

Procena subjektivne razumljivosti CVC logatoma u ambijentu pravoslavne crkve u prisustvu babble šuma

Zoran Milivojević, Violeta Stojanović, Zoran Veličković, Dijana Kostić
Visoka tehnička škola strukovnih studija
Niš, Srbija

zoran.milivojevic@vtsnis.edu.rs violeta.stojanovic@vtsnis.edu.rs zoran.velickovic@vtsnis.edu.rs koricanac@yahoo.com

Sažetak — U ovom radu izvršena je procena subjektivne razumljivosti govora u ambijentu pravoslavne crkve u prisustvu babble šuma za neke vrednosti odnosa signal/šum. U prvom delu rada su istaknute specifičnosti srpskog jezika i uticaj šuma na razumljivost govora. U drugom delu rada opisan je eksperiment, MOS test razumljivosti CVC logatoma, na osnovu koga je izvršena procena razumljivosti govora. Rezultati eksperimenta su prikazani tabelarno i grafički. Na osnovu izvršene komparativne analize donosi se zaključak o razumljivosti govora u ambijentu pravoslavne crkve.

Ključne reči - razumljivost govora; razumljivost logatoma; vreme reverberacije; babble šum; odnos signal/šum; MOS test.

I. UVOD

Razumljivost govora je važan parametar kako za prostornu akustiku (učionice, pozorišta, crkve i sl.), tako i za komunikacione sisteme (telefonija, radiofonija i sl.). Kvalitet audio sistema preko koga se prenosi govorni signal, ocenjuje se stepenom njegove razumljivosti, vernosti i prirodnosti [1].

Ocenjivanje subjektivne razumljivosti govora vrši se na osnovu razumljivosti logatoma (jednoznačne reči bez smisla) koja predstavlja standard za egzaktno merenje razumljivosti govora [2]. Za potrebe testiranja koristi se, najčešće, lista logatoma tipa CVC, CCV, VCV, CV, VC, CCVC, CVCC, (C – konsonant, V – vokal). Prva istraživanja subjektivne razumljivosti govora započeli su Fletcher i Steinberg, 1929. god., a zatim su ih sledili Egan 1944., Miller i Nicely, 1955., House 1965., Voiers 1977. i dr. Pregled testova ocenjivanja za razumljivost govora su dali Pols 1991. i Steeneken 1992.

U radu je data procena subjektivne razumljivosti govora u prisustvu babble šuma, za neke vrednosti odnosa signal/šum, SNR (*eng. Signal Noise Ratio*), u objektu pravoslavne crkve. Procena je vršena na osnovu rezultata MOS (*eng. Mean Opinion Score*) testa razumljivosti CVC logatoma. Postupak merenja je obavljen tako što su najpre formirane baze: a) logatoma, b) impulsnih odziva crkve, c) simuliranih akustičkih signala crkve i d) simuliranih akustičkih signala sa superponiranim babble šumom za unapred definisane odnose SNR. Efekat akustičkog ambijenta crkve simuliran je, pomoću programa Matlab, konvolucijom govornog signala sa snimljenim impulsnim odzivima crkve. Simulacija efekta ljudi koji govore i ometaju originalni signal ostvarena je superpozicijom babble šuma i generisanog akustičkog signala crkve. Testiranje slušalaca je vršeno MOS testom otvorenog tipa pri čemu slušalac beleži reprodukovane logatome onako

kako ih čuje. Nakon analize dobijenih rezultata i komparacije sa međunarodnim standardom IEC 60268-16:2011 donosi se zaključak o oceni razumljivosti govora u crkvi.

Organizacija rada je sledeća: u Sekciji II date su karakteristične osobine srpskog jezika i uticaj šuma na razumljivost govora. U Sekciji III prikazani su rezultati i analiza MOS testa razumljivosti logatoma tipa CVC. Sekcija IV je zaključak.

