

# Primjena skrivenih Markovljevih modela u prepoznavanju govora u šapatu

Jovan Galić, Milan Popović  
Elektrotehnički fakultet  
Banja Luka, Republika Srpska, BiH  
jgalic@etfbl.net, millanpopovic@hotmail.com

Branko Marković  
Visoka škola tehničkih strukovnih studija,  
Čačak, Srbija  
brankomarko@yahoo.com

Đorđe T. Grozdić, Slobodan T. Jovičić  
Elektrotehnički fakultet  
Centar za unapređenje životnih aktivnosti  
Beograd, Srbija  
djordjegrozdic@gmail.com, jovicic@etf.rs

**Sadržaj**—U radu su predstavljeni prvi rezultati istraživanja prepoznavanja riječi izgovorenih šapatom, primjenom skrivenih Markovljevih modela, obučeni na govor uobičajenog intenziteta. Prikazana je govorna baza, namjenski formirana za ovo istraživanje. Preliminarni rezultati pokazuju pad performansi sistema za prepoznavanje govora u šapatu u odnosu na prepoznavanje govora uobičajenog intenziteta.

**Ključne riječi**—prepoznavanje govora; šapat; skriveni Markovljevi modeli; govorna baza

## I. UVOD

Govorna komunikacija predstavlja čovjeku najprirodniji i napogodniji način komunikacije. Šapat, kao neuobičajen i rjeđe korišćen vid verbalne komunikacije, se koristi radi ostvarivanja diskretne ili intimne atmosfere u razgovoru. Takođe, može da bude i posljedica zdravstvenih problema (prehlada, kijavica ili hronično oboljenje laringealnih struktura vokalnog trakta). Nije rijetka ni upotreba šapata u kriminalnim aktivnostima, najčešće radi sakrivanja identiteta govornika.

Osnovna karakteristika šapata je odsustvo laringealnih vibracija i šumna struktura govornog signala. Utvrđeno je da su formanti na nižim frekvencijama u šapatu pomjereni ka višim frekvencijama, a spektralni nagib je dosta ravniji nego u normalnom govoru [1], [2], [3]. Pored navedenog, šapat ima dosta nižu energiju u odnosu na normalni govor [4]. Šapat ne posjeduje osnovnu frekvenciju, intonacione konture, a samim tim i mnogo drugih prozodijskih obilježja. Zbog svega navedenog, šapat predstavlja problem u govornim tehnologijama, posebno u prepoznavanju i sintezi govora, kao i identifikaciji govornika. Međutim, i pored nešto povećanog napora u percepciji, razumljivost šapata u kontinualnom govoru je veoma dobra. Prepoznavanje fonema u logatomima tipa konsonant-vokal-konsonant je preko 80%, kod svih vokala i preko 2/3 konsonanata [5], što je veoma dobra logatomska razumljivost, kojoj odgovara rečenička razumljivost od preko 95%. S druge strane, u šapatu je otežano prepoznavanje nelingvističkih informacija koje se tiču

polu govornika, doba starosti, emocionalnog stanja, identiteta, itd.

Postoje razni pristupi, tehnike i metode prepoznavanja govora. Te tehnike su najčešće zasnovane na algoritmima HMM (eng. *Hidden Markov Models*), DTW (eng. *Dynamic Time Warping*), neuralnim mrežama i njihovim hibridnim rješenjima [6].

Rad je koncipiran na sljedeći način. U odjeljku 2 je dat opis govorne baze koja je posebno formirana za ovakvu vrstu istraživanja. U odjeljku 3 je dat kratki opis programskog paketa korišćenog u istraživanju. U odjeljku 4 su prikazani rezultati eksperimenta i data njihova diskusija, dok je u odjeljku 5 dat zaključak sa pravcima budućih istraživanja. Cilj istraživanja je određivanje procenta prepoznavanja sistema za automatsko prepoznavanje govora, obučenog na govor uobičajenog intenziteta, ukoliko se primijeni na izolovane riječi izgovorene šapatom.

