

# Razumljivost govora u ambijentu ranih refleksija Gausovog i difuznog šuma

Zoran Milivojević, Zoran Veličković

Visoka tehnička škola strukovnih studija  
Niš, Srbija  
zoran.milivojevic@vtsnis.edu.rs  
zoran.velickovic@vtsnis.edu.rs

Dijana Kostić

“Geonais” Niš  
Niš, Srbija  
koricanac@yahoo.com

**Sažetak**— U prvom delu rada analiziran je efekat razumljivosti u prisustvu ambijentalnih smetnji (refleksija, Gausov šum i šum ambijenta). U drugom delu rada opisan je eksperiment u kome je testirana razumljivost govora kod izgovaranja rečenica na srpskom jeziku u prisustvu snažne prve refleksije, Gausovog šuma i šuma ambijenta. Rezultati eksperimenta su prikazani tabelarno i grafički. Na kraju je izvršena komparativna analiza razumljivosti i izvršena klasifikacija razumljivosti u odnosu na standard IEC 60268-16: 2011.

**Gljučne reči** - Razumljivost; Refleksije, Difuzioni šum; STOI; SMST baza.

## I. UVOD

Kod razmene informacija putem govora u ambijentu gde postoji više izvora akustičkih smetnji, dolazi do degradacije govornog akustičkog signala. Degradacije govornog signala mogu biti u tolikoj meri da se kod slušaoca pojavi semantički problem, odnosno problem nerazumljivosti poruke. Degradacije govornog signala mogu da dovedu do nerazumevanja : a) pojedinih reči, b) grupe reči, kao i c) čitavih rečenica. Analizom problema razumljivosti bavili su se mnogi autori, koristeći pri tom različite metode u cilju procene stepena razumljivosti [1] - [4]. Korišćeni su eksperimenti u kojima je, na osnovu specijalno dizajniranih reči, grupa reči ili rečenica, vršeno testiranje razumljivosti testne grupe slušaoca. Test reči mogu biti: a) smislene [5] ili b) reči bez smisla u određenom jeziku (logatomi [6]). Test rečenice mogu biti konstruisane u: a) semantičkom ili b) sintaksnom obliku [7]. Semantički oblik podrazumeva rečenice iz svakodnevnog govora, dok rečenice u sinaksnom obliku predstavljaju rečenice sa ispravnim sintaksnom strukturom ali, ne nužno, semantički korektne. Takve rečenice se formiraju pomoću računara korišćenjem reči iz specijalnih baza reči, prema slučajnom zakonu. Kreirane su matrične baze reči za švedski [8], španski [9], ruski [10], poljski [11], danski [12], holandski [13] i srpski jezik [14].

Razumljivost se, u numeričkom smislu, izražava procentualno u opsegu 0-100% ili pomoću srednje, odnosno MOS (engl. *Mean Opinion Score*), ocene. U cilju simulacije realnih ambijentalnih uslova, odnosno simuliranja akustičkih maskirajućih efekata, govornim signalima superponirane su različite akustičke smetnje, kao što su: a) šum ambijenta [15],

b) prostorni efekti (reverberacija, eho, rane i kasne refleksije), c) smetnje izazvane prisustvom drugih osoba (Babble noise) [1], d) industrijskih [16] i muzičkih smetnji [17], [18], e) prirodnih pojava kao što su kiša [19], vetar [20] i dr.

