

Performanse razumljivosti Srpskog MST-a u prisustvu akustičkog Gausovog šuma

Zoran Milivojević, Dijana Kostić

Visoka Tehnička Škola Strukovnih Studija
Niš, Srbija

zoran.milivojevic@vtsnis.edu.rs, koricanac@yahoo.com

Darko Brodić

Univerzitet Beograd, Tehnički fakultet u Boru
dbrodic@tf.bor.ac.rs

Sažetak — U prvom delu rada opisana je SMST baza koja se sastoji od rečenica izgovorenih na srpskom jeziku. SMST baza namenjena je testiranju razumljivosti reči i rečenica kod osoba sa normalnim i neoštećenim sluhom. U drugom delu rada opisan je eksperiment koji je sproveden u cilju utvrđivanja razumljivosti rečenica u prisustvu akustičkog Gausovog šuma. Rezultati su prikazani tabelarno i grafički. Na kraju su prikazani komparativni rezultati razumljivosti govornog signala degradiranog Gausovim i Babble šumom.

Ključne reči: Razumljivost govora; SMST; Gausov šum; Babble šum.

I. UVOD

Kod savremenih multimedijalnih informacionih komunikacionih sistema važan zadatak je preneti govorni signal. Prilikom snimanja i prenosa javljaju se različite degradacije govornog signala (akustičke ambijentalne smetnje prilikom snimanja, kao i kod slušanja, nelinearnosti na prenosnom putu, superponiranje različitih vrsta šumova i sl.), tako da na prijemnoj strani govorni signal može biti degradiran u tolikoj meri da ne bude razumljiv. Zbog toga je potrebno za razne govorne signale, kao i za smetnje koje se javljaju, vršiti testiranje razumljivosti.

Za potrebe testiranja razumljivosti govornog signala razvijeni su različiti testovi: a) testiranje logatomima i b) test rečenicama (svakodnevne rečenice i matrični rečenični test) u zavisnosti od različitih parametara (različitih reverberacionih uslova, odnosa signal/šum, različitih vrsta šuma i sl.). Rezultati ovih testova daju svoj doprinos u različitim oblastima: elektronici, lingvistici, medicini ...

Logatomi test predstavlja testiranje razumljivosti govora pomoću jednoznačnih reči bez smisla. U srpskom jeziku postoje nekoliko tipa logatoma: Afrikati, Frikativi, Plozivi, Nazali, Laterali, CVC i CCV logatomi. Za ispitivanje razumljivosti govora najčešći korišćen tip logatoma je CVC ili CCV [1],[2].

Rečenični testovi se dele na: a) testove koji koriste svakodnevne rečenice (Plomp rečenični test [3]) i b) rečenice koje se formiraju iz baze reči organizovane u formi matrice gde su kolone tipovi reči (ime, glagol, broj, pridev i objekat), takozvan matrični rečenični test.

Test koji koristi svakodnevne sastoji se od kratkih svakodnevnih rečenica. Nedostatak ovog testa je što se rečenice mogu lako zapamtiti ili se može naslutiti značenje reči iz konteksta rečenice.

Matrični rečenični testovi su testovi sa fiksnom sintaksnom strukturom (ime, glagol, broj, pridev i objekat), ali zbog velikog broja kombinacija reči, rečenica ne mora uvek biti smisljena. Matrični testovi razvijeni su za sledeće jezike: švedski [4], španski [5], francuski [6], engleski [7], poljski [8].

Za potrebe testiranja reazumljivosti rečenica i reči, autori su formirali SMST bazu (**eng.** *Serbian Matrix Sentence Test*), po ugledu na originalni matrični test koji je razvio Hegerman 1982 za švedski jezik [4].

U ovom radu izvršen je eksperiment sa ciljem ocene performansi SMST-a (razumljivost reči i rečenica) u prisustvu akustičkog Gausovog šuma. Eksperiment je organizovan tako da su po slučajnom zakonu formirane rečenice iz SMST baze kojima je superponiran Gausov šum. Nakon toga je uz pomoć degradiranog govornog signala za SNR= {-10, -9, -8, -7, -5, -2, 0, 2} dB izvršen MOS test razumljivosti.

