

Razumljivost govora u prisustvu ranih refleksija, sirena smetnje i difuznog šuma

Zoran Milivojević

Akademija tehničko-vaspitačkih strukovnih studija
Niš, Srbija
e-mail: zoran.milivojevic@akademijanis.edu.rs

Dijana Kostić

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

Marina Milivojević

Dom Zdravlja Brus
Brus, Srbija
e-mail: marina.milivojevic@gmail.com

Bojan Prlinčević

Akademija strukovnih studija kosovsko-metohijska,
Leposavić, Srbija.
e-mail: bojan.prlincevic@akademijakm.edu.rs

Sažetak— U prvom delu rada analizirana je razumljivost govora u prisustvu ambijentalnih smetnji (refleksija Ar, difuzni šum DN i šum sirene SN). U drugom delu rada opisan je eksperiment u kome je testirana razumljivost govora korišćenjem rečenica iz Srpskog Matričnog rečeničnog testa (SMST) u prisustvu ranih refleksija, difuznog šuma i šuma od sirene. Rezultati eksperimenta su prikazani tabelarno i grafički. Na kraju je izvršena komparativna analiza razumljivosti sa rezultatima sličnih testova i izvršena klasifikacija razumljivosti u odnosu na standard IEC 60268-16: 2011.

Ključne reči - Razumljivost; Rane refleksije, Difuzioni šum; STOI;

I. UVOD

Važna karakteristika kod govora je razumljivost (engl. *intelligibility*). Akustički ambijentalni uslovi mogu dovesti do degradiranja govornog signala, i, samim tim, do smanjenja razumljivosti. Degradacije govornog signala mogu da dovedu do delimičnog ili potpunog nerazumevanja reči i/ili celih rečenica. Kod zatvorenog akustičkog prostora (sobe, učionice, sportske hale, pozorišta...) na razumljivost utiču akustičke karakteristike prostora: višestruke refleksije u formi ranih refleksija [1] [2], vreme reverberacije [3] i sl. Drugi akustički izvori dodatno degradiraju razumljivost: a) šum od strane većeg broja osoba (Babble [4], cocktail party [5]), b) muzičke smetnje [6], industrijski šum [7], i dr. Kod otvorenih prostora (ulice, trgovi, sportski stadioni, amfiteatri na otvorenom) degradaciju razumljivosti realizuju ambijentalne akustičke smetnje (vetar [8], kiša [9]), saobraćajne i industrijske smetnje, i dr. U cilju procene razumljivosti sprovode se razne metode testiranja, gde se pomoću specijalnih reči i rečenica vrši testiranje govora ili se testiraju osobe sa problemom sluha.

Metode testiranja se mogu svrstati u: a) subjektivne i b) objektivne metode testiranja. Kod subjektivnih metoda razumljivost se testira od strane većeg broja ispitanika, koji glasno izgovaraju rečenicu kako su čuli. Na taj način je

moгуće statistički odrediti faktor razumljivosti. Kod velike testne grupe računaju se srednje ocene razumljivosti pomoću MOS testa (Mean Opinion Score). Kod objektivnih metoda testiranja koriste se posebno dizajnirani programi, gde se na osnovu statističkih parametara govornog signala procenjuje indeks razumljivosti. Matrične baze kreirane su za švedski [11], španski [12], ruski [13], poljski [14], danski [15], holandski [16]. Autori su kreirali SMST (Serbian Matrix Sentence Test) bazu [7]. Generisanje rečenica je po slučajnom zakonu i to do 100000 rečenica sa definisanom sintaksnom strukturom.