U ovom radu analiziran je efekat ranih refleksija na razumljivost. Razumljivost je testirana eksperimentalnim putem. Eksperiment je sproveden u prostorijama Visoke tehničke škole strukovnih studija u Nišu. U eksperimentu su korisnom akustičkom test signalu superponirane akustičke smetnje sa promenljivim intenzitetom. Superponirane su smetnje nastale kao posledica: a) prostornih uticaja (rane refleksije), b) ambijentalnog difuznog šuma (eng. *Diffusion Noise*, DN) i c) izvora Gausovog (eng. *Gaussian Noise*, GN) šuma iz jedne tačke prostora. Test rečenice formirane su pomoću računara na osnovu reči iz srpske matrične baze SMST (eng. *Serbian Matrix Sentence Test*), kreirane i testirane od strane autora ovog rada [14]. Akustički test signal sa superponiranim smetnjama testiran je objektivnom metodom pomoću STOI (engl. *Short Term Objective Intelligibility*) algoritma [21]. Kao rezultat testiranja generiše se dSTOI koeficijent sa vrednostima 0 - 1, gde 0 predstavlja apsolutnu nerazumljivost, a 1 apsolutnu razumljivost. Testiranje je sprovedeno prema binauralnom konceptu gde se razumljivost proverava posebno za levo i desno uvo. U drugom delu rada su prikazani rezultati razumljivosti, dobijeni eksperimentalnim putem, pomoću tabela i grafika. Sprovedena je komparativna analiza eksperimentalnih rezultata sa rezultatima razumljivosti u prisustvu Babble šuma [22]. Poređenjem sa standardom IEC 60268-16:11 [23] došlo se do zaključka o stepenu razumljivosti.

Organizacija rada je sledeća. U sekciji 2 opisani su govorni signal, rane refleksije, Gausov i difuzni šum. Sekcija 3 opisuje eksperiment, njegove rezultate i analizu istih. Sekcija 4 je zaključak.

## II. RANE REFLEKSIJE

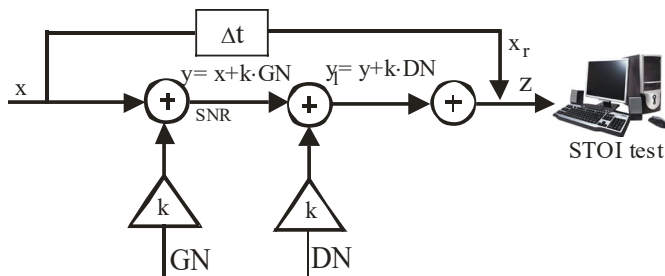
U zatvorenom prostoru do slušaoca, pored direktnog talasa, dolazi i niz talasa nakon refleksije od zidova, poda, tavanice i drugih objekata u prostoru. Kao posledica prisustva reflektovanih talasa dolazi do smanjenja razumljivosti. Za razumevanje je značajan direktan talas od izvora do mesta prijema i, dominantno, prvi reflektovani talasi (rane refleksije).

Značajne su refleksije koje stižu do prijema sa kašnjenjem do 50 ms u odnosu na direktni talas. Različite studije su pokazale da talasi sa kašnjenjem do 35 ms povećavaju efekat glasnoće a ne utiču na razumevanje, čak doprinose poboljšanju razumljivosti. Reflektovani talasi sa kašnjenjem preko 35 ms dovode do degradacije akustičkog signala na prijemu tako da se smanjuje razumljivost. U ovom radu je eksperimentalno analiziran efekat rane refleksije talasa sa kašnjenjem od 50 ms.

### III. EKSPERIMENTALNI REZULTATI I ANALIZA

#### A. Eksperiment

Sproveden je eksperiment sa ciljem određivanja degradacije govornog akustičkog signala u prisustvu: a) rane refleksije, b) ambijentalnog difuznog šuma i c) šuma iz jedne tačke prostora. Blok šema, kojim su opisani koraci u eksperimentu prikazana je na Sl.1. Kao test signal  $x$  korišćene su rečenice iz srpske SMST test baze sa sintaksnom strukturom: ime, glagol, broj, pridev i imenica.



Slika 1. Blok šema eksperimenta.