Testiranje je obavljeno nad testnom grupom koja je formirana od studenata Visoke Tehničke Škole Strukovnih Studija u Nišu. Testiranje slušalaca sprovedeno je MOS testom otvorenog tipa pri čemu su slušaoci glasno ponavljali reprodukovanu rečenicu onako kako su je čuli. Na osnovu komparativne analize rezultata sa Babble šumom i komparacije sa Međunarodnim standardom IEC 60268-16:2011 izveden je zaključak o razumljivosti govora.

Organizacija rada je sledeća. U sekciji II opisana je baza SMST-a i njena fonetska struktura. Sekcija III opisuje eksperiment, rezultate ekperimenta i analizu istih. U sekciji IV se na osnovu komparativne analize rezultata dolazi do zaključka.

II. SMST BAZA

Matrični test formiran je od 10 rečenica jednake sintaksne strukture: ime, glagol, broj, pridev i objekat, obrazujući na ovaj način matricu $N=10^5$. Kombinacijom elemenata redova i kolona matrice možemo dobiti 100000 novih rečenica.

SMST baza koja sadrži 10 imena, 10 glagola, 10 brojeva, 10 prideva i 10 objekata, prikazana je u Tabeli I [9].

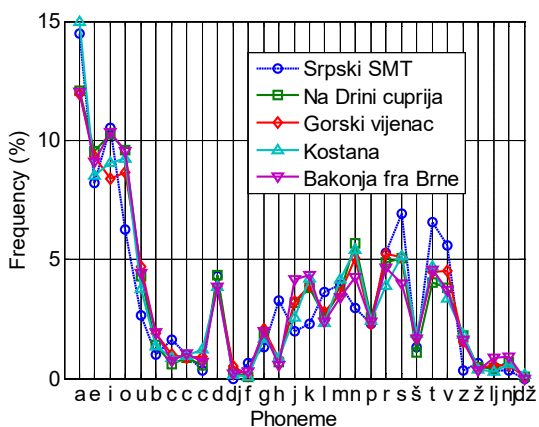
TABELA I. SRPSKI MATRIČNI TEST

Ime	Glagol	Broj	Pridev	Imenica
Slaviša	čuva	osamsto	skupih	knjiga
Tamara	kupuje	sedamdeset	crvenih	jabuka
Gordana	ima	šesnaest	čistih	stolova
Tomislav	pravi	petnaest	velikih	kuća
Jasmina	želi	deset	jeftinih	rukavica
Miodrag	vidi	devet	različitih	brodova
Marina	daje	osam	žutih	stolica
Dragoslav	prodaje	sedam	novih	fotelja
Miroslav	briše	šest	starih	ormara
Danica	voli	pet	lepih	cvetova

Govorni materijal i fonetska struktura

Govorni materijal formiran je iz baze Srpskog matričnog testa prikazanog u Tabeli I. Računar je prema slučajnom zakonu birao elemente matrice i formirao rečenicu, čija je kombinacija bila nepoznata i ispitivaču i slušaocu.

Procenat učestanosti pojavljivanja fonema prikazana je na Sl. 1. Fonetska struktura Srpskog MST prikazana je grafički sa (-o-). Za analizu fonetske učestalosti fonema, kao reference, analizirani su: romani "Na Drini Čuprija", autora Ive Andrića (-□-) i "Bakonja fra Brne" Sime Matavulja (-∇-), ep "Gorski vjenac" (-◇-), autora Petra Petrovića Njegoša i drama "Koštana" (-Δ-), autora Bore Stankovića.



Slika 1. Fonetska učestanost pojavljivanja fonema Srpskog MST (-o-), "Na Drini čuprija" (-□-), "Gorski vjenac"(-◇-), "Koštana"(-Δ-), (-∇-) "Bakonja fra Brne"

Grafički prikaz rezultata na Sl. 1, pokazuje da postoji velika korelacija učestanosti pojavljivanja foneme. Na osnovu ovoga može se smatrati da SMST baza ima dobru fonetsku strukturu.

III. EKSPERIMENTALNI REZULTATI I ANALIZA

Ekperimentalnim putem ispitana je razumljivost rečenica iz SMST baze u prisustvu akustičkog Gausovog šuma.