Autori ovog rada koristili su STOI (Short-Time Objective Intelligibility measure) algoritam opisan u radu [10] za objektivno testiranje matričnih test rečenica (SMST Matrična baza na srpskom jeziku). Razumljivost se izražava procentualno u opsegu 0 - 100 %, i vrši klasifikacija razumljivosti na osnovu standarda IEC 60268-16: 2011 [17]. Analiziran je efekat akustičke smetnje koja nastaje od dejstva sirene (Sirene Noise, SN). Zvukovi sirene ozbiljno degradiraju razumljivost govorne komunikacije u blizini sirene (ambulantna, policijska vatrogasna vozila,...) [18]. Smanjenje efekta SN često se obavlja filtriranjem pomoću Notch filtra [19]. Realizovan je eksperiment u kome je simulirano složeno akustičko okruženje gde je govornom signalu x, superponiran reflektovani govorni signal, zatim superponiran SN za razne vrednosti SNR. Nakon toga je dodat difuzni šum. Nad ovako degradiranim signalom sprovedeno je objektivno testiranje razumljivosti pomoću binauralnog STOI algoritma. Rezultat testiranja je faktor razumljivosti za levo i desno uvo. Rezultati su prikazani tabelarno i grafički. Nakon toga je izvršena komparativna analiza rezultata sa razumljivošću kod degradacije nekim drugim akustičkim signalima. Klasifikacija razumljivosti je obavljena upoređivanjem dobijenih rezultata sa standardom IEC 60268-16: 2011.

Organizacija rada je sledeća. U sekciji II opisan je eksperiment, prikazani rezultati i izvršena komparativna analiza. Sekcija III je zaključak.

II. EKSPERIMENTALNI REZULTATI I ANALIZA

A. Eksperiment

Realizovan je eksperiment u kome je testirana razumljivost Test rečenica na srpskom jeziku u ambijentalnim uslovima: a) ranih refleksija govornog signala (AR), b) akustičke smetnje prouzrokovane policijskom sirenom (SN) i c) difuznog pozadinskog šuma (DN).

B. Algoritam testiranja

Algoritam realizacije eksperimenta i kreiranja test signala realizovan je u sledećim koracima:

Ulaz: Test rečenica na srpskom jeziku kreirana iz SMST matrične baze x , Vremensko kašnjenje Δt , SNR

Izlaz: Koeficijent razumljivosti dSTOI za levo (dSTOI_L) i desno (dSTOI_R) uvo.

Korak 1: Kreiranje reflektovanog signala sa ranim refleksijama
 $x_r = A_r \cdot x(t - \Delta t)$.

Korak 2: Govornom signalu x se superponira signal sirene SN sa projektovanim SNR: $y = x + k \cdot SN$.

Korak 3: Superponiranje difuznog šuma $y_1 = y + l \cdot DN$.

Korak 4: Formiranje konačnog test signala

$$Z = y_1 + x_r = x + A_r \cdot x(t - \Delta t) + k \cdot SN + l \cdot DN.$$

Korak 5: Testiranje razumljivosti primenom binauralnog STOI algoritma (koeficijent razumljivosti dSTOI za levo i desno uvo)

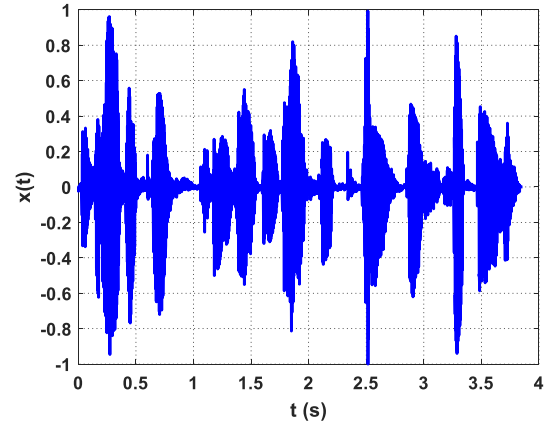
Razumljivost je testirana za: a) $\Delta t = (0, 10, 25, 50)$ ms, b) SNR = (0, -2, -5) dB; b) ugao pobude govornog signala $\phi_x = 0^\circ$; c) ugao pobude reflektovanog signala $\phi_r = 0^\circ$; d) ugao difuznog šuma $\phi_{DN} = 0 : 5 : 360^\circ$, e) amplituda refleksije $A_r = 1$.