## II. OPIS GOVORNE BAZE

Za potrebe istraživanja najprije je kreirana govorna baza. Baza je snimana i obrađivana godinu dana i kompletirana je nedavno. Govorna baza je označena sa Whi-Spe i sastoji se od 50 riječi: 14 brojeva, 6 boja i 30 riječi. Riječi su preuzete iz govorne baze GEES [7], te zadovoljavaju osnovne jezičke kriterijume srpskog jezika (distribucija fonema, slogovna kompozicija, akcenatska struktura, konsonantski skupovi). Cijelu bazu je 10 govornika (5 ženskih i 5 muških), izgovaralo 10 puta normalnim govorom i šapatom. U svakoj sesiji subjekti su izgovarali cijelu bazu u kontinuitetu. U izgovoru svake riječi je ručno segmentiran početak/kraj i sačuvan kao posebni fajl. Znači, bazu sačinjava 5000 fajlova riječi izgovorenih normalnim govorom i 5000 fajlova riječi izgovorenih šapatom. Baza je snimana u Visokoj školi tehničkih strukovnih studija u Čačku, u laboratorijskim uslovima, kvalitetnim mikrofonom i u adaptiranoj akustičkoj komori smještenoj u prostoriji laboratorije. U tabeli 1 su dati sadržaji pojedinih riječi iz baze Whi-Spe.

TABELA I. SADRŽAJ RIJEČI IZ BAZE WHI-SPE

	Boje		Balansirane riječi	Riječi	
	Oznaka	Riječ		Oznaka	Riječ
Boje	Boja1	/bela/	Rijec1	/Mirko/	
	Boja2	/žuta/	Rijec2	/žurka/	
	Boja3	/crna/	Rijec3	/Petar/	
	Boja4	/crvena/	Rijec4	/demonstracije/	
	Boja5	/plava/	Rijec5	/standard/	
	Boja6	/zelena/	Rijec6	/pijaca/	
Brojevi	Broj1	/nula/	Rijec7	/padavine/	
	Broj2	/jedan/	Rijec8	/ponedeljak/	
	Broj3	/dva/	Rijec9	/godina/	
	Broj4	/tri/	Rijec10	/predstava/	
	Broj5	/četiri/	Rijec11	/komputeri/	
	Broj6	/pet/	Rijec12	/inostranstvo/	
	Broj7	/šest/	Rijec13	/drvo/	
	Broj8	/sedam/	Rijec14	/Mirjana/	
	Broj9	/osam/	Rijec15	/more/	
	Broj10	/devet/	Rijec16	/kiša/	
	Broj11	/deset/	Rijec17	/zgrade/	
	Broj12	/sto/	Rijec18	/klinici/	
	Broj13	/hiljadu/	Rijec19	/Milan/	
	Broj14	/milijon/	Rijec20	/rezultati/	
		Rijec21	/telefon/		
		Rijec22	/svetlo/		
		Rijec23	/prozor/		
		Rijec24	/ruke/		
		Rijec25	/lokal/		
		Rijec26	/ključ/		
		Rijec27	/sunce/		
		Rijec28	/pare/		
		Rijec29	/sef/		
		Rijec30	/blok/		

Snimci su u *mono* tehnici, sa frekvencijom odmjerenja 22kHz. Baza je dosnimavana u više navrata jer su kontrolama kvaliteta snimaka ustanovljene različite greške, subjektivne i objektivne prirode. Kod normalnog govora najčešće je dolazilo do pogrešnog izgovora date riječi ili pogrešne artikulacije pojedinog glasa u trenutku izgovora riječi (što je česta pojava i u svakodnevnom govoru), dok je pri samom snimanju bio izražen efekat akustičkog udara u mikrofon koji je bio u datom trenutku preblizu usta govornika. Ipak, najveći broj loših snimaka se odnosio na riječi izgovorene šapatom. Pošto se ovakva baza šapata po prvi puta formira za srpski jezik, a cilj je bio dobijanje regularno izgovorenih riječi šapatom, treba navesti najtipičnije greške kod izgovora i snimanja šapata [8]:

- suviše tih izgovor šapata (značajno maskiran šumom),
- suviše naglašen (izforsiran do izobličjenja) šapat,
- probijanje zvučnosti kod izgovora šapata,
- omisija glasova,
- direktno duvanje u mikrofon (akustički udar) i
- neregularnost rada artikulatora (pojava stridensa, trenutak odlepljivanja jezika od nepca i sl.).