II. RAZUMLJIVOST GOVORA

Za uspešnu komunikaciju potrebno je da se ostvari odgovarajuća razumljivost kako bi kombinacija glasova u slogovima i rečima bila prepoznatljiva. Razumljivost govora je pojam koji označava tačnost kojom slušalac prima sadržaj govora koji sluša. Jedan od važnih parametara koji utiče na razumljivost govora kod prostorne akustike je vreme reverberacije [1]. Što je manje vreme reverberacije to je razumljivost govora veća.

Razumljivost govora zavisi od: a) fizičkog kvaliteta prenosa govornog signala kroz akustički prenosni sistem (sadržina govora, preciznost izgovora glasova i dikcija govornika, koncentracija slušaoca na govor), b) psihološkog aspekta (uznemiravanje, skretanje pažnje), c) fiziološkog aspekta (dejstvo prekomernog intenziteta zvuka sa fizičkim oštećenjima čula sluha) i d) aspekta preciznosti primljenih zvučnih informacija.

U govoru različiti fonemi (glasovi) nose različitu energiju. U srpskom jeziku postoje 30 fonema [3]. Oni se dele na: a) vokale (samoglasnici): a, e, i, o, u; b) konsonante, sonante-glasnike: v, j, l, lj, m, n, nj, r; i c) konsonante, opstruente-suglasnike: svi ostali glasovi. Fonemi mogu biti zvučni i bezvučni. Vokali su zvučni, dok konsonanti mogu biti i zvučni i bezvučni. Korelacija između razumljivosti logatoma i razumljivosti govora prikazana je u Tabeli I. [4]. Na razumljivost govora utiču konsonati. Pored njih na razumljivost govora utiču i šumovi.

Šum je nepravilno oscilovanje u kojem nema ni stalnih frekvencija ni stalnih amplituda. Dve specifične vrste signala šuma koje se često koriste u audiotehnici za testiranje uređaja i sistema su beli i roze šum, dok se za govor koristi babble šum.

Babble šum se karakteriše brojem N govornika od kojih potiče [5]. Smatra se da je jedan od najboljih šumova za maskiranje govora. U poređenju sa belim šumom, babble šum

TABELA I. KORELACIJA IZMEĐU RAZUMLJIVOSTI LOGATOMA I RAZUMLJIVOSTI GOVORA

Razumlj. logatoma (%)	0-67	67-78	78-87	87- 94	94-100
Razumlj. govora	Loša	Slaba	Prihvatljiva	Dobra	Odlična

nudi veću efikasnost kada je u pitanju kamufliranje glasa i mogućnost da se iskoristi maskiranje na nižim nivoima.

Ako se govor subjekata modeluje kao redosled izgovaranja glasova, babble šum koji dolazi od više subjekata može biti posmatran kao raslojavanje fonema i perioda tišine individualnih subjekata. Za ekperimentalne analize razumljivosti govora koristi se babble šum od većeg broja subjekata da bi se maskiralo izgovaranje logatoma i sagledao uticaj babble šuma na konačnu razumljivost govora.

Odnos signal/šum (S/N) jedan je od parametara koji utiče na razumljivost govora. On predstavlja meru koja se koristi u nauci i inženjerstvu za poređenje nivoa željenog signala i nivoa šuma u pozadini. Za beli šum je dovoljno da odnos nivoa govora i šuma bude svega nekoliko decibela da bi razumljivost bila zadovoljavajuća. Što je ovaj odnos veći razumljivost je bolja sve do vrednosti odnosa signal/ šum od 25 dB. Iznad ovog iznosa razumljivost ne zavisi od odnosa signal/šum. Za prostorije kod kojih je vreme reverberacije $T \geq 1,6$ s odnos S/N treba da bude najmanje 25 dB.

III. EKSPERIMENTALNI REZULTATI I ANALIZA

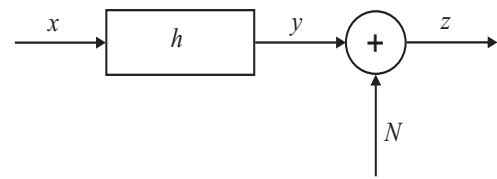
A. Eksperiment

U radu je izvršena ocena razumljivosti govora za pravoslavnu crkvu „Sv.Velikomučenik Prokopije“ u Katunu (Niš, Srbija) u prisustvu babble šuma za SNR = (-10, -5, 0, 10, 20, 30) dB primenom MOS testa razumljivosti logatoma tipa CVC.