3.1. Eksperiment

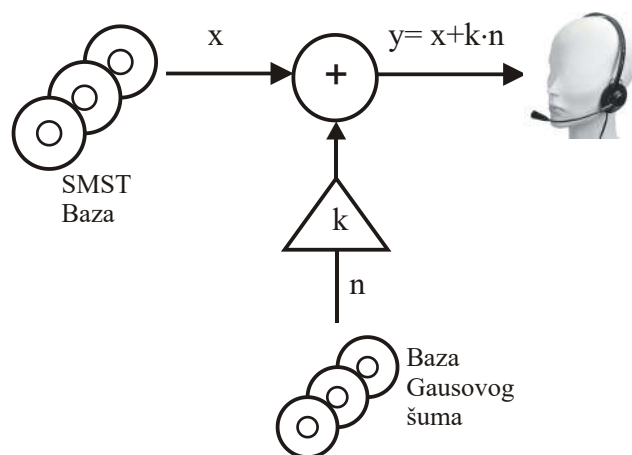
Eksperiment se sprovodi nad govornim signalom x , materijala iz SMST baze kome je superponiran Gausov šum n

(Sl.2). Generisani govorni signal $y=x+k\cdot n$ se formira za različite odnose signal/šum i to u opsegu $SNR = -10\div 2$ dB.

U drugom delu eksperimenta sproveden je MOS test nad testnom grupom. Test je sproveden u akustički izolovanom prostoru, a slušaoci su govorni signal slušali preko slušalica. Programski se prema slučajnom zakonu formira jedna od 100000 različitih rečenica. Slušalac glasno ponavlja šta je čuo, odnosno ono što je on razumeo. Ispitivač tabelarno beleži ispravnost odgovora na nivou rečenice i na nivou pojedinih vrsta reči (imena, glagola, broja, prideva i objekata).

Razumljivost je ocenjivanja na nivou razumljivosti reči i rečenica. Samo pravilno ponovljenja reč kao i cela rečenica se računala u tačan odgovor. Nad dobijenim rezultatima sprovedena je statistička analiza sa ciljem određivanja razumljivosti reči i rečenica. Razumljivost je predstavljena procentualno u opsegu od 0÷100%.

Na kraju analize razumljivosti prikazana je komparativna analiza razumljivosti govora u prisustvu Gausovog (iz ovog rada) i Babble šuma (rezultati iz [10]). Komparativnom analizom razumljivosti u prisustvu Gausovog šuma i Babble šuma zaključuje se da je razumljivost manja kod Babble šuma. Ovo je očekivani rezultat imajući u vidu raspodele energije u spektru gde je energija Babble šuma koncentrisana u opsegu od 100÷1000Hz, gde se inače nalaze značajne komponente analiziranog govornog signala, dok je energija Gausovog šuma ravnomerno raspoređena u celom frekventom opsegu.



Slika 2. Blok šema generisanog signala y , formiranog za testiranje razumljivosti MOS testom.

Na Sl. 2 grafički je prikazana blok šema generisanog govornog signala y . Čist govorni signal generisan iz SMST baze predstavljen je sa x , signal n prestavlja generisan Gausov šum dobijen iz baze Gausovog šuma, a koeficijent za određivanje vrednosti parametra SNR obeležen je sa k .

3.2. Baze

U eksperimentu se koriste:

a) SMST baza sa rečima izgovorenim na srpskom jeziku, formirana na Visokoj Tehničkoj Školi Stukovnih Studija u Nišu.

b) Baza Gausovog šuma (srednja vrednost $\mu=0$, varijansa $\sigma^2=1$)

c) Testna grupa formirana od studenata Visoke tehničke škole strukovnih studija u Nišu, starosti od 19÷32 godine (prosečne starosti $\mu= 22$ godina, varijanse $\sigma^2=7$)