Za ovo preliminarno istraživanje iskorišćen je dio baze sa šapatom, za svih 10 govornika.

### III. SOFTVERSKI PAKET I GOVORNA OBILJEŽJA

Za dobijanje preliminarnih rezultata smo koristili programski paket AlfaNumCASR, sistem za prepoznavanje kontinualnog govora koji se bazira na prepoznavanju fonema

u kontekstu, nezavisan od govornika. U mnogim dijelovima paket je kompatibilan sa poznatim paketom HTK (eng. *The Hidden Markov Model Toolkit*). Uspješno funkcioniše na malim i srednjim rječnicima (do par stotina riječi) i za normalni govor ima procenat prepoznavanja preko 98% za snimke telefonskog kvaliteta, odnosno preko 99% za snimke studijskog kvaliteta [9]. Detaljniji opis modula za obuku i prepoznavanje se može naći u radovima [10] i [11]. Sistem koristi tzv. *semi continuous HMM* pristup koji podrazumijeva da modeli pojedinih fonema u kontekstu nemaju svoje, i samo svoje, smješe, već postoji skladište svih smješa koje su nastale obukom, a svaki model posjeduje listu indeksa onih smješa koje mu pripadaju. Ovaj pristup obezbjeđuje veliku vremensku uštedu u procesu prepoznavanja. Koristi prozorovanje *Hamming*-ovim prozorom trajanja 30ms sa pomjerajem trajanja 10ms. Iz svakog prozorovanog segmenta su izdvojeni 12 prvih Mel frekvencijskih kepstralnih koeficijenata (eng. *Mel Frequency Cepstral Coefficients – MFCC*) i energija (13. koeficijent), čemu su dodati njihov prvi i drugi izvod. Na taj način se dobilo 39 MFCC koeficijenata po prozoru. Obuka je vršena na govornoj bazi telefonskog kvaliteta sa govorom uobičajene glasnoće (nije obuhvaćen šapat), a testirana na opisanoj bazi snimljenoj u studijskim uslovima, sa riječima izgovorenim šapatom. Modeli za prepoznavanje fonema u kontekstu su obučeni tako da obuhvataju i muške (njih je bilo više) i ženske glasove, a testiranja su vršena posebno na ženskim, a posebno na muškim govornicima. Obučeni modeli za prepoznavanje fonema u kontekstu omogućuju dalji razvoj sistema za prepoznavanje neograničenog rječnika, a ovdje su iskorišćeni samo za istraživanje koliko, kao takvi, mogu da prepoznaju izolovano izgovorene riječi u šapatu.

### IV. REZULTATI EKSPERIMENTA

Eksperiment je organizovan na sljedeći način. Govorni fajlovi svih riječi određene grupe (boje, brojevi i balansirane riječi) za određenog govornika(-cu) (GM1 do GM5 i GZ1 do GZ5) i određeni izgovor (1 do 10) su ručno unošeni u programski paket, koji je kao izlaz davao tekstualni fajl sa rezultatima prepoznavanja. Ti rezultati su bilježeni u pomoćne tabele u formi matrica konfuzije. Iz tih tabela su se prebrojavali rezultati radi određivanja broja uspješnih prepoznavanja (od nula do deset) za određenog govornika(-cu) i određenu riječ. Ti rezultati su prikazani u tabelama 2 (za muške govornike) i 3 (za ženske govornike). U posljednjoj koloni tabela je prikazan procenat prepoznavanja za određenu riječ, dok je u posljednjoj vrsti prikazan procenat prepoznavanja za određenog govornika(-cu). Na presjeku posljednje vrste po grupama i posljednje kolone (boldovane vrijednosti) je prikazan procenat prepoznavanja po grupama riječi. Iz tabela 2 i 3 se može vidjeti da određene riječi imaju nultu tačnost prepoznavanja, dok, s druge strane, neke riječi imaju maksimalnu tačnost prepoznavanja. Na Sl.1 je prikazan procenat prepoznavanja za muške govornike, i to u zavisnosti grupe riječi (a), odnosno pojedinačnog govornika (b). U ovom slučaju, procenat prepoznavanja po grupama je najmanji za balansirane riječi (51,3%), a najveći za riječi boja (59%).