Govornom signalu  $x$  superponira se Gausov šum tako da se dobija signal  $y$  sa odnosom signal/šum  $SNR = \{-5, -2, 0\}$  dB. Signalu  $y$  superponira se difuzni šum koji dovodi do  $SNR = \{-5, -2, 0\}$ . Nakon toga vrši se sabiranje signala sa šumom sa reflektovanim signalom  $x_r$  koji se generiše u kolu za kašnjenje sa  $\Delta t$ . U cilju realizovanja binauralnog testiranja korišćene su sledeće vrednosti: a) ugao govornog signala  $\phi = 0^\circ$ , b) ugao refleksije  $\phi_h = 0^\circ$ , c) ugao difuznog šuma  $\phi_{DN} = 0 : 5 : 360^\circ$ , d) amplituda refleksije  $A_r=1$ , e) vreme kašnjenja između direktnog zvuka i refleksije  $\Delta t = \{0, 10, 25, 50\}$  ms. Testiranja su obavljena za slučajeve kada su  $SNR$  usled dejstva Gausovog šuma ( $SNR_{GN}$ ) i  $SNR$  usled dejstva difuznog šuma ( $SNR_{DN}$ ): a) jednaki i b) različiti. Razumljivost je testirana STOI algoritmom [9] za oba uva posebno. Rezultati su prikazani tabelarno i grafički. Nakon toga sprovedena je analiza razumljivosti na osnovu koeficijenta dSTOI.

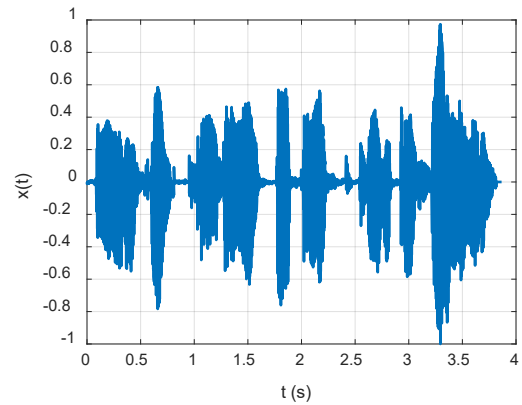
#### B. Baza reči

SMST baza reči izgovorenih na srpskom jeziku je sastavljena od 50 reči (po deset za ime, glagol, pridev, broj i imenicu). Kombinacijom se dobija 100000 rečenica. Matrične rečenice korišćene za testiranje razumljivosti dobijene su od reči iz SMST baze opisane u [14]. Reči su izgovorene na srpskom jeziku od strane profesionalnog ženskog spikera, a snimljene su u studiju "Banker radia" u Nišu. Snimanje je

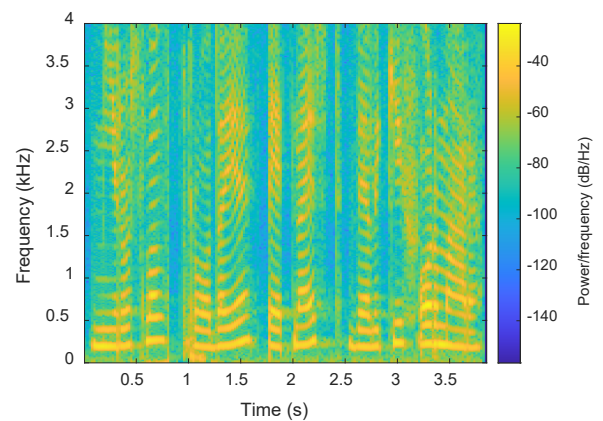
obavljeno sa  $f_s = 44100$  Hz i 16 bps. Snimci su u wav formatu arhivirani na hard disku.