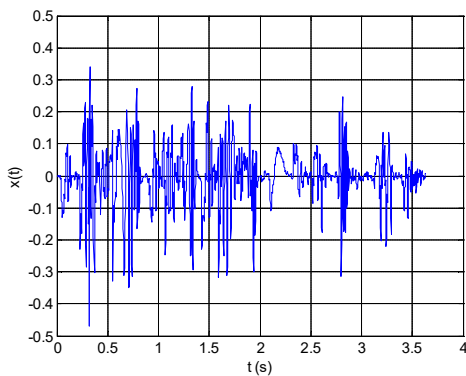
a) SMST baza

Baza SMST-a iz koje je dobijen govorni materijal za testiranje snimljena je u prostorijama Visoke Tehničke Škole Strukovnih Studija u Nišu, pomoću odgovarajuće standardne opreme. Snimljen govorni materijal pamćen je u wav obliku na hard disk računara. Semplovanje govornog signala vršeno je na frekvenciji $f_s=44.1\text{kHz}$ i 16bps [9]. Nakon programskog odabira rečenice po slučajnom zakonu za testiranje razumljivosti pomoću MOS testa, rečenica se ispisivala na ekranu koji je bio vidljiv samo ispivaču. Na Sl. 3 prikazan je vremenski oblik jedne od rečenica (“Gordana vidi devet žutih cvetova”) koje su korišćene u ovom eksperimentu. Na Sl. 4 prikazana je spektralna karakteristika rečenice.

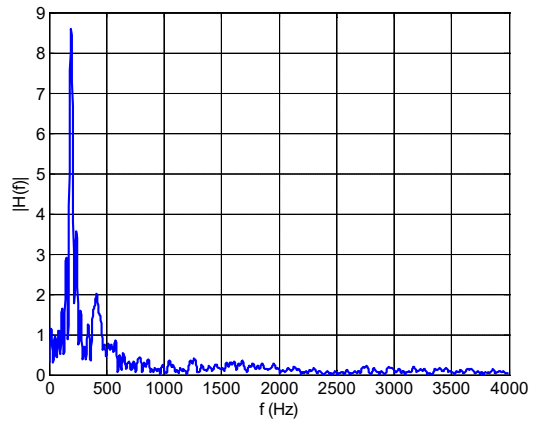
b) Baza Gausovog šuma

Formiranje baze sekvenci Gausovog šuma, čiji je broj jednak broju rečenica i koji su superponirani rečenicama po slučajnom odabiru. Na Sl. 5 prikazan je vremenski oblik Gausovog šuma za segment trajanja 32ms, dok je spektralna karakteristika ovog šuma prikazana na Sl. 6.

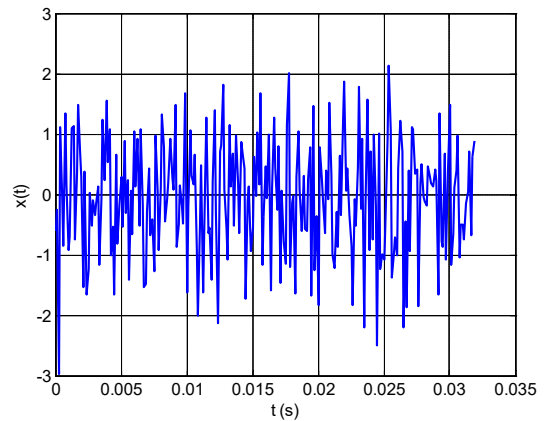
U cilju komparativne analize na Sl. 7. prikazan je vremenski oblik, za segment od 32ms, Babble šuma, tipa N8 (istovremeni govor 8 govornika), dok je na Sl. 8 prikazana njegova spektralna karakteristika.



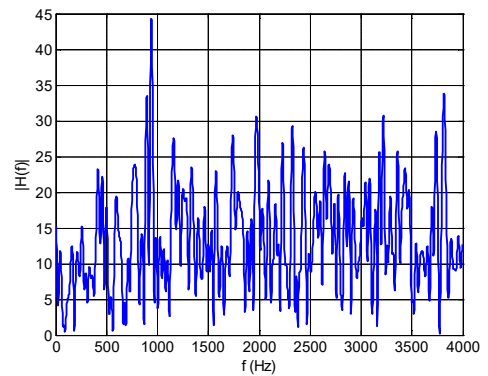
Slika 3. Vremenska karakteristika rečenice.



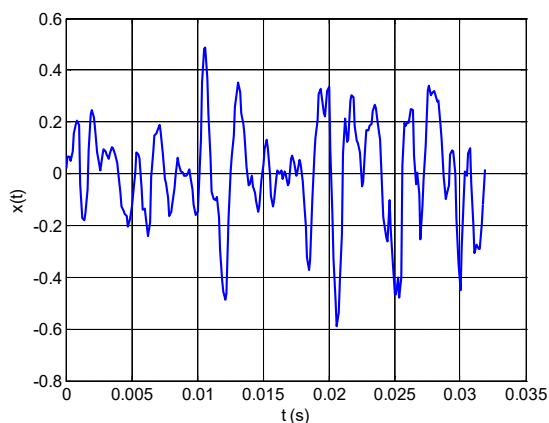
Slika 4. Spektralna karakteristika rečenice.