Primer primene algoritma. Na Sl. 1 prikazan je vremenski oblik govornog signala x za rečenicu „*Dragoslav čuva deset čistih fotelja*“, dok je njegov spektrogram prikazan na Sl. 6. Spektrogram kreiranog reflektovanog govornog signala sa ranim refleksijama objašnjenog u koraku 1, prikazan je na Sl.7. Vremenski oblik signala šuma sirene SN prikazan je na Sl. 2 i Sl.3, dok je njegov spektar/spektrogram prikazan na Sl.5, Sl.8 i Sl.9 (Korak 2). Na Sl.3 prikazan je vremenski oblik, a na Sl.10 spektrogram signala za testiranje razumljivosti Z , dobijenog reflektovanjem govornog signala i superponiranjem difuznog šuma i šuma sirene (Korak 4).

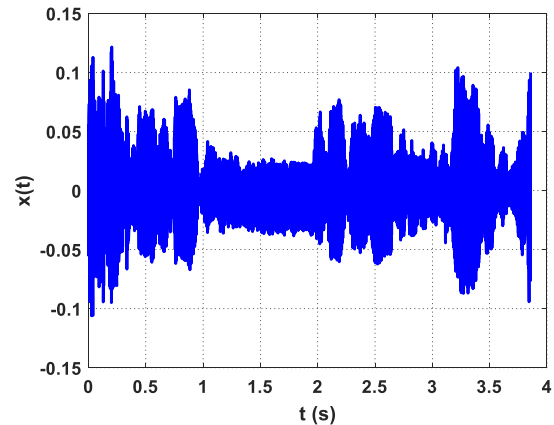
C. SMST matrična baza rečenica

SMST Matrična baza rečenica na srpskom jeziku formirana je od izolovanih reči tipa: *ime, glagol, pridev, broj i imenica*. Za svaki tip snimljeno je po deset reči. Reči su izgovorene od strane profesionalnog ženskog spikera. Snimanje je obavljeno 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. Generisanje test rečenica sa strukturom *ime, glagol, pridev, broj i imenica* kreira se programski prema slučajnom izboru pojedinih reči. Na taj način može da se formira 100000 rečenica. Primer jedne od rečenica „*Dragoslav čuva deset*

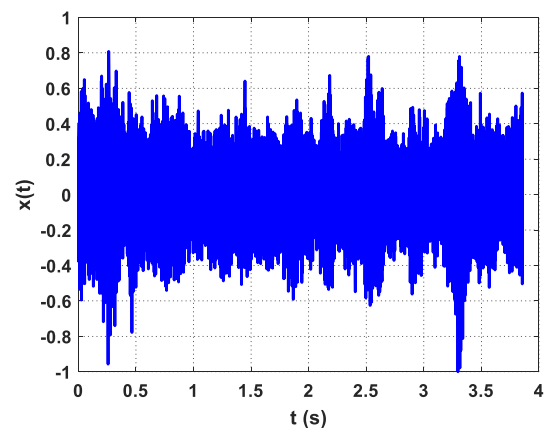
čistih fotelja“, prikazan je na Sl.1 u vremenskom obliku, dok je njen spektrogram prikazan na Sl.7.



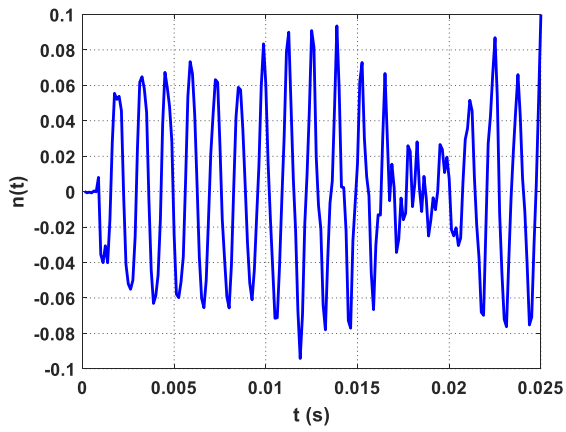
Slika 1. Vremenski oblik govornog signala „*Dragoslav čuva deset čistih fotelja*“.