Zapremina crkve je $V = 1659.68 \text{ m}^3$, ukupna unutrašnja površina $S = 646.48 \text{ m}^2$. Vreme reverberacije je $RT = 2.05$ s.

Blok dijagram signala korišćen za MOS test razumljivosti prikazan je na Sl. 1. gde je: x čist govorni signal, h impulsni

odziv crkve, y generisan akustički signal i z generisan akustički signal sa superponiranim babble šumom.



Slika 1. Blok dijagram formiranja signala za MOS test razumljivosti logatoma tipa CVC.

Merenje razumljivosti govora se sastoji u tome da se napravi lista logatoma koji se emituju sa uređaja za reprodukciju slušaocima. Na ulaznoj strani sistema prenosa izgovaraju se logatomi, a na prijemu se beleži prepoznati akustički signal. Razumljivost logatoma, LI , izračunava se prema formuli:

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

gde je: $NEWL$ – broj tačno zapisanih logatoma i NLS – broj izgovorenih logatoma.

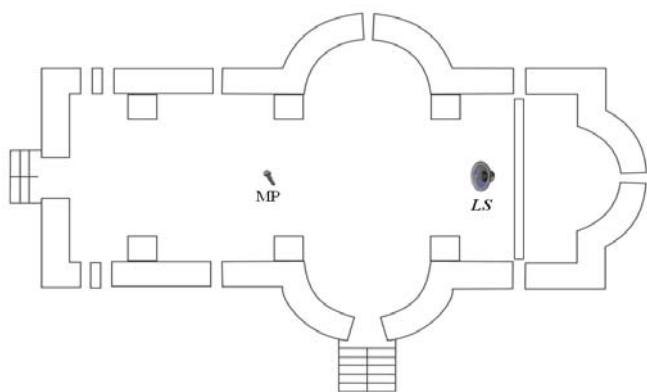
MOS test razumljivosti sproveden je sa 34 slušalaca (17 ženskih i 17 muških, godina starosti od 18 – 25) koji su zapisivali fonetski balansirane logatome koje čuju (Tabela II). MOS test sadrži listu od 3 tipa fonetski balansiranih logatoma CVC. Rezultati su sagledani na bazi tačnih i netačnih odgovora.

B. Baza

Bazu eksperimenta čine: **1)** baza originalnog signala x : a) napravljena je test lista od 90 reči (tri tipa fonetski balansiranih logatoma CVC od po 30 reči). b) snimljeno je 6 govornika (3 muška i 3 ženska, godine starosti od 18 – 25) koji su čitali test listu. **2)** baza impulsnih odziva crkve: bazu podataka čine wav fajlovi koji su dobijeni snimanjem akustičkih impulsnih odziva pomoću softverskog paketa EASERA. Snimanja su izvršena u mernoj tački MP koja je prikazana na Sl. 2.

TABELA II. TIPOVI TESTIRANIH FONETSKI BALANSIRANIH CVC LOGATOMA

red. br.	CVC tip 1			CVC tip 2			CVC tip 3		
	x	y	z SNR = (-10, 10) dB	x	y	z SNR=(-5, 20) dB	x	y	z SNR=(0, 30)dB
1.	KAN	FELJ	SAS	NIT	TES	NOK	HED	ČET	GAV
2.	VUD	TOC	NOJ	SAL	JER	POS	MENJ	TID	PIN
3.	MEN	TIŠ	VEN	NJEK	TAR	DUČ	TUK	ŽEĆ	PIT
4.	VAB	JAN	MEĐ	JEČ	KET	NAT	KIG	PED	ZAŠ
5.	DAV	ROZ	VAT	BIN	PIJ	NIV	JAV	JOR	TIJ
6.	PIL	GAD	ŠOR	DEV	DAM	LOZ	MOS	ŠER	MIS
7.	DIK	RON	VOD	NOG	RAJ	SUN	SEP	NAH	NAN
8.	ROJ	SET	LUN	SAP	LAR	ZOD	POL	MAR	TAJ
9.	VOK	RIDŽ	MUL	KER	DAG	VAN	KOT	KIS	DIR
10.	ŽIČ	ZONJ	JOS	POB	RIL	BEČ	SIS	LOR	NIS



Slika 2. Položaj merne tačke MP i zvučnog izvora LS u crkvi tokom snimanja impulsnog odziva.