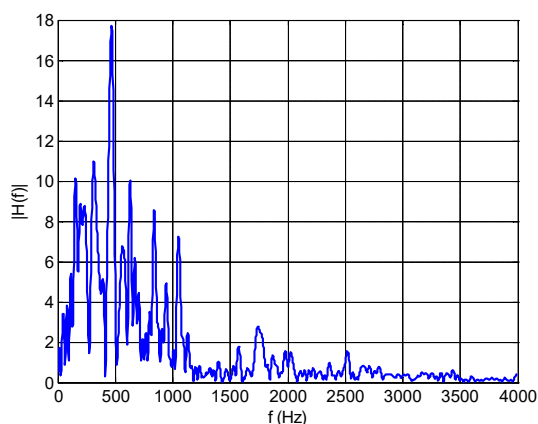
Slika 5. Vremenska karakteristika Gausovog šuma.



Slika 6. Spektralna karakteristika Gausovog šuma.



Slika 7. Vremenska karakteristika Babble šuma.



Slika 8. Spektralna karakteristika Babble šuma.

c) Testna grupa

Testna grupa je formirana od studenata Visoke tehničke Škole Strukovnih Studija u Nišu (15 muških i 15 ženskih osoba), starosti od 19÷32.

Testiranje razumljivosti je vršeno pomoću MOS testa koji se koristi za subjektivno merenje razumljivosti. Slušaocu je reprodukovani govorni signal sa Gausovim šumom putem slušalica. Slušalac ponavlja glasno šta je čuo, odnosno ono što misli da je razumeo. Ispitivač na osnovu ispisane rečenice na ekranu, tabelarno beleži ispravnost odgovora na nivou rečenice i na nivou pojedinih vrsta reči (imena, glagola, broja, prideva i objekata).

3.3. Rezultati

U Tabeli II prikazana je razumljivost cele rečenice kao i pojedinih reči u prisustvu Gausovog šuma za SNR. U Tabeli III prikazana je razumljivost pojedinih reči i rečenica u prisustvu Babble šuma za $SNR = \{-3, -2, -1, 0, 1, 3, 5\}$ dB [10]. U Tabeli IV prikazana je razumljivost vrste reči u prisustvu Gausovog i Babble šuma [10].

Na slikama od 9÷15 grafičari prikazane razumljivosti: a) Imena (Sl. 9), b) Glagola (Sl. 10), c) Broja (Sl. 11), d) Prideva (Sl. 12), e) Objekta (Sl. 13), f) vrste reči (Sl. 14) i g) Cele rečenice (Sl. 15).

TABELA II. RAZUMLJIVOST REČI I REČENICA U PRISUSTVU GAUSOVOG ŠUMA

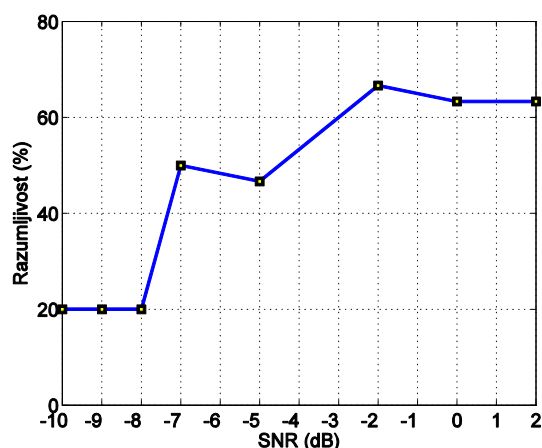
SNR [dB]	Razumljivost [%]					
	Ime	Glagol	Broj	Pridev	Objekat	Cela rečenica
2	63.33	63.33	76.67	76.67	73.33	26.66
0	63.33	50	70	63.33	53.33	3.33
-2	66.67	40	63.33	60	53.33	20
-5	46.67	16.67	53.33	56.67	33.33	0
-7	50	20	33.33	20	23.33	6.67
-8	20	23.33	23.33	16.67	30	6.67
-9	20	20	23.33	13.33	20	0
-10	20	13.33	10	3.33	13.33	0

