concentration in speech, age, etc.

Intelligibility of speech in a language implies intelligibility logatom. The correlation between intelligibility logatom and intelligibility of speech is shown in Table I [6].

TABLE I
THE CORRELATION BETWEEN THE INTELLIGIBILITY LOGATOMS (IL) AND SPEECH INTELLIGIBILITY (SI).

***IL* (%)**	0÷34	34÷65	65÷85	85÷95	95÷100
***SI* (%)**	bad	poor	fair	good	excellent

In a speech different phonemes (sounds) carry a different energy. Higher participation in speech (in terms of time excuse) has vocals than consonants. Also, they carry more energy than consonants. However, for intelligibility speech more important are consonants.

There are 30 phonemes (votes) [7] in the Serbian language. Phonemes by their common traits can be divided into groups: a) according to the noise level: silent (to the vocals, there are 5 and the noise (to the consonants, there are 25), b) according to the frequency of occurrence in the Serbian language phonetically balanced phoneme type CVC and the type of CCV and c) uniformly balanced logatoms all of which are equally represented in the Serbian language and forms are CVCV. The vocals are audible, while consonants may be audible and voiceless.

This paper presents the evaluation of speech intelligibility in the Orthodox Church on the basis of Mean Opinion Score (MOS) test intelligibility CCV logatom, in the presence of Babble noise, for some values of the SNR. The measurement was performed by creating a base of logatom, the impulse response of the church, the church of simulated acoustic signals and simulated acoustic signal with superimposed Babble noise the predefined relationships SNR. The effect of the acoustic environment of the church was simulated, using Matlab, using convoluting speech signal with recorded impulse responses of the church. Simulation of the effect of people who speak and disrupt the original signal is generated

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

by superposition Babble noise and acoustic signals generated church. Testing was performed by the open MOS test and listeners were wrote down reproduced logatoms as they hear them. After analyzing the results and comparisons with the International Standard IEC 60268-16: 2011 brings to a conclusion of the evaluation of speech intelligibility in the church.

The organization of work is as follows: Section II explains the experiment and the results are presented. Section III presents the analysis results of the test MOS intelligibility CCV logatom basis on MOS test intelligibility CCV logatom in the presence of Babble noise, for some values of the SNR. Section IV is the conclusion.

II. Experiments

In the paper we made the evaluation of speech intelligibility for the Orthodox Church "Holy Martyr Procopius" in Katun (Nis, Serbia) using MOS test intelligibility logatom type CCV, in the presence of Babble noise for SNR = {-10, -5, 0, 10, 20, 30}dB.

Volume and the total internal surface of the Orthodox Churches are V = 1659.68 m^3 and S = 646.48 m^2. Reverberation time RT = 2.06 s.

For intelligibility of the MOS test signals are formed which block diagram shown in Fig. 1 where: **x** pure speech signal, **h** the impulse response of the church, **y** generated acoustic signal and **z** generated acoustic signal with superimposed noise Babble **BNN**.

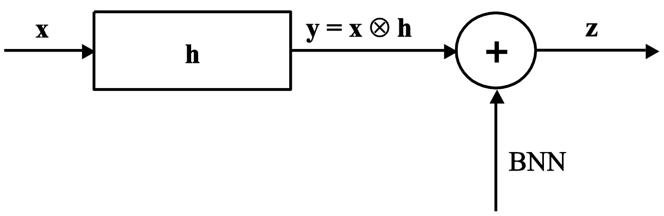


Fig. 1. Block diagram of the signal formed by MOS test intelligibility logatom type CCV

A MOS intelligibility test was conducted with 30 listeners (15 female and 15 male, age 18 ÷ 25), which are recorded phonetically balanced logatoms output from the player. The MOS test contains a list of 3 types of phonetically balanced logatom CCV (Table II). The results are analyzed on the basis of correct and incorrect written answers.

Intelligibility logatom, *LI*, in a measurement point for each MP for each listener in the presence of Babble noise with specific values SNR is calculated according to Eq. 1:

$$LI(\%) = \frac{NEWL}{NLS} \cdot 100, \tag{1}$$

where: *NEWL* - the number exactly written logatom and *NLS* - the number of words spoken logatom. Calculated the average value of each type of intelligibility logatom and all logatom for all listeners in the MP for a specific value SNR.

A. The Basis

The base of the experiment consists of: **1)** the base of the original signal **x**: recorded 6 speakers (3 men and 3 women, age 18 ÷ 25) who read test made the list of the three types phonetically balanced logatom CCV of 30 words; **2)** a base impulse response of the church: the database includes wav files that were obtained by recording the acoustic impulse response using the software package EASERA. Recordings were made at measuring point MP, which is shown in Fig.2.