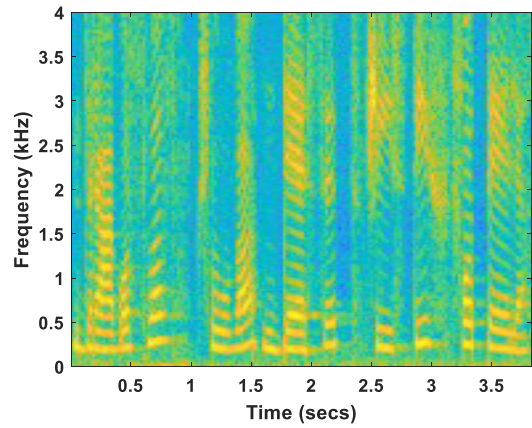
Slika 2. Vremenski oblik šuma tipa SN.



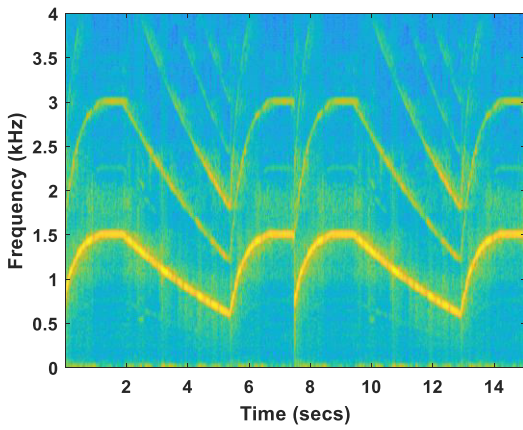
Slika 3. Vremenski oblik test signala Z.



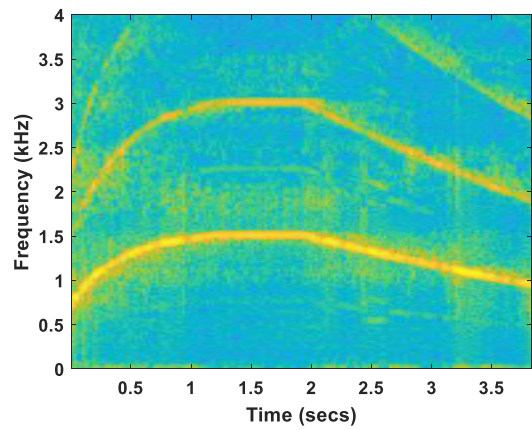
Slika 4. Detalj SN u vremenskom obliku.



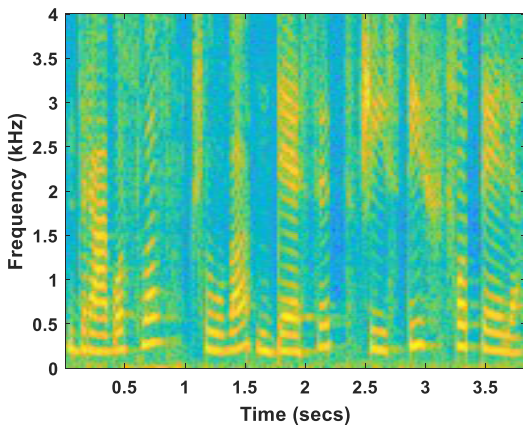
Slika 7. Spektrogram reflektovanog govornog signala.