#### C. Rezultati

U tabeli I prikazane su razumljivosti rečenica u prisustvu Gausovog i difuznog šuma za  $SNR_{GN} \neq SNR_{DN}$ . U tabeli II prikazane su razumljivosti rečenica u prisustvu Gausovog i difuznog šuma za  $SNR_{GN} = SNR_{DN}$ . U tabeli III prikazane su srednje vrednosti razumljivosti. Primer vremenskog oblika signala za rečenicu "Tamara ima deset žutih cvetova" prikazan je na Sl.2, dok je spektrogram prikazan na Sl.3. Na Sl.4. prikazan je vremenski oblik difuznog šuma. Na Sl.5. prikazan je spektar, a na Sl.6. spektrogram difuznog šuma. Test signal sa superponiranim reflektovanim signalom, Gausovim i difuznim šumom prikazan je na Sl.7. dok je njegov spektrogram prikazan na Sl.8. Razumljivost na levom uvu za  $SNR_{GN} \neq SNR_{DN}$  prikazana je na Sl.9. Razumljivost na levom uvu za  $SNR_{GN} = SNR_{DN}$  prikazana je na Sl.10. Razumljivost na desnom uvu za  $SNR_{GN} \neq SNR_{DN}$  prikazana je na Sl.11. Razumljivost na desnom uvu za  $SNR_{GN} = SNR_{DN}$  prikazana je na sl. Sl.12.



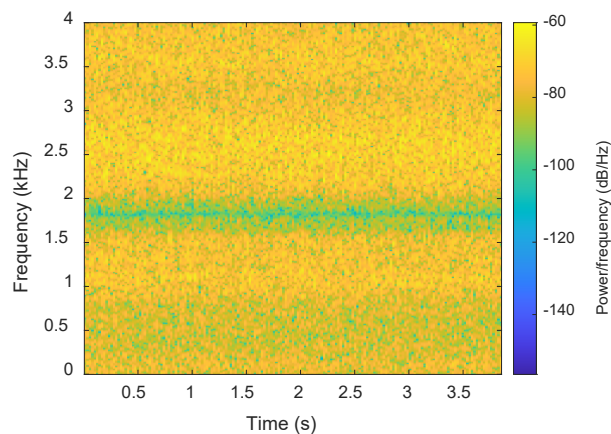
Slika 2. Vremenski oblik govornog signala formiranog iz SMST baze.



Slika 3. Spektrogram govornog signala formiranog iz SMST baze.

TABELA I. RAZUMLJIVOST REČENICA U PRISUSTVU GAUSOVOG I DIFUZNOG ŠUMA (SNR<sub>GN</sub> ≠ SNR<sub>DN</sub>, SNR<sub>DN</sub>=0dB)

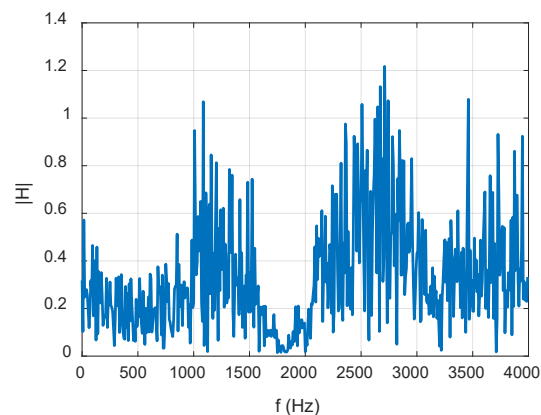
Razumljivost					
SNR (dB)	dSTOI	Δt (ms)			
		0	10	25	50
0	dSTOI <sub>L</sub>	0,70996	0,61199	0,59227	0,53778
	dSTOI <sub>R</sub>	0,71445	0,61679	0,5999	0,53541
-2	dSTOI <sub>L</sub>	0,70301	0,62272	0,59633	0,54483
	dSTOI <sub>R</sub>	0,69761	0,62573	0,59649	0,54794
-5	dSTOI <sub>L</sub>	0,72199	0,60535	0,59791	0,54812
	dSTOI <sub>R</sub>	0,72428	0,60382	0,592	0,54684



Slika 5. Spektrogram difuznog šuma – DN.