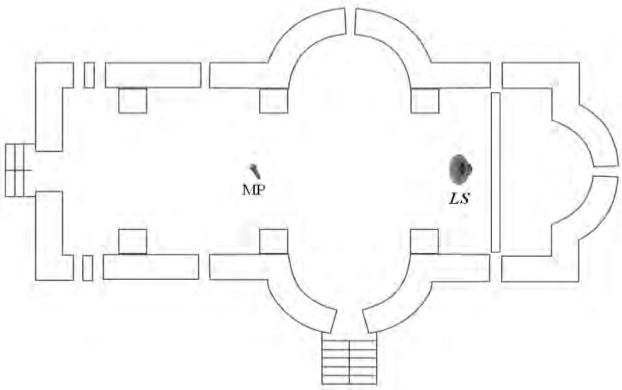


Fig. 2. The position of measuring point MP and sound source LS in the church during the recording impulse response.

Measuring point MP is 7m away from the sound source LS. 7 measurements were carried out, making a total of 7 wav files. Impulse response measurement was performed by the excitation sweep signal for a period of 6 s. The Sampling frequency is f_s = 44.1 kHz. The measurement of the impulse response of the church was performed in accordance with ISO 3382 [8]. **3)** the base of a simulated acoustic signal: the effect of the acoustic environment of the church was simulated, using Matlab, convolution (signal **y**) of the speech signal from the base **x**, with recorded impulse responses of the church to the measuring point MP. **4)** base of Babble noise BN: Babble noise was recorded with 8 speakers (BN8), 4 male and 4 female (age 18 ÷ 25) (signal **BNN**) and **5)** of the base of a simulated acoustic signal with superimposed BN8 noise for SNR = {-10, -5, 0, 10, 20, 30} dB, (signal **z**).

B. The Results

Table II shows the intelligibility logatom CCV type 1, type 2 and type 3. Table III shows the middle values of intelligibility logatom CCV type 1, type 2, type 3 and the middle value of the intelligibility of all types logatom observed for pure signal to the input signal x and the impulse response of the church, y. Figs. (3 – 5) shows the intelligibility of CCV logatom type 1 SNR = {- 10, 10} dB, type 2 for SNR = {- 5, 20} dB and type 3 for SNR = {0, 30} dB, respectively. Fig. 6. shows the intelligibility of all analyzed CCV logatom on these values SNR. On figures at the abscissa are ordinal numbers that are shown in Table II.

TABLE II
INTELLIGIBILITY CCV LOGATOM.

CCV type 1										
ord. num.	**1.**	**2.**	**3.**	**4.**	**5.**	**6.**	**7.**	**8.**	**9.**	**10.**
logatom	BLE	ŠLJA	DNJE	BRI	VRE	PLJE	RNO	DRU	TRA	GDI
x (%)	66.67	77.78	100	100	100	22.23	100	100	100	88.89
logatom	ČNU	SMI	TNI	FRA	PRO	SKO	SKE	PRI	VRA	STU
y (%)	0	0	55.56	77.78	55.56	77.78	55.56	77.78	88.89	22.23
logatom	SPO	BRE	PRU	TRO	GLE	GRU	BRA	TNA	VNE	ŠKO
$z\|_{SNR=-10dB}$	0	0	11.11	0	0	0	11.11	0	0	0
$z\|_{SNR=10dB}$	0	44.44	55.56	88.89	55.56	66.67	55.56	33.33	0	11.11
CCV type 2										
ord.num.	**1.**	**2.**	**3.**	**4.**	**5.**	**6.**	**7.**	**8.**	**9.**	**10.**
logatom	SHO	DRE	KSI	ČKU	BRE	HVA	ZNA	ŠKE	KNI	TLA
x (%)	81.82	100	100	81.82	100	90.91	100	90.91	90.91	81.82
logatom	KRI	VNO	DRA	SLO	ČNO	SLI	SKI	PRU	SJI	DRI
y (%)	54.55	9.09	72.73	9.09	18.18	45.45	54.55	18.18	0	0
logatom	ŠLU	PNO	DRO	GRA	DRU	TNE	DRA	SRE	PRE	SNA
$z\|_{SNR=-5dB}$	0	0	54.55	18.18	45.45	0	45.45	18.18	45.45	0
$z\|_{SNR=20dB}$	0	0	45.45	45.45	63.64	0	63.64	54.55	81.82	63.64
CCV type 3										
ord.num.	**1.**	**2.**	**3.**	**4.**	**5.**	**6.**	**7.**	**8.**	**9.**	**10.**
logatom	TRA	ŽDO	TRI	SPI	KTE	ČNA	KLE	PNE	SVA	ZLA
x (%)	70	30	90	80	80	30	60	40	90	100
logatom	TVA	MNE	SLU	TSO	SPE	DNE	SRA	SNO	SKO	ŠLU
y (%)	10	0	0	30	50	0	20	0	20	10
logatom	STE	DNO	PRA	ŠTE	KSA	MNA	SMA	PRI	STI	SVE
$z\|_{SNR=0dB}$	0	0	20	0	0	0	0	20	0	20
$z\|_{SNR=30dB}$	70	0	40	50	0	0	0	50	40	10