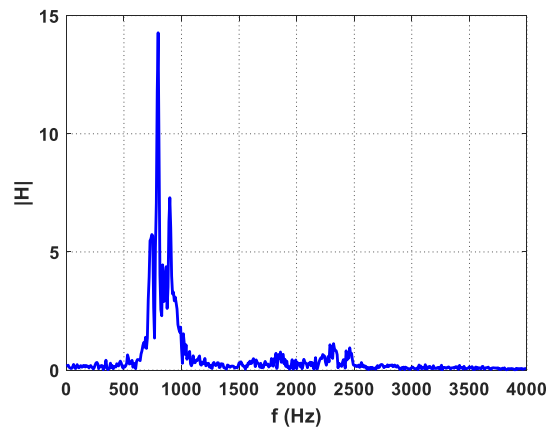
Slika 5. Spektar SN.



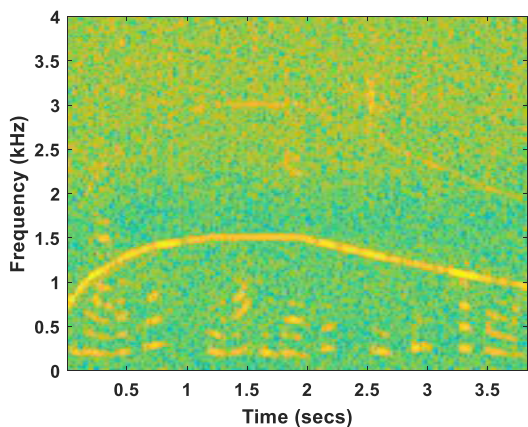
Slika 8. Spektrogram SN.



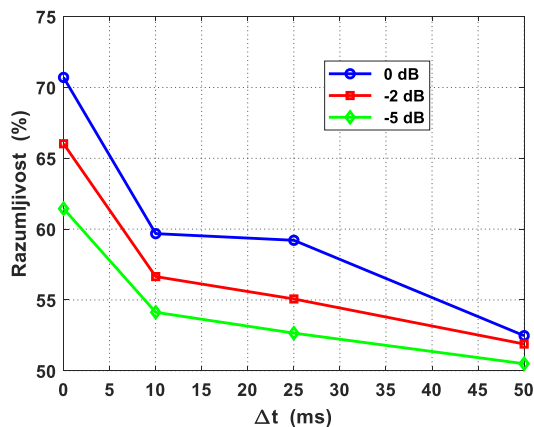
Slika 6. Spektrogram govornog signala.



Slika 9. Detalj SN u spektralnom obliku.



Slika 10. Spektrogram generisanog test signala Z.



Slika 11. Razumljivost na levom uhu za SNR_{SN} = SNR_{DN}.

D. Rezultati

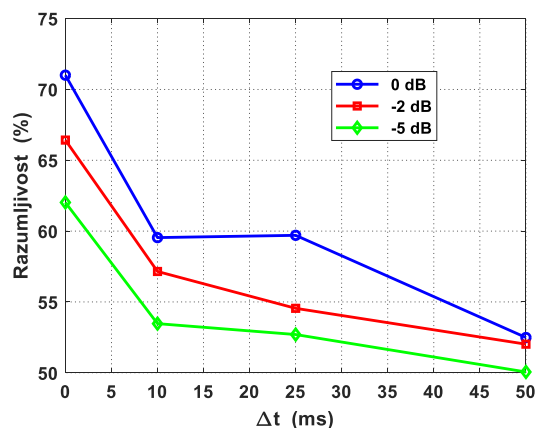
U tabeli I prikazane su razumljivosti rečenica u prisustvu ranih refleksija, difuznog šuma i šuma sirene, za uslov da je $SNR_{SN} = SNR_{DN}$, dok su u tabeli II prikazani rezultati za uslov da je $SNR_{SN} \neq SNR_{DN}$. Razumljivost na levom uvu za $SNR_{SN} = SNR_{DN}$ prikazana je na Sl.11, a za $SNR_{SN} \neq SNR_{DN}$ prikazana je na Sl.13. Razumljivost na desnom uvu za $SNR_{SN} = SNR_{DN}$ prikazana je na Sl.12, dok je za uslov $SNR_{SN} \neq SNR_{DN}$ prikazana na Sl.14.