TABELA II. RAZUMLJIVOST REČENICA U PRISUSTVU GAUSOVOG I DIFUZNOG ŠUMA (SNR<sub>GN</sub> = SNR<sub>DN</sub>)

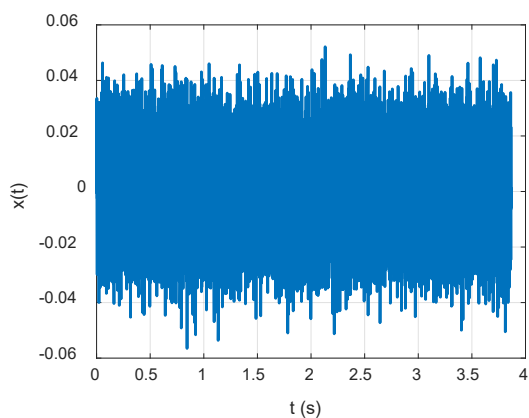
Razumljivost					
SNR (dB)	dSTOI	Δt (ms)			
		0	10	25	50
0	dSTOI <sub>L</sub>	0,70996	0,61199	0,59227	0,53778
	dSTOI <sub>R</sub>	0,71445	0,61679	0,5999	0,53541
-2	dSTOI <sub>L</sub>	0,65603	0,5814	0,57993	0,5437
	dSTOI <sub>R</sub>	0,65727	0,58054	0,57787	0,54501
-5	dSTOI <sub>L</sub>	0,63326	0,5525	0,50757	0,5437
	dSTOI <sub>R</sub>	0,63009	0,55556	0,50976	0,54501



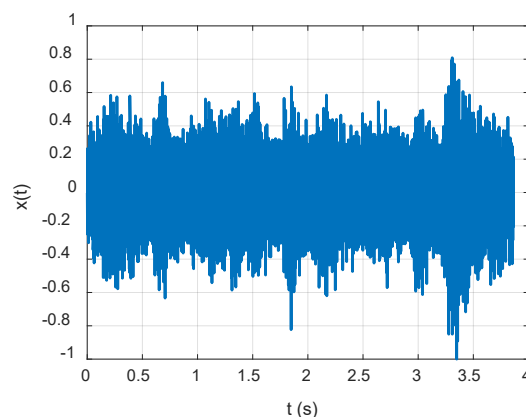
Slika 6. Spektar difuznog šuma – DN.

TABELA III. SREDNJA VREDNOST RAZUMLJIVOST REČENICA U PRISUSTVU GAUSOVOG I DIFUZNOG ŠUMA

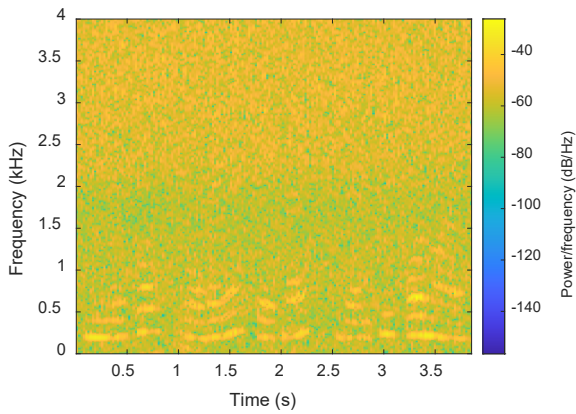
Razumljivost			
dSTOI	SNR (dB)		
	0	-2	-5
dSTOI <sub>L</sub>	0,613	0,6167	0,6183
dSTOI <sub>R</sub>	0,616	0,6169	0,6167
dSTOI <sub>L</sub>	0,613	0,59026	0,5519
dSTOI <sub>R</sub>	0,616	0,59017	0,55927
dSTOI <sub>L</sub>	0,56358	0,53814	0,47130



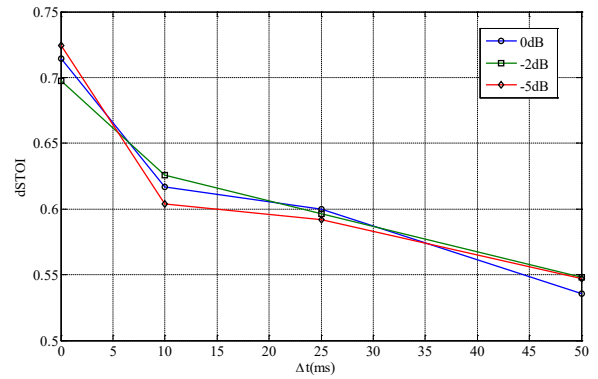
Slika 4. Vremenski oblik difuznog šuma – DN.