Merna tačka MP je 7 m udaljena od zvučnog izvora LS. Izvršeno je 7 merenja, što ukupno čini 7 wav fajlova. Merenje impulsnog odziva izvršeno je pomoću pobudnog sweep signala u trajanju od 6 s. Frekvencija odabiranja je $f_s = 44.1$ kHz. Postupak merenja impulsnih odziva u crkvi obavljen je u skladu sa standardom ISO 3382 [6]. **3)** baza simuliranih akustičkih signala: efekat akustičkog ambijenta crkve simuliran je, pomoću programa Matlab, konvolucijom (signal y) govornog signala iz baze x , sa snimljenim impulsnim odzivima

crkve h u mernoj tački MP. **4)** baza babble šuma BN (eng. *Babble Noise*): snimljen je babble šum sa 8 govornika (BN8), 4 muška i 4 ženska (starosti od 18 do 25), (signal N). **5)** baza simuliranih akustičkih signala sa superponiranim BN8 šumom za SNR = (-10, -5, 0, 10, 20, 30) dB, (signal z).

C. Rezultati

U Tabeli III. prikazane su razumljivosti logatoma CVC tipa 1, tipa 2 i tipa 3. U Tabeli IV. prikazane su srednje vrednosti razumljivosti logatoma CVC tipa 1, tipa 2, tipa 3 i srednja vrednost razumljivosti svih tipova logatoma posmatrana za čist signal x i za ulazni signal sa impulsnim odzivom crkve, y .

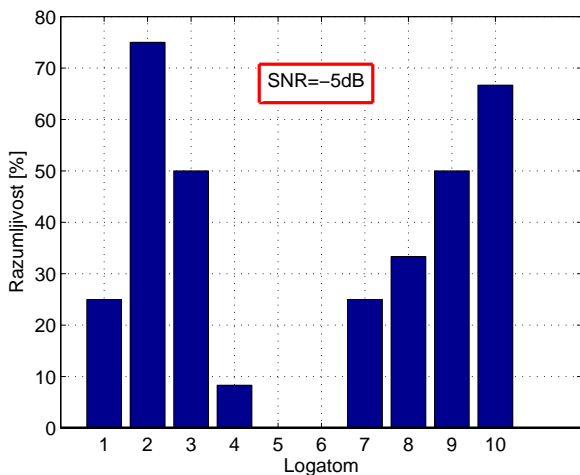
Na Slici 3. i Slici 5. prikazana je razumljivost CVC logatoma tipa 2 u zavisnosti od SNR = -5 dB i SNR = 20 dB respektivno. Na Slici 4. i Slici 6. prikazana je razumljivost CVC logatoma tipa 3 u zavisnosti od SNR = 0 dB i SNR = 30 dB respektivno. Na Slici 7. prikazana je zavisnost razumljivosti CVC logatoma tipa 2 i tipa 3 od SNR = (-5, 0, 20, 30) dB. Na ovim slikama na apscisama su redni brojevi logatoma koji su dati u Tabeli II.

TABELA III. RAZUMLJIVOST CVC LOGATOMA (%).