TABELA I. RAZUMLJIVOST REČENICA U PRISUSTVU ŠUMA TIPA SIRENA I DIFUZNOG ŠUMA ($SNR_{SN} = SNR_{DN}$)

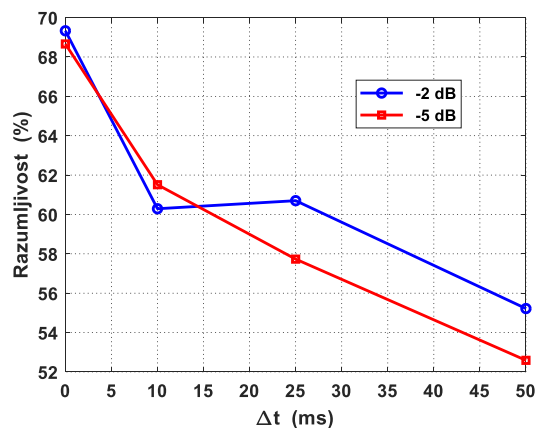
Razumljivost						
SNR (dB)	dSTOI	Δt (ms)				μ
		0	10	25	50	
0	dSTOIL	70,70	59,68	59,21	52,48	60,52
	dSTOIR	71,00	59,54	59,70	52,49	60,67
-2	dSTOIL	66,02	56,65	55,07	51,89	57,41
	dSTOIR	66,42	57,15	54,55	52,02	57,53
-5	dSTOIL	61,44	54,13	52,66	50,50	54,68
	dSTOIR	62,02	53,47	52,70	50,06	54,56

TABELA II. RAZUMLJIVOST REČENICA U PRISUSTVU ŠUMA TIPA SIRENA I DIFUZNOG ŠUMA ($SNR_{SN} \neq SNR_{DN}$)

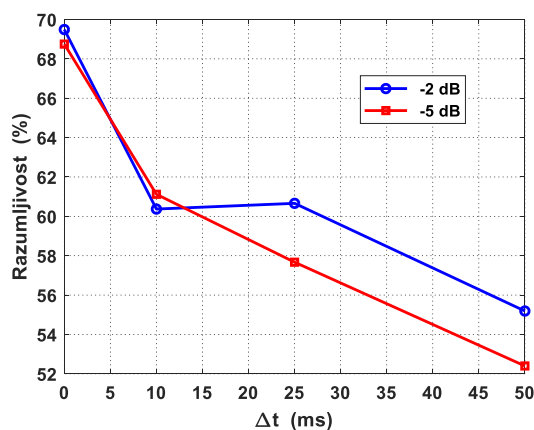
Razumljivost						
SNR (dB)	dSTOI	Δt (ms)				μ
		0	10	25	50	
-2	dSTOIL	69,33	60,29	60,70	55,22	61,39
	dSTOIR	69,49	60,37	60,66	55,19	61,43
-5	dSTOIL	68,66	61,51	57,73	52,59	60,13
	dSTOIR	68,75	61,12	57,67	52,40	59,98



Slika 12. Razumljivost na desnom uhu za SNR_{SN} = SNR_{DN}.



Slika 13. Razumljivost na levom uhu SNR_{SN} ≠ SNR_{DN}.



Slika 14. Razumljivost na desnom uhu $SNR_{SN} \neq SNR_{DN}$.

E. Analiza rezultata

Analizirajući rezultate prikazane u tabelama I-II i Sl. 11-14 može se zaključiti da se granica razumljivosti rečenica u zavisnosti do vremena kašnjenja $\Delta t = \{0, 10, 25, 50\}$ ms kreće:

- za $SNR=-2dB$ od 55,22 - 69,33% za levo uvo i od 55,19 - 69,49% za desno uvo ($SNR_{SN} \neq SNR_{DN}$),
- za $SNR=-5dB$ od 52,59 - 68,66% za levo uvo i od 52,40 - 68,75% za desno uvo ($SNR_{SN} \neq SNR_{DN}$),
- za $SNR=0dB$ od 52,48 - 70,70% za levo uvo i od 52,49 - 71,00% za desno uvo ($SNR_{SN} = SNR_{DN}$),
- za $SNR=-2dB$ od 51,89 - 66,02% za levo uvo i od 52,02-66,42% za desno uvo ($SNR_{SN} = SNR_{DN}$),
- za $SNR=-5dB$ od 50,50-61,44% za levo uvo i od 50,06-62,02% za desno uvo ($SNR_{SN} = SNR_{DN}$)

Analizirajući rezultate prikazane za srednju vrednost u tabelama I i II za $SNR = \{-5, -2, 0\}$ dB, ne uzimajući u obzir vreme kašnjenja zaključuje se da je razumljivosti rečenica:

- za $SNR=0dB$ 60,52% za levo uvo i 60,67% za desno uvo ($SNR_{SN} = SNR_{DN}$),
- za $SNR=-2dB$ 57,41% levo uvo i 57,53% za desno uvo ($SNR_{SN} = SNR_{DN}$),
- za $SNR=-5dB$ 54,68% levo uvo i 54,56% za desno uvo ($SNR_{SN} = SNR_{DN}$),
- za $SNR=-2dB$ 61,39% levo uvo i 61,43% za desno uvo ($SNR_{SN} \neq SNR_{DN}$),
- za $SNR=-5dB$ 60,13% levo uvo i 59,98% za desno uvo ($SNR_{SN} \neq SNR_{DN}$),

Upoređivanjem dobijenih rezultata sa standardom IEC 60268-16: 2011, dolazi se do zaključka da razumljivost pripada klasifikaciji loše razumljivosti (0 ÷ 89 %). Poređenjem sa rezultatima radova u kome je sagledan uticaj rane refleksije, difuznog šuma u kombinaciji sa pink [20], Gausovim [21] šumom, odnosno šumom tipa aplauz [22], kola [23] i

rezultatima ovog rada došlo se do zaključka da su dobijeni rezultati u istim granicama, kao i rezultati predhodnih ispitivanja (od 44,08 - 63,77%, posmatrajući srednju vrednost). Takođe se i ovde uočava bolja razumljivost na desnom uhu u odnosu na levom, u većini dobijenih rezultata.

III. ZAKLJUČAK

U ovom radu testirana je razumljivost govora u prisustvu ranih refleksija u kombinaciji SN i difuznim šumom, za vrednosti $SNR = \{-5, -2, 0\}$ dB, pomoću matičnih rečenica iz SMST baze. Rezultati su upoređeni sa rezultatima gde je vršen eksperiment sa kombinacijom ranih refleksija, difuznog šuma i pink šuma [20], Gausovog šuma [21], šuma tipa aplauz [22], i šuma tipa kola [23]. Komparativna analiza je pokazala da je najmanja razumljivost govornog signala u prisustvu šuma tipa aplauz što se objašnjava raspodelom energije u spektru. Komparativnom analizom dobijenih eksperimentalnih rezultata i klasifikacijom iz standarda IEC 60268-16: 2011, dolazi se do zaključka da razumljivost pripada klasi loše razumljivosti (0 ÷ 89 %), ali i da je prisutna bolja razumljivost za desno uvo.