TABELA III. RAZUMLJIVOST REČI I REČENICA U PRISUSTVU BABBLE ŠUMA [10]

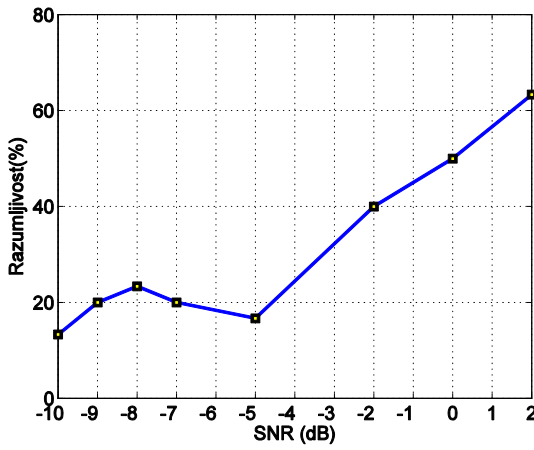
SNR [dB]	Razumljivost [%]					
	Ime	Glagol	Broj	Pridev	Objekat	Cela rečenica
5	40	50	90	90	63.33	20
3	60	60	60	70	50	20
1	50	56.67	60	50	33.33	10
0	53.33	36.67	53.33	46.67	30	6.67
-1	40	20	46.67	36.67	33.33	3.33
-2	40	10	36.67	20	16.67	6.67
-3	46.67	30	33.33	23.33	20	0
-5	26.67	6.67	10	3.33	6.67	0

TABELA IV. RAZUMLJIVOST REČI U PRISUSTVU GAUSOVOG ŠUMA I BABBLE ŠUMA [10]

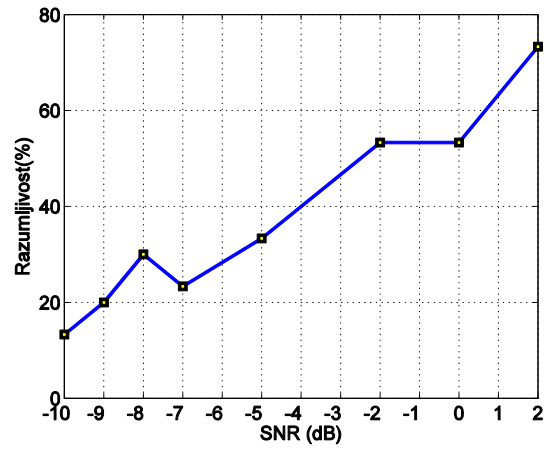
Razumljivost (%)	Ime	Glagol	Broj	Pridev	Objekat
Gausov šum	43.75	30.83	44.17	38.75	37.50
Babble šum	71.61	55.48	78.06	68.39	50.97



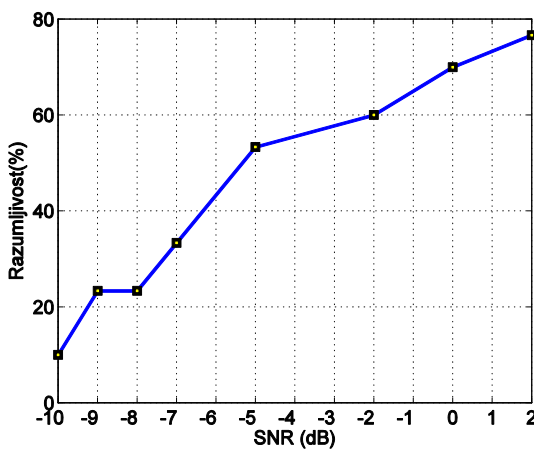
Slika 9. Razumljivost vrste reči "Ime" u zavisnosti od SNR