TABELA II. USPJEH U PREPOZNAVANJU RIJEČI MUŠKIH GOVORNIKA

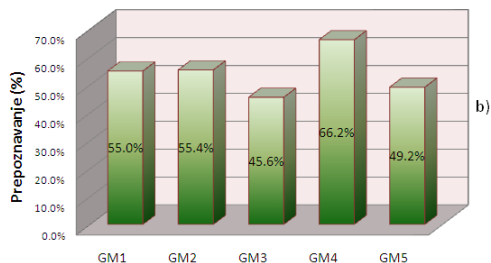
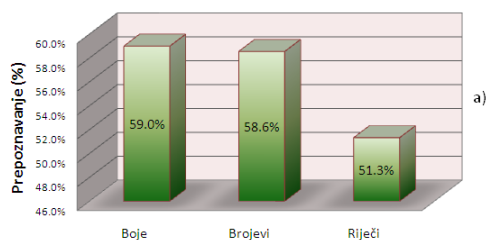
		GM1	GM2	GM3	GM4	GM5	Uspjeh[%]
Boje	Boja1	7	3	2	2	4	<b>36</b>
	Boja2	8	5	3	6	8	<b>60</b>
	Boja3	3	6	4	10	7	<b>60</b>
	Boja4	5	6	3	10	4	<b>56</b>
	Boja5	7	9	9	8	7	<b>80</b>
	Boja6	7	9	5	9	1	<b>62</b>
	Uspjeh [%]	<b>61,7</b>	<b>63,3</b>	<b>43,3</b>	<b>75,0</b>	<b>51,7</b>	<b>59,0</b>
Brojevi	Broj1	10	9	9	10	8	<b>92</b>
	Broj2	9	9	6	9	3	<b>72</b>
	Broj3	7	10	7	10	2	<b>72</b>
	Broj4	1	1	1	2	0	<b>10</b>
	Broj5	2	1	0	8	9	<b>40</b>
	Broj6	0	0	0	5	4	<b>18</b>
	Broj7	9	9	9	10	10	<b>94</b>
	Broj8	9	9	7	9	7	<b>82</b>
	Broj9	10	9	10	8	0	<b>74</b>
	Broj10	4	4	2	0	0	<b>20</b>
	Broj11	0	0	0	0	0	<b>0</b>
	Broj12	8	4	10	4	6	<b>64</b>
	Broj13	9	10	10	10	9	<b>96</b>
	Broj14	7	10	9	10	7	<b>86</b>
	Uspjeh [%]	<b>60,7</b>	<b>60,7</b>	<b>57,1</b>	<b>67,9</b>	<b>46,4</b>	<b>58,6</b>
Riječi	Rijec1	10	10	10	10	10	<b>100</b>
	Rijec2	8	9	8	8	8	<b>82</b>
	Rijec3	0	0	0	2	0	<b>4</b>
	Rijec4	3	7	6	9	7	<b>64</b>
	Rijec5	6	7	2	9	4	<b>56</b>
	Rijec6	4	6	0	10	7	<b>54</b>
	Rijec7	1	2	1	2	1	<b>14</b>
	Rijec8	3	8	2	9	0	<b>44</b>
	Rijec9	7	3	2	3	1	<b>32</b>
	Rijec10	2	3	1	10	7	<b>46</b>
	Rijec11	3	0	0	7	5	<b>30</b>
	Rijec12	4	7	7	8	4	<b>60</b>
	Rijec13	8	6	0	6	5	<b>50</b>
	Rijec14	8	10	10	10	9	<b>94</b>
	Rijec15	10	10	10	10	10	<b>100</b>
	Rijec16	9	9	7	10	10	<b>90</b>
	Rijec17	3	7	6	3	4	<b>46</b>
	Rijec18	0	0	0	4	1	<b>10</b>
	Rijec19	7	8	7	7	8	<b>74</b>
	Rijec20	8	7	3	6	6	<b>60</b>
	Rijec21	2	1	1	0	0	<b>8</b>
	Rijec22	9	8	6	7	6	<b>72</b>
	Rijec23	1	2	0	1	4	<b>16</b>
	Rijec24	10	7	10	10	10	<b>94</b>
	Rijec25	10	10	10	8	9	<b>94</b>
	Rijec26	0	2	1	6	5	<b>28</b>
	Rijec27	6	0	2	5	1	<b>28</b>
	Rijec28	4	0	0	5	1	<b>20</b>
	Rijec29	5	4	9	5	7	<b>60</b>
	Rijec30	2	1	1	1	0	<b>10</b>
Uspjeh [%]	<b>51,0</b>	<b>51,3</b>	<b>40,7</b>	<b>63,7</b>	<b>50,0</b>	<b>51,3</b>	
Uspjeh [%]	55,0	55,4	45,6	66,2	49,2	54,3	