LITERATURA

- I. Arweiler, J. Bucholz, "The influence of spectral characteristics of early reflections on speech intelligibility," J. Acoust. Soc. Am, 130, pp. 996-1005, 2011.
- J. S. Bradley and H. Sato, M. Picard, "On the importance of early reflections for speech in rooms," J. Acoust. Soc. Am. 113 (6), June 2003
- Z. Milivojević, D. Kostić, Z. Veličković, D. Brodić, "Serbian sentence matrix test for speech intelligibility measurement in different reverberation conditions," UNITEH Gabrovo, 2016
- D. Kostić, Z. Milivojević, Z. Veličković, "The influence of early reflections and Babble noise on the intelligibility of speech signal," ICEST 2019, Ohrid, Macedonia, pp. 150-153, 2019.
- E. C. Cherry, "Some experiments on the recognition of speech, with one and with two ears," J. Acoust. Soc. Am. 25, 1953, pp. 975-979.
- 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.
- 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.
- Z. Milivojević, D. Kostić, Z. Veličković, D. Brodić, "Efekat dejstva akustičke smetnje tipa vetar na razumljivost govora," INFOTEH Jahorina, pp. 268-272, 2018.
- D. Kostić, Z. Milivojević, Z. Veličković, "Influence of the rain noise to intelligibility in Serbian language," UNITEH Gabrovo, 2018.
- C. H. Taal, R.C. Hendriks, R. Heusdens and J. Jansen, "An Algorithm for Intelligibility Prediction of Time-Frequency Weighted Noise Speech," in IEEE Transactions on Audio, Speech, and Language Processing, vol.19, no. 7, pp. 2125-2136, Sept. 2011, doi: 10.1109/TASL.2011.2114881.
- B. Hagerman, "Sentences for testing speech intelligibility in noise," Scand Audio, Vol. 11, pp. 79-87, 1982.
- S. Hochmuth et al, "A Spanish matrix sentence test for assessing speech reception thresholds in noise," Int. J. Audiol. 51(7) 536-544), 2012.
- M. Boboshko, A. Warzybok, M. Zokoll, N. Maltseva, "RUMatrix test: construction, evaluation and clinical validation," Otorhinolaryngologia Hungarica. Vol. 59, N2, 2013.
- E. Ozimek, A. Warzybok, D. Kutzner, "Polish sentence matrix test for speech intelligibility measurement in noise," Int. J. Audiol. 49:444-454, 2010.

- [15] 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.
- [16] 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.
- [17] 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
- [18] H. Ding, L. Jing, Q. Xiaojun, X. Boling, "An adaptive speech enhancement method for siren noise cancellation," Applied Acoustics, Vol. 65 (2004), pp. 385–399.
- [19] B. Horvei, G.E. Ottesen, J. Svean, "Cancellation of siren sound in emergency vehicle communication equipment," In: Proc. IEEE Norwegian Signal Processing Symp., Tromso, Norway, May, 1997. pp. 26–9.
- [20] D. Kostić, M. Milivojević, Z. Milivojević, "Evaluation of the influence of Pink Noise (PN) on the intelligibility of speech in the Serbian language, using Serbian Matrix Sentence Test (SMST) base on STOI algorithm," Unitech Gabrovo, pp. 206-2011, 2020.
- [21] Z. Milivojević, Z.Veličković, D. Kostić, "Razumljivost govora u ambijentu ranih refleksija, Gausovog i difuznog šuma," INFOTEH Jahorina, pp. 104-108, 2020.
- [22] Z. Milivojević, Z.Veličković, D. Kostić, M. Milivojević, Speech intelligibility in the Serbian language in the presence of ambient noise of applause, INFOTEH Jahorina, pp- 2021.
- [23] M. Milivojević, D. Kostić, Z. Milivojević, "Influence of interference on speech intelligibility in a car using the binaural method", SED Užice, pp-, 2021 .

ABSTRACT

The first part of the paper analyzed the intelligibility of speech in the presence of ambient noise (reflection Ar , diffuse noise DN and siren noise SN). The second part of the paper described an experiment in which speech intelligibility was tested using sentences from the Serbian Matrix Sentence Test (SMST) in the presence of early reflections, diffuse noise and siren noise. The results of the experiment are presented in tables and graphically form. At the end, a comparative analysis of intelligibility was performed with the results of similar tests and a classification was performed in relation of standard IEC 60268-16: 2011.

Speech intelligibility in the presence of early reflections, siren noise and diffuse noise
 Zoran Milivojević, Dijana Kostić, Marina Milivojević, Bojan Prlinčević