CVC tipa 1										
logatomi	KAN	VUD	MEN	VAB	DAV	PIL	DIK	ROJ	VOK	ŽIČ
x	50	80	40	30	20	50	80	60	80	80
logatomi	FELJ	TOC	TIŠ	JAN	ROZ	GAD	RON	SET	RIDŽ	ZONJ
y	0	0	0	30	60	20	0	50	0	0
logatomi	SAS	NOJ	VEN	MEĐ	VAT	ŠOR	VOD	LUN	MUL	JOS
z (SNR = -10dB)	20	0	0	0	0	0	0	0	0	0
z (SNR = 10dB)	70	0	50	10	60	0	20	0	10	20
CVC tipa 2										
logatomi	NIT	SAL	NJEK	JEČ	BIN	DEV	NOG	SAP	KER	POB
x	100	100	100	100	75	75	100	100	100	25
logatomi	TES	JER	TAR	KET	PIJ	DAM	RAJ	LAR	DAG	RIL
y	91.67	8.33	8.33	66.67	0	33.33	8.33	0	0	0
logatomi	NOK	POS	DUČ	NAT	NIV	LOZ	SUN	ZOD	VAN	BEČ
z (SNR = -5dB)	25	75	50	8.33	0	0	25	33.33	50	66.67
z (SNR = 20dB)	41.67	66.67	50	33.33	0	0	33.33	41.67	66.67	83.33
CVC tipa 3										
logatomi	HED	MENJ	TUK	KIG	JAV	MOS	SEP	POL	KOT	SIS
x	58.33	0	50	41.67	50	100	66.67	66.67	100	100
logatomi	ČET	TID	ŽEĆ	PED	JOR	ŠER	NAH	MAR	KIS	LOR
y	83.33	16.67	8.33	0	8.33	66.67	8.33	25	100	0
logatomi	GAV	PIN	PIT	ZAŠ	TIJ	MIS	NAN	TAJ	DIR	NIS
z (SNR = 0dB)	25	58.33	58.33	100	0	91.67	0	75	8.33	50
z (SNR = 30dB)	25	66.67	100	100	0	91.67	0	83.33	8.33	58.33

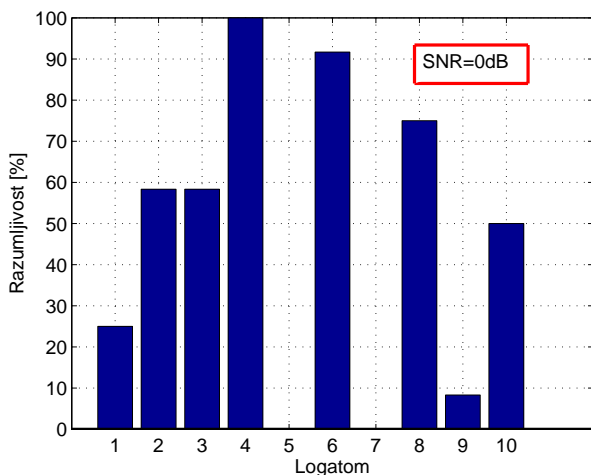
TABELA IV. SREDNJA VREDNOST RAZUMLJIVOSTI LOGATOMA CVC (%)

Signali	SNR (dB)	Vrsta logatoma			
		CVC tip1	CVC tip 2	CVC tip 3	\overline{CVC}
x	∞	57	87.50	63.33	69,27
y	∞	16	21.60	31.67	23,09
z	-10	2	/	/	
	-5	/	33.34	/	
	0	/	/	46.67	
	10	24	/	/	
	20	/	41.67	/	
	30	/	/	53.33	

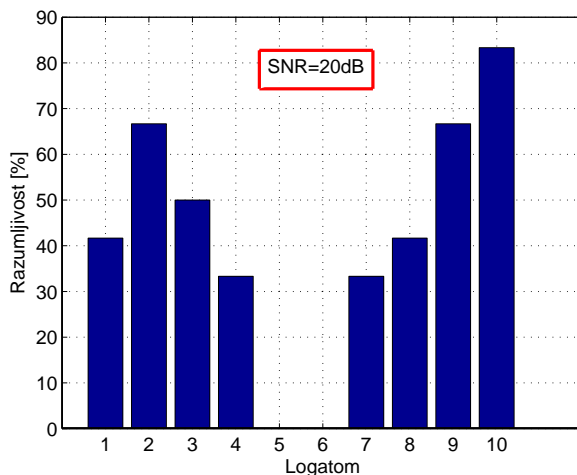


Slika 3. Razumljivost logatoma CVC tipa 2 za SNR = -5dB.