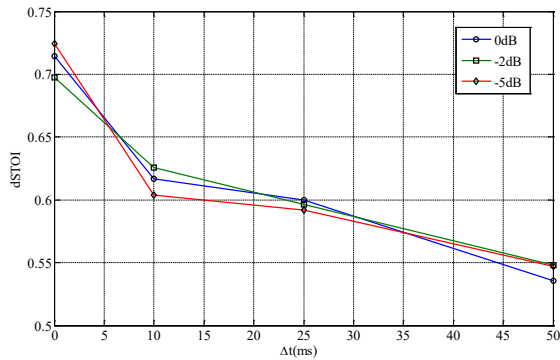
Slika 7. Vremenski oblik generisanog test signala z.



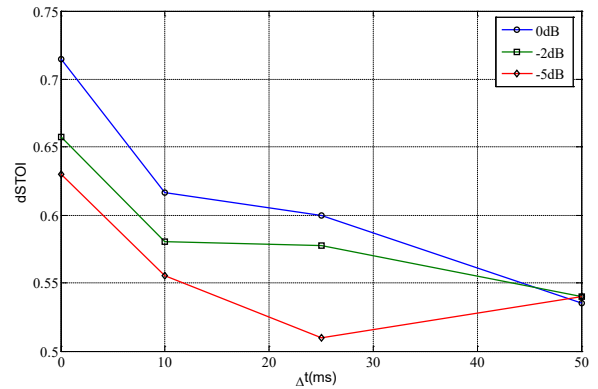
Slika 8. Spektrogram generisanog test signala z.



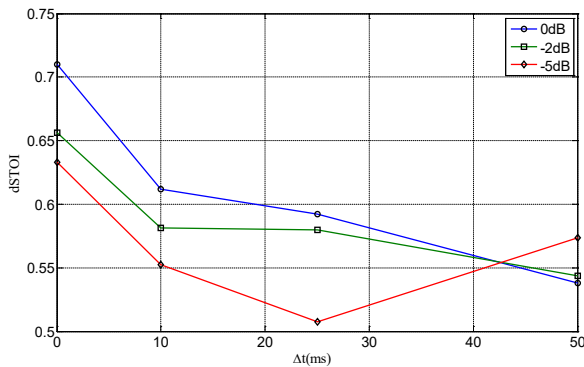
Slika 11. Razumljivost na desnom uhu kada je različit SNR.



Slika 9. Razumljivost na levom uhu, kad je različit SNR.



Slika 12. Razumljivost na desnom uhu kada je isti SNR.



Slika 10. Razumljivost na levom uhu, kad je isti SNR.

#### D. Analiza rezultata

Analizirajući rezultate prikazane u tabelama I-II i Sl. 9-12 može se zaključiti da se razumljivost rečenica za  $\Delta t = \{0, 10, 25, 50\}$  ms kreće u granicama:

- za  $\text{SNR}=0\text{dB}$  od 0,53778 - 0,70996 za levo uvo, i 0,53541- 0,71445 za desno uvo ( $\text{SNR}_{\text{GN}} \neq \text{SNR}_{\text{DN}}$ )
- za  $\text{SNR}=-2\text{dB}$  od 0,54483 - 0,70301 za levo uvo i 0,54794- 0,69761 za desno uvo ( $\text{SNR}_{\text{GN}} \neq \text{SNR}_{\text{DN}}$ ),
- za  $\text{SNR}=-2\text{dB}$  od 0,5437 - 0,65603 za levo uvo i 0,5401- 0,65727 za desno uvo ( $\text{SNR}_{\text{GN}} = \text{SNR}_{\text{DN}}$ ),
- za  $\text{SNR}=-5\text{dB}$  od 0,54812 - 0,72199 za levo uvo i 0,54684- 0,72428 za desno uvo ( $\text{SNR}_{\text{GN}} \neq \text{SNR}_{\text{DN}}$ ),
- za  $\text{SNR}=-2\text{dB}$  od 0,50757 - 0,63326 za levo uvo i 0,50976- 0,63009 za desno uvo ( $\text{SNR}_{\text{GN}} = \text{SNR}_{\text{DN}}$ ),