Procenat prepoznavanja po pojedinačnim govornicima je najmanji za govornika GM3 (45,6%), a najveći za govornika GM4 (66,2%). Na Sl. 2 je prikazan procenat prepoznavanja za ženske govornike u zavisnosti od grupe riječi (a) i pojedinačnog govornika (b).

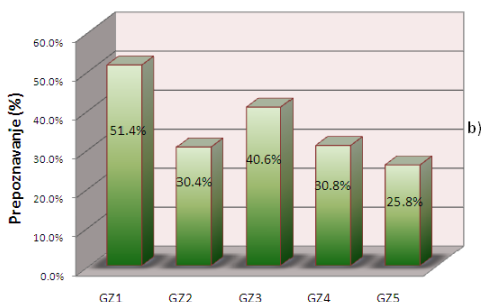
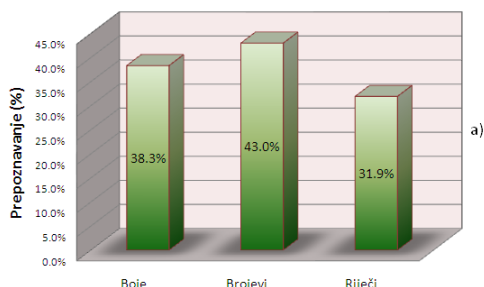
TABELA III. USPJEH U PREPOZNAVANJU RIJEČI ŽENSKIH GOVORNIKA