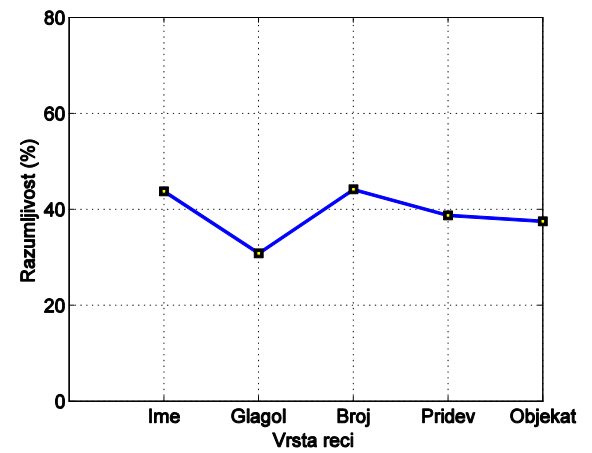
Slika 10. Razumljivost vrste reči "Glagol" u zavisnosti od SNR



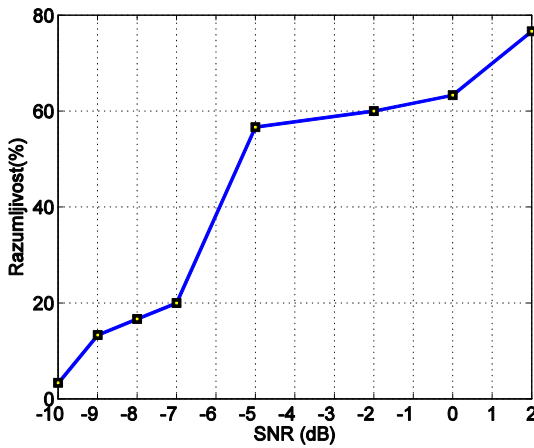
Slika 13. Razumljivost tipa reči "Objekat" u zavisnosti od SNR



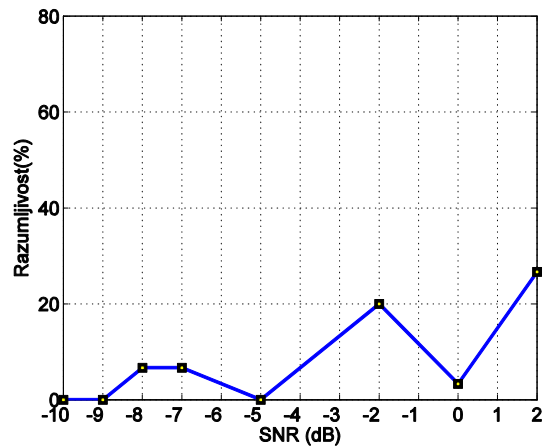
Slika 11. Razumljivost tipa reči "Broj" u zavisnosti od SNR



Slika 14. Razumljivost različitih vrste reči



Slika 12. Razumljivost tipa reči "Pridev" u zavisnosti od SNR



Slika 15. Razumljivost celih rečenica u zavisnosti od SNR

3.4. Analiza rezultata

Analizirajući rezultate sagledane u Tabeli III i prikazane na Sl. od 7÷13. dolazi se do zaključka da je razumljivost tipa reči sledeća:

- a) **Ime** - najbolja za -2 dB (66.67%), a najlošija za -8,-9 i -10 dB (20%),
- b) **Glagol** - najbolja za 2 dB (63.33%), a najlošija za -10 dB (13.33%),
- c) **Broj** - najbolja za 2 dB (76.67%), a najlošija za -10 dB (10%),
- d) **Pridev** - najbolja za 2 dB (76.67%), a najlošija za -10 dB (3.33%),
- e) **Objekat** - najbolja za 2 dB (73.33%), a najlošija za -10 dB (13.33%),

Analizirajući rezultate sagledane u Tabeli IV i prikazane na Sl. 14 dolazi se do zaključka da je razumljivost tipa reči (bez obzira na SNR) sledeća:

- a) **Broj** (44.17%)
- b) **Glagol** (30.83%)

Analizirajući rezultate sagledane u Tabeli 3 i prikazane na slici 15 dolazi se do zaključka da je razumljivost rečenica sledeća:

- a) najbolja za 2dB (26.66%), a
- b) najlošija -5, -9 i -10 dB (0%).

Rezultati dobijeni za 2dB za Glagol (63.33%), Broj (76.67%), Pridev (76.67%) i Objekat (73.33%) odgovaraju hipotezi [11] da prilikom merenja razumljivosti govora u prisustvu šuma kod osoba sa normalnim neoštećenim sluhom razumljivost bude 50%, za SNR=2dB.