Analizirajući rezultate prikazane u tabeli III može se zaključiti da je srednja vrednost razumljivosti rečenica za  $\text{SNR} = \{-5, -2, 0\}$  dB:

- za  $\text{SNR}=0\text{dB}$  0,613 za levo uvo i 0,616 za desno uvo ( $\text{SNR}_{\text{GN}} \neq \text{SNR}_{\text{DN}}$ ),

- b) za  $SNR=-2dB$  0,6167 levo uvo i 0,6169 za desno uvo ( $SNR_{GN} \neq SNR_{DN}$ ),
- c) za  $SNR=-5dB$  0,6183 levo uvo i 0,6167 za desno uvo ( $SNR_{GN} \neq SNR_{DN}$ ),
- d) za  $SNR=-2dB$  0,59026 levo uvo i 0,59017 za desno uvo ( $SNR_{GN} = SNR_{DN}$ ),
- e) za  $SNR=-2dB$  0,55919 levo uvo i 0,55927 za desno uvo ( $SNR_{GN} = SNR_{DN}$ ),

Razumljivost kod kombinacije rane refleksije, Gausovog šuma i difuznog šuma (rezultati iz ovog rada) je veća u odnosu na kombinaciju rane refleksije i Babble šuma, i to: a) 1,03 puta za 0 dB, b) 1,12 puta za -2dB i c) 1,36 puta za -5dB. Ovi podaci ukazuju na veliku destruktivnost Babble šuma, što se objašnjava raspodelom energije u spektru koja je jednaka sa raspodelom govornog signala.

Upoređivanjem sa standardom IEC 60268-16: 2011, dolazi se do zaključka da razumljivost pripada klasifikaciji loše razumljivosti (0 ÷ 89 %). Kao zaključak nameće se da je uticaj rane refleksije u kombinaciji sa Gausovim i difuznim šumom uzročnik smanjenja razumljivosti.

#### IV. ZAKLJUČAK

U ovom radu testirana je razumljivost govora u prisustvu ranih refleksija u kombinaciji sa Gausovim i difuznim šumom, za vrednosti  $SNR = \{-5, -2, 0\}$  dB, pomoću rečenica iz SMST baze. Rezultati su upoređeni sa rezultatima iz [22], gde je vršen eksperiment sa kombinacijom ranih refleksija i Babble šuma. Komparativna analiza je pokazala da je manja razumljivost govornog signala u prisustvu Babble šuma, što se objašnjava raspodelom energije u spektru. Upoređivanjem vrednosti razumljivosti sa klasifikacijom iz standarda IEC 60268-16: 2011, dolazi se do zaključka da razumljivost u svim slučajevima testiranja pripada klasi loše razumljivosti.

#### LITERATURA

- [1] A. K. Nabelek, L. Robinette, "Influence of the precedence effect on word identification by normal hearing and hearing - impaired subject," J. Acous. Soc. Am, vol. 63, pp. 187-194, 1978.
- [2] J. Lohner, J. Burger "The influence of reflections on auditorium acoustics," J. Sound Vib, 1, pp. 426-454, 1964.
- [3] I. Arweiler, J.M. Bucholz, "The influence of spectral characteristics of early reflections on speech intelligibility," J. Acoust. Soc. Am, 130, pp. 996-1005, 2011.
- [4] G. Soudre, N. Popplwell, J. Bradley, "Combined effect of early reflections and background noise on speech intelligibility," J. Sound Vib 135, pp. 123-133, 1989.
- [5] J.E. Clark, "Four PB word lists for Australian English", Aust. J. Audiology 3(1), 21-31, 1981.
- [6] 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.
- [7] R. Plomp, A. M. Mimpen, "Improving the reliability of testing the speech reception threshold for sentences", Audiology, 18, pp. 43-52, 1979.