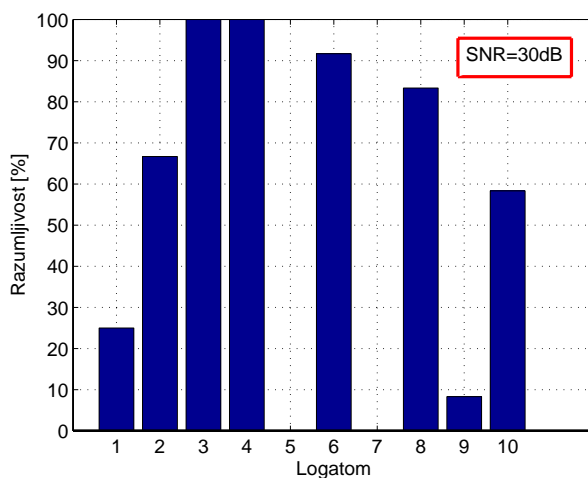
Slika 4.



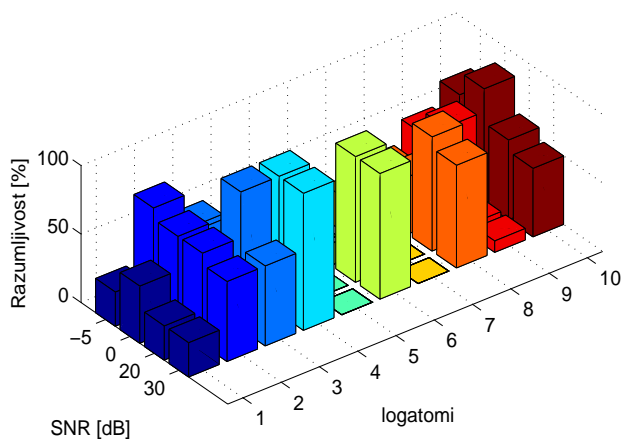
Slika 5. Razumljivost logatoma CVC tipa 3 za SNR = 0dB.



Slika 6. Razumljivost logatoma CVC tipa 2 za SNR = 20dB.



Slika 7. Razumljivost logatoma CVC tipa 3 za SNR = 30dB.



Slika 8. Zavisnost razumljivost logatoma tipa CVC od SNR-a.

D. Analiza rezultata

1. Na osnovu razumljivosti logatoma tipa CVC po signalima (Tabela III. i Slike (3 - 7) zaključuje se da: a) 33.33 % logatoma CVC predstavljeni čistim govornim signalom ima odličnu razumljivost, 13.33 % prihvatljivu, 6.67 % slabu a 46.67 % ima lošu razumljivost; b) čak 90 % logatoma CVC predstavljeni ulaznim signalom sa impulsnim odzivom crkve, ima lošu razumljivost, 3.33 % prihvatljivu, 6.67 % ima odličnu razumljivost; c) svi CVC logatomi predstavljeni ulaznim signalom sa impulsnim odzivom crkve i babble šumom sa SNR = (-10, -5, 0, 10, 20, 30) dB imaju najveći procenat loše razumljivosti: 100 %, 90 %, 70 %, 90%, 90 % i 60%, respektivno. d) 10% slabe razumljivosti imaju sigali sa impulsnim odzivom crkve i babble šumom sa SNR = (- 5, 0, 10) dB, e) prihvatljivu razumljivost ima 10 % signala sa impulsnim odzivom crkve i babble šumom sa SNR = (20, 30) dB; f) dobru razumljivost ima 10 % logatoma CVC prikazanih signalom sa impulsnim odzivom crkve i babble šumom sa SNR = (0, 30) dB i g) odličnu razumljivost ima 10 % i 20% logatoma CVC prikazani signalom sa impulsnim odzivom crkve i babble šumom sa SNR = (0, 30) dB, respektivno.