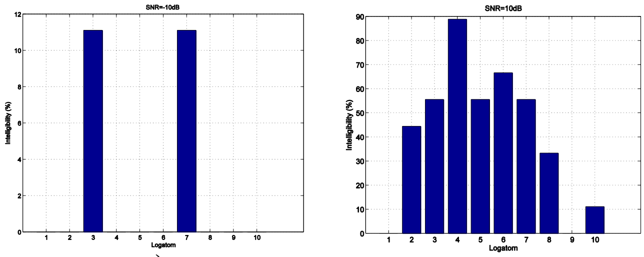
TABLE III
THE MIDDLE VALUE INTELLIGIBILITY CCV LOGATOMS.

Signal	**SNR (dB)**	**Type of logatom**			
		CCV type 1	**CCV type 2**	**CCV type 3**	$\overline{CCV}$
x (%)	∞	85.56	91.82	67	81.46
y (%)	∞	51.11	27.27	14	30.79
z (%)	-10	1.11	/	/	
	-5	/	22.73	/	
	0	/	/	6	
	10	41.11	/	/	
	20	/	41.82	/	
	30	/	/	26	

III. THE RESULTS ANALYSIS

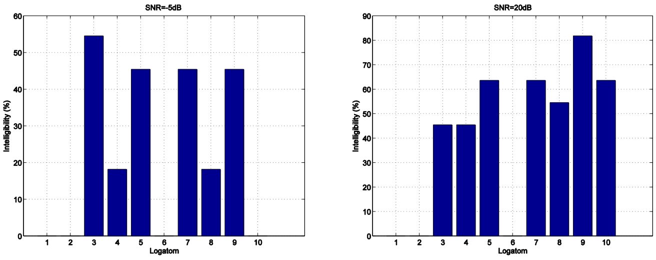
1. Based on the results for the user CCV logatom after signals (Table II and Figures (3 - 6) concluded that: a) 46.67% logatom CCV presented by pure speech signal has excellent, at 23.33% good, and fair, and 6.67% have a bad intelligibility; b) 60% logatom represented CCV the input signal with the impulse response of the church has a bad, 23.33% poor, 13.33% fair and 3.33% have a good intelligibility; c) all CCV logatoms represented by the input signal with the impulse response of the church and Babble noise with SNR = {-10, -5, 0 dB} have the highest percentage of bad intelligibility: 100%, 60% and 100%, respectively; 40% of these signals with the SNR = -5 dB are poor intelligibility; d) CCV logatoms signal presented in the impulse response of the church and Babble noise with SNR = 10 dB has 10% good and fair intelligibility a 40% poor and bad intelligibility; e) logatoms signal presented to the impulse response of the church and Babble noise with SNR = 20 dB have the highest percentage of poor intelligibility, 60%, followed by bad with 30% and fair intelligibility with 10%; f) logatoms signal presented in the impulse response of the church and Babble noise with SNR=30dB has the highest percentage of poor intelligibility, 50%, then bad, 40% and fair intelligibility only 10%.

2. Based on the middle value intelligibility logatom CCV by signals (Table III.) concludes that: a) for all CCV logatoms presented pure input signal intelligibility is fair (81.46%). The only logatoms CCV type 2, with middle value intelligibility



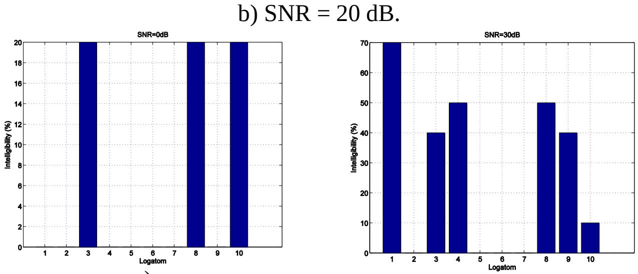
a) b)

Fig. 3. Intelligibility CCV logatom type 1: a) SNR = -10 dB and b) SNR = 10 dB



a) b)

Fig. 4. Intelligibility CCV logatom type 2: a) SNR = -5 dB and b) SNR = 20 dB.



a) b)

Fig. 5. Intelligibility CCV logatom type 3: a) SNR = 0 dB and b) SNR = 30 dB.