		GZ1	GZ2	GZ3	GZ4	GZ5	Uspjeh[%]
Boje	Boja1	4	1	0	0	0	<b>10</b>
	Boja2	10	9	10	5	7	<b>82</b>
	Boja3	1	2	4	3	5	<b>30</b>
	Boja4	7	5	6	1	6	<b>50</b>
	Boja5	1	3	1	1	0	<b>12</b>
	Boja6	8	1	9	1	4	<b>46</b>
	Uspjeh [%]	<b>51,7</b>	<b>35,0</b>	<b>50,0</b>	<b>18,3</b>	<b>36,7</b>	<b>38,3</b>
Brojevi	Broj1	9	4	5	7	5	<b>60</b>
	Broj2	10	8	8	7	4	<b>74</b>
	Broj3	10	9	10	7	2	<b>76</b>
	Broj4	0	0	1	0	0	<b>2</b>
	Broj5	4	1	1	0	0	<b>12</b>
	Broj6	0	0	0	3	0	<b>6</b>
	Broj7	9	10	10	9	10	<b>96</b>
	Broj8	9	6	8	6	3	<b>64</b>
	Broj9	7	5	2	1	1	<b>32</b>
	Broj10	5	1	0	0	0	<b>12</b>
	Broj11	0	0	0	0	0	<b>0</b>
	Broj12	3	8	6	8	1	<b>52</b>
	Broj13	8	4	10	8	2	<b>64</b>
	Broj14	9	6	4	6	1	<b>52</b>
	Uspjeh [%]	<b>59,3</b>	<b>44,3</b>	<b>46,4</b>	<b>44,3</b>	<b>20,7</b>	<b>43,0</b>
Riječi	Rijec1	10	10	10	10	10	<b>100</b>
	Rijec2	9	6	10	7	9	<b>82</b>
	Rijec3	0	0	0	0	0	<b>0</b>
	Rijec4	7	0	4	1	0	<b>24</b>
	Rijec5	7	0	9	2	0	<b>36</b>
	Rijec6	4	0	0	0	0	<b>8</b>
	Rijec7	3	0	0	0	0	<b>6</b>
	Rijec8	2	0	0	0	0	<b>4</b>
	Rijec9	5	0	1	2	0	<b>16</b>
	Rijec10	0	0	0	0	0	<b>0</b>
	Rijec11	4	0	0	0	0	<b>8</b>
	Rijec12	5	0	3	2	0	<b>20</b>
	Rijec13	0	0	0	0	0	<b>0</b>
	Rijec14	10	4	10	10	10	<b>88</b>
	Rijec15	10	6	7	10	10	<b>86</b>
	Rijec16	10	10	10	10	10	<b>100</b>
	Rijec17	1	1	0	0	0	<b>4</b>
	Rijec18	2	0	0	0	1	<b>6</b>
	Rijec19	10	1	1	6	2	<b>40</b>
	Rijec20	7	1	4	0	0	<b>24</b>
	Rijec21	0	0	0	0	0	<b>0</b>
	Rijec22	2	7	3	2	0	<b>28</b>
	Rijec23	0	0	0	0	0	<b>0</b>
	Rijec24	10	8	10	10	5	<b>86</b>
	Rijec25	10	7	6	3	9	<b>70</b>
	Rijec26	4	0	10	1	7	<b>44</b>
	Rijec27	3	0	1	2	0	<b>12</b>
	Rijec28	0	0	0	0	0	<b>0</b>
	Rijec29	8	8	9	3	5	<b>66</b>
	Rijec30	0	0	0	0	0	<b>0</b>
Uspjeh [%]	<b>47,7</b>	<b>23,0</b>	<b>36,0</b>	<b>27,0</b>	<b>26,0</b>	<b>31,9</b>	
Uspjeh [%]	51,4	30,4	40,6	30,8	25,8	35,8	

Procenat prepoznavanja po grupama je najmanji za balansirane riječi (31,9%), a najveći za riječi brojeva (43%). Procenat prepoznavanja po pojedinačnim govornicima je najmanji za govornika GZ5 (25,8%), a najveći za govornika GZ1 (51,4%).



Slika 1. Procenat prepoznavanja za a) grupe riječi i b) pojedinačne govornike kod muških govornika



Slika 2. Procenat prepoznavanja za a) grupe riječi i b) pojedinačne govornike kod ženskih govornika