2. Na osnovu srednje vrednosti razumljivosti logatoma CVC po signalima (Tabela IV.) zaključuje se da: a) je za sve CVC logatome koji su predstavljeni čistim ulaznim signalom razumljivost slaba (69.27%). Naime, samo logatomi CVC tip 2, imaju srednju vrednost razumljivosti koja pripada grupi dobre razumljivosti (procenat srednje razumljivosti im je 87.5 %) dok ostali logatomi imaju srednju vrednost razumljivosti koja ih svrstava u loše razumljive (57 % logatomi CVC tip 1 i 63.33 % logatomi CVC tip 3); b) srednja vrednost razumljivosti logatoma CVC tip 1 (16%), tip 2 (21.6 %) i tip 3 (31.67 %) koji su predstavljeni ulaznim signalom sa impulsnim odzivom crkve pripada vrednostima loše razumljivosti. To ima za posledicu i da srednja vrednost razumljivosti za sve ove signale (23.09 %) pripada opsegu vrednosti koja klasifikuje ovaj signal u loše razumljive.

3) svi CVC logatomi predstavljeni signalima sa impulsnim odzivom i babble šumom sa nekom od izabranih vrednosti SNR = (-10, -5, 0, 10, 20, 30) dB imaju srednje vrednosti razumljivosti koje ih svrstavaju u loše razumljive. Najveći procenat razumljivosti imaju signali sa SNR = 30 dB, 53.33% , a najmanju signali sa SNR = -10 dB, 2 %.

Sagledavanjem rezultata eksperimenta zaključuje se da najveća vrednost procenta razumljivosti CVC logatoma predstavljenih pomoću čistog govornog signala 46.67 %, klasifikuje razumljivost govora u pravoslavnoj crkvi kao lošu razumljivost. Zbog toga je bilo i očekivano da će razumljivost CVC logatoma analizirana preko signala sa impulsnim odzivom i babble šumom sa vrednostima SNR = (-10, -5, 0, 10, 20, 30) dB biti u opsegu vrednosti koji pripada lošoj razumljivosti.

IV. ZAKLJUČAK

Na osnovu komparacije uporednih vrednosti za razumljivosti logatoma i govora datih Međunarodnim standardom IEC 60268 – 16 i rezultata dobijenih MOS testom za procenu subjektivne razumljivosti logatoma tipa CVC u prisustvu babble šuma za pravoslavnu crkvu zaključuje se da najveći procenat razumljivosti, 53.33 %, ima signal sa impulsnim odzivom i babble šumom sa SNR = 30 dB koji predstavlja logatome CVC tipa 3. Međutim, ovaj procenat razumljivosti pripada opsegu loše razumljivosti.

Ovom analizom pokazano je da je razumljivost logatoma tipa CVC, a samim tim i razumljivost govora, u ambijentu srpske pravoslavne crkve loša, što se i očekivalo zbog velike vrednosti vremena reverberacije $RT = 2.05$ s.

LITERATURA

- [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] Lj.Subotić, D.Sredojević, I.Bjelaković, "Fonetika i fonologija:ortoepska i ortografska norma standardnog srpskog jezika", Filozofski fakultet, Novi Sad.
- [4] 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.
- [5] Nitish Krishnamurthy, 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] ISO 3382: "1997, Acoustic - Measurement of the Reverberation Time of Rooms with Reference to Other Acoustical Parameters", 1997.

ABSTRACT

In this paper we made an evaluation of subjective intelligibility of speech in the ambience of the Orthodox Church in the presence of babble noise for some values of the signal noise ratio. First part of the paper highlights the specifics of the Serbian language and the influence of noise on speech intelligibility. Second part of the paper describes the experiment, MOS test intelligibility of CVC logatom, on which basis is made an evaluation of intelligibility. The results of the experiment are presented in tables and graphs. Based on the comparative analysis we have brought the conclusion of the speech intelligibility in the ambience of the Orthodox Church.

EVALUATION OF THE SUBJECTIVE SPEECH INTELLIGIBILITY OF LOGATOMS CVC IN THE AMBIENCE OF ORTHODOX CHURCH IN THE PRESENCE OF BABBLE NOISE

Zoran Milivojević, Violeta Stojanović, Zoran Veličković,
Dijana Kostić