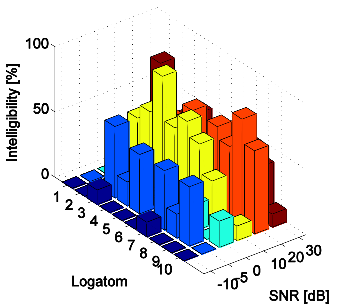


Fig. 6. Intelligibility CCV logatom for SNR = {-10, -5, 0, 10, 20, 30} dB.

that belongs to the good intelligibility while logatoms CCV type 1 have a fair intelligibility (85.56%), and logatoms CCV type 3 have a bad intelligibility (67%); b) middle value intelligibility logatom CCV type 1 (51.11%), type 2 (27.27%) and type 3 (14%) represented by the input signal with the impulse response of the church belongs to the values of bad intelligibility. This has the effect that the middle value of intelligibility of all these signals (30.79%) belongs to the range of values which classifies the signal into bad intelligibility.

3. All CCV logatoms presented with impulse response signals and Babble noise with one of the selected values SNR = {-10, -5, 0, 10, 20, 30} dB have a middle value of intelligibility, which puts them in a bad intelligibility. The highest percentage of bad intelligibility is presented logatoms signals with SNR = 20 dB, 41.82%. The lowest percentage of bad intelligibility is presented logatoms signals with SNR = -10 dB, 1.11%.

Based on the results of the experiment it is concluded that the greatest value of the percentage of intelligibility CCV logatom presented with a pure speech signal, 46.67%, classifies speech intelligibility in the Orthodox Church as an excellent intelligibility. However, already the largest percentage of intelligibility CCV logatom presented through the signals from the impulse response of the church, 60% classify speech intelligibility in the Orthodox Church as a bad intelligibility. Bad speech intelligibility in the Orthodox Church confirmed by the percentage of intelligibility logatom presented with impulse response signals and Babble noise with values SNR = {-10, -5, 0, 10, 20, 30} dB, which was expected.

IV. Conclusion

This paper is based on a comparison of the results obtained MOS test for the evaluation of subjective intelligibility logatom CCV type in the presence of Babble noise from the Orthodox Church and the comparable value for the intelligibility logatom and intelligibility of speech given to the International Standard IEC 60268-16.

Based on the above performed analysis showed that CCV logatoms represented signals with impulse response and Babble noise with values SNR = {-10, -5, 0, 10, 20, 30} dB have a middle value of subjective intelligibility 6% ÷ 41.82%. On this basis, it is concluded that the intelligibility of speech in the ambience of the Orthodox Church is *bad*. The reason for bad speech intelligibility in the ambience of the Orthodox Church is definitely the huge value of the reverberation time.

References

[1] H. Kuttruff, Room acoustics, fifth edition, Spoon Press, 2009.

[2] H. J. M. Steeneken, "On measuring and predicting speech intelligibility", Academisch Proefschrift, Soesterberg, 1992.

[3] Peutz, V. M. A. "Articulation Loss of Consonants as a criterion for Speech Transmission in a Room", J. AUDIO – ENG. SOC. Vol 19, p. 915 - 919, 1971.

[4] Steeneken, H. J. M., "Apsyhical Method for measuring Speech Transmission Quality", J. ACOUST. SOC. AM., Vol 19, 1980.

[5] N. Krishamurthy, Student member, IEEE, and John H.L. Hansen, Fellow IEEE, "Babble Noise: Modeling, Analysis and Applications", IEEE Transactions on audio, speech and language processing, vol 17, no.7 September 2009.

[6] 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.

[7] Lj. Subotić, D. Sredojević, I. Bjelaković, Fonetika i fonologija: ortoepska i ortografska norma standardnog srpskog jezika, Filozofski fakultet, Novi Sad, 2012.

[8] ISO 3382: 1997, Acoustic - Measurement of the Reverberation Time of Rooms with Reference to Other Acoustical Parameters, 1997.