Vršeći komparativnu analizu sa Međunarodnim standardom IEC 60268-16:2011, dolazi se do zaključka da je razumljivost rečenica loša (26.66%). Posmatrajući prema vrsti reči razumljivost je slaba za **Broj** (76.67%), **Pridev** (76.67%) i **Objekat** (73.33%), a loša za **Ime** (66.67%) i **Glagol** (63.33%).

Upoređujući sa rezultatima iz [10] prikazanim u Tabeli III i IV primećuje se da je na vrednostima koje su sagledane u oba rada SNR = {-5, -2, 0}dB značajno bolja razumljivost u prisustvu Gausovog šuma. Ovo je očekivani rezultat imajući u vidu raspored energije u spektru Gausovog (Sl.6.) i Babble (Sl. 8.)šuma. Raspodela energije u spektru kod Gausovog šuma je u celom opsegu ($f_0=0 \div F_s/2$), dok je raspodela kod Babble šuma koncentrisana u opsegu od 100÷1000Hz, kakva je i kod govornih signala iz SMST baze. Samim tim za iste vrednosti SNR-a Babble šum je više degradirajući.

IV. ZAKLJUČAK

Na osnovu sprovedenog MOS testa razumljivosti performansa (reči i rečenica) SMST baze u prisustvu

akustičkog Gausovog šuma dolazi se do zaključka da je razumljivost celih rečenica loša, dok je razumljivost pojedinačnih vrsta reči slaba i prihvatljiva (>50%). U zavisnosti od vrste šuma (Gaus ili Babble) koji maskira govorni signal – rečenicu, izvodi se zaključak da je razumljivost manja u prisustvu Babble šuma zbog maskiranja u spektru. Ovakav zaključak može se reći da je očekivan s obzirom da Gausov šum ima ravnomerno raspoređenu energiju, dok Babble šum ima energiju raspoređenu od 100÷1000 Hz.

LITERATURA

- [1] Z. Milivojević, V. Stojanović, Z. Veličković, D. Kostić, "Procena subjektivne razumljivosti CVC logatoma u ambijentu pravoslavne crkve u prisustvu babble šuma", INFOTEH-JAHORINA Vol. 15, pp 383-387, March 2016.
- [2] 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, Makedonija, pp. 153-156, 2016
- [3] R. Plomp & A.M. Mimpen, "Improving the reliability of testing the speech reception threshold for sentences". Audiology, 18, 43–52, 1979.
- [4] B. Hagerman, "Sentences for testing speech intelligibility in noise", Scand Audio, Vol. 11, pp. 79-87, 1982.
- [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) 536-544, 2012.
- [6] H. Luts et al., FIST: a French sentence test for speech intelligibility in noise. International Journal of Audiology, 47:373-374, 2008.
- [7] C. Cattermole, Comparison of Speech Recognition in Noise With Sound Field And Earphone Presentation. MSc thesis, University of Southampton, 2003
- [8] E. Ozimek, A. Warzybok, D. Kutzner, "Polish sentence matrix test for speech intelligibility measurement in noise", Int. J. Audio. 49:444-454, 2010Z.
- [9] Z. Milivojević, D. Kostić, Z. Veličković, D. Brodić, "Serbian sentence matrix test for speech intelligibility measurement in different reverberation conditions". UNITEH Gabrovo, pp. - , 2016
- [10] Z. Milivojević, D. Kostić, Z. Veličković, D. Brodić, "Performance of Serbian Matrix Sentence Test in presence of the Babble noise," in press.
- [11] M. Killion, "New Thinking on Hearing in Noise: A Generalized Articulation Index", Seminars in Hearing, 23(1), pp 57-75, 2002.

ABSTRACT

First part of the paper describes SMST base that consists sentences spoken in Serbian language. SMST base is created to test intelligibility of words and sentences on people with normal and undamaged hearing. Second part of the paper describes the experiment which was performed in aim to determine the intelligibility of sentences in the presence of acoustic Gaussian noise. The results are presented in tables and graphs. At the end of the paper it was performed comparative analysis of the results intelligibility of the speech signal degraded by Gaussian and Babble noise.

PERFORMANCE INTELLIGIBILITY SERBIAN MST-A IN THE PRESENCE OF ACOUSTIC GAUSSIAN NOISE

Zoran Milivojević, Dijana Kostić and Darko Brodić