## V. ZAKLJUČAK

Prikazani eksperimentalni rezultati prepoznavanja šapata pomoću skrivenih Markovljevih modela predstavljaju nastavak istraživanja prepoznavanja šapata kao specifičnog oblika verbalne komunikacije. Motiv za ovakvu vrstu istraživanja se prevashodno ogleda u veoma visokoj razumljivosti šapata. Rezultati prikazani u ovom radu pokazuju značajno smanjenje performansi sistema za automatsko prepoznavanje kontinualnog govora baziranog na HMM ukoliko se umjesto kontinualnog govora normalnog intenziteta koriste izolovano izgovorene riječi u šapatu. Dalja istraživanja prepoznavanja šapata će se ogledati u sinergiji

skrivenih Markovljevih modela sa neuralnim mrežama i (ili) DTW algoritmom, kao i izbor optimalnih akustičkih obilježja koji će doprinijeti poboljšanju prepoznavanja govora u šapatu. Tekuća istraživanja sa neuralnim mrežama [8] pokazuju slične rezultate, a u toku su i eksperimenti sa DTW algoritmom.

## ZAHVALNICA

Ovim putem iskazujemo posebnu zahvalnost preduzeću "AlfaNum – Govorne tehnologije" i prof. Vladi Deliću na ustupljenom softveru potrebnom za istraživanje.

## LITERATURA

- [1] T. Ito, K. Takeda, F. Itakura, "Analysis and Recognition of Whispered speech," *Speech Communication*, 2005, pp.129-152.
- [2] C. Zhang, J.H.L. Hansen, "Analysis and classification of Speech Mode: Whisper through Shouted," *Interspeech 2007*, 2007, pp. 2289-2292.
- [3] S.T. Jovičić, "Formant feature differences between whispered and voiced sustained vowels," *ACUSTICA - Acta Acoustica*, vol. 84, no. 4, 1998, pp. 739-743.
- [4] I. Eklund, H. Traunmuller, "Comparative Study of Male and Female Whispered and Phonated Versions of the Long Vowels of Swedish," *Phonetica*, 1996, pp. 1-21.
- [5] S.T. Jovičić, M.M. Đorđević, "Percepcija fonema u šapatu: identifikacija i konfuzija", poglavlje u knjizi S.T. Jovičić, M.Sovilj (ed.): *Govor i jezik: interdisciplinarna istraživanja*, II; Centar za unapređenje životnih aktivnosti i Institut za eksperimentalnu fonetiku i patologiju govora, 2008.
- [6] J. Holms, W. Holms, "Speech Synthesis and Recognition," Taylor & Francis, London, 2001.
- [7] S.T. Jovičić, Z. Kašić, M. Đorđević, M. Vojnović, M. Rajković, J.Savković, "Korpus psiho-emosivnog govora u srpskom jeziku," poglavlje u knjizi S.T. Jovičić, M. Sovilj (urednici): *Govor i jezik: interdisciplinarna istraživanja srpskog jezika*, I; 2004, str. 20-45.
- [8] Đ.T. Grozdić, B. Marković, J. Galić, S.T. Jovičić, "Primena neuralnih mreža u prepoznavanju govora u šapatu," Telekomunikacioni forum TELFOR, 2012, pp. 728-731
- [9] [http://alfanum.ftn.uns.ac.rs/frameset\\_results\\_srpski.htm](http://alfanum.ftn.uns.ac.rs/frameset_results_srpski.htm), posjećeno januara 2013. godine.
- [10] D. Pekar, R. Obradović, V. Delić, "Programski paket AlfaNumCASR sistem za prepoznavanje kontinualnog govora" Digitalna obrada govora i slike DOGS, 2002, pp. 49-56
- [11] N. Jakovljević, D.Pekar, "Description of Training Procedure for AlfaNum Continuous Speech Recognition System", EUROCON, 2005, pp.1646-1649

## ABSTRACT

This paper presents the results of an experimental research of recognition of whispered speech based on application of Hidden Markov Models (HMM), trained with normally phonated speech. The paper also describes the speech database of words that were spoken in whispered and normal manner, which was especially created for this study. Preliminary results show decline of performance for whispered speech compared to normally phonated speech.

## APPLICATION OF HIDDEN MARKOV MODELS IN WHISPERED SPEECH RECOGNITION

Jovan Galić, Milan Popović, Branko Marković,  
Đorđe T. Grozdić, Slobodan T. Jovičić