- [8] B. Hagerman, "Sentences for testing speech intelligibility in noise", Scand Audio, Vol. 11, pp. 79-87, 1982.
- [9] S. Hochmuth et al, "A Spanish matrix sentence test for assessing speech reception thresholds in noise", Int. J. Audiol. 51(7) 536-544, 2012.
- [10] M. Boboshko, A. Warzybok, M. Zokoll, N. Maltseva, "RUMatrix test: construction, evaluation and clinical validation", Otorhinolaryngologia Hungarica. Vol. 59, N 2, 2013.
- [11] E. Ozimek, A. Warzybok, D. Kutzner, "Polish sentence matrix test for speech intelligibility measurement in noise", Int. J. Audiol. 49:444-454, 2010.
- [12] K. Wagener, J. Lignel Josvassen, R. Ardenkjær, "Design, optimization and evaluation of a Danish sentence test in noise", International Journal of Audiology, Volume 42, 2003 - Issue 1.
- [13] R. Houben et al, "Development of a Dutch matrix sentence test to assess speech intelligibility in noise", Int J Audiol. 2014 Oct;53(10):760-3.
- [14] Z. Milivojević, D. Kostić, Z. Veličković, D. Brodić, "Serbian sentence matrix test for speech intelligibility measurement in different reverberation conditions", UNITEH Gabrovo, 2016.
- [15] Z. Milivojević, D. Kostić, D. Brodić, "Performanse razumljivosti Srpskog MST-a u prisustvu akustičkog Gausovog šuma", INFOTEH Jahorina, pp. 327-332, 2017.
- [16] Z. Milivojević, D. Kostić, D. Brodić, "The Influence of Industrial Noise on the Performance of Speech Intelligibility Serbian Sentence Matrix Test", ICEST Niš, pp., 2017.
- [17] D. Kostić, Z. Milivojević, Z. Veličković, "The influence of musical noise, type major and minor chord, to the intelligibility of speech in Serbian language", UNITEH Gabrovo, pp. II 212-217, 2017.
- [18] Z. Milivojević, D. Kostić, Z. Veličković, D. Brodić, "Influence of the musical noise, type major accord, on the comorensibility of speech in the Serbian language", SED 2017, Užice, Srbija
- [19] D. Kostić, Z. Milivojević, Z. Veličković, "The influence of early reflections and Babble noise on the intelligibility of speech signal", ICEST 2019, Ohrid, Macedonia.
- [20] Z. Milivojević, D. Kostić, Z. Veličković, D. Brodić, "Efekat dejstva akustičke smetnje tipa vetar na razumljivost govora", INFOTEH Jahorina, pp., 2018.
- [21] Z. Milivojević, D. Kostić, Z. Veličković, "The optimization of the STOI algorithm parameters in presence of the WGN", ICEST 2018, Sozopol, Bulgaria.
- [22] D. Kostić, Z. Milivojević, Z. Veličković, "The influence of early reflections and Babble noise on the intelligibility of speech signal", ICEST 2019, Ohrid, Macedonia.
- [23] 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.

#### ABSTRACT

The first part of the paper analyzes the effect of intelligibility in the presence of ambient noise (reflection, Gaussian noise and ambient noise). The second part of the paper describes an experiment in which speech intelligibility was tested when pronouncing sentences in Serbian in the presence of first reflection, Gaussian noise, and ambient noise. The results of the experiment are presented in tables and graphs. Finally, a comparative intelligibility analysis was performed and the intelligibility classification was performed based on IEC 60268-16: 2011.

#### Speech intelligibility in the ambience of early reflections, Gaussian and diffuse noise

Zoran Milivojević, Dijana Kostić, Zoran Veličković