

UDK 811.163.42'34
811.163.42'282:811.163.42'242
811.163.42-055.1'282:811.163.42-055.2'282
Izvorni znanstveni članak
Prihvaćeno za tisak 17. 4. 2014.

Gordana Varošaneć-Škarić
Filozofski fakultet
gvarosan@ffzg.hr

Iva Pavić
Filozofski fakultet
ipavic2@ffzg.hr

Gabrijela Kišiček
Filozofski fakultet
gkisicek@ffzg.hr

Indeksi sličnosti i različitosti kod govornika hrvatskoga jezika u nefiltriranim i filtriranim uvjetima

Svrha istraživanja bila je usporedba akustičko–statističkih postupaka indeksa sličnosti (R) i različitosti (SDDD) na temelju dugotrajnoga prosječnoga spektra govora (LTASS), dobivenih u kontroliranim uvjetima na temelju standardiziranih tekstova i spontanoga govora govornika hrvatskoga jezika (N = 86) iz osam većih hrvatskih gradova. Gradovi su birani ciljano kao predstavnici štokavskoga (Zadar, Šibenik, Dubrovnik), čakavskoga (Rijeka, Pula), kajkavskoga (Zagreb, Čakovec) te čakavsko–štokavskoga varijeteta (Split). Montirani su uzorci (oko 60 s) za svakoga govornika u oba govorna uvjeta te su uspoređivani LTASS u nefiltriranim (0 – 10 kHz) i filtriranim uvjetima (0,8 – 4 kHz). Uspoređivani su indeksi R i SDDD unutar svakoga govornika između čitanog i spontanoga govora (unutargovorne razlike) i između različitih govornika (međugovorne razlike), posebno za muškarce i žene.

Rezultati unutargovornih razlika pokazuju da se prosječne vrijednosti indeksa R u nefiltriranim uvjetima kod muškaraca i žena kreću od 0,94 kod muških glasova pri čitanju do 0,98 kod žena u oba uvjeta. Rezultati međugovornih razlika pokazuju niže prosječne vrijednosti indeksa R u istim uvjetima: od 0,86 u spontanom do 0,94 u čitačem govoru kod žena. Prosječne vrijednosti indeksa R u filtriranim uvjetima (unutargovorno) kreću se od 0,83 kod žena i muškaraca u spontanom govoru do 0,95 kod žena i muškaraca pri čitanju. Prosječne vrijednosti indeksa sličnosti (međugovorno) znatno su niže, od 0,57 u muškome spontanom govoru do 0,9 kod žena u čitanju.

Prosječne vrijednosti indeksa SDDD u nefiltriranim uvjetima (unutargovorno) kreću se unutar nižih vrijednosti, od 2,27 kod žena do 3 kod muškaraca pri čitanju. Indeks SDDD (međugovorno) pokazuje prosječno više vrijednosti, od 4,75 u ženskome čitačem govoru i muškome spontanom govoru do 5,12 kod muškaraca pri čitanju. U filtriranim uvjetima (unutargovorno) indeks SDDD kreće se u rasponu nižih vrijednosti, od 2,14 kod muškaraca pri čitanju do 3,01 kod žena u spontanom govoru. Očekivano, rezultati međugovorno pokazuju prosječno više vrijednosti indeksa SDDD, od 3,06 za ženski do 4,71 za muški čitaći govor.

Rezultati unutargovornih varijacija pokazuju statistički značajne razlike indeksa sličnosti kod žena ($p < 0,0001$) i muškaraca ($p < 0,05$) u spontanom i čitaćem govoru. Rezultati međugovornih varijacija pokazuju značajne razlike indeksa sličnosti kod muških ($p < 0,0001$ pri čitanju i $p < 0,0001$ u spontanom govoru) i ženskih govornika ($p < 0,0001$ pri čitanju i $p < 0,0001$ u spontanom govoru) te značajne razlike indeksa SDDD kod muških ($p < 0,0001$ pri čitanju i $p < 0,0001$ u spontanom govoru) i ženskih govornika ($p < 0,0001$ pri čitanju i $p < 0,0001$ u spontanom govoru).

Rezultati ovoga istraživanja pokazuju da su akustičko–statistički postupci indeksa sličnosti i različitosti dobri kao pomoćna metoda kod prepoznavanja govornika u forenzično–fonetskim vještačenjima. Također, rezultati pokazuju da se u forenzično–fonetskim slučajevima ne smiju zanemariti okogovorni uvjeti.

1. UVOD

1.1. Značenje akustičkih mjerenja u prepoznavanju govornika

Od početaka praktične primjene fonetske znanosti u forenzične svrhe tijekom 20. stoljeća znanstvenici propituju značenje akustičkih postupaka pri fonetskom vještačenju. Početne lakonske tvrdnje da su akustičke slike govora, tj. spektrogrami »otisci glasa« poput *otisaka prstiju* u forenzici, srušene su znanstvenim pristupom akademskih fonetičara koji su bili svjesni višedimenzionalnosti govora te velikih razlika u okogovornim uvjetima. Dakle, mit o 100%–tnoj sigurnosti akustičkih podataka pri prepoznavanju govornika srušen je već u 50–im i ranim 60–im godinama, a i poslije znanstvenici upozoravaju da je neodgovorno tvrditi kako su spektrogrami »otisci prstiju«, primjerice Baldwin i French (1990). Oni se zalažu za akustičke metode koje su znatno razlikovnije i primjenjivije u forenzičnoj fonetici, primjerice frekvencijska struktura F0 i viših formantata, koja je do danas tradicionalna akustička metoda neizostavna u forenzičnim istraživanjima i kao potpora slušnoj fonetskoj analizi. Razvojem digitalnih računalnih programa unaprijedeno je i ubrzano mjerenje akustičkih parametara glasa i govora i ona su danas, možda više nego ikada, nezaobilazna u forenzičnoj fonetici, ali još nisu odlučujuća u fonetskom vještačenju, tj. pri prepoznavanju glasova govornika. Prijepori oko njihove samostalne uloge pri prepoznavanju prisutni su među vodećim forenzičnim fonetičarima u praksi i u recentnim znanstvenim radovima. Neki automatskom prepoznavanju glasova daju prednost, poput Künzela (2010, 2013), uz uzimanje svih mogućih uvjeta u izračune, dok se drugi koriste metodom automatskog prepoznavanja jedino uz obvezno slušno prepoznavanje s pomoću razvijenih percepcijskih fonetskih protokola, poput Holliena (2002) i Frencha (2013), jer su svjesni da osobine vokalnog trakta imaju ograničenja u akustičkoj fonetici, bez obzira na to što pojedini rezultati automatskog prepoznavanja obećavaju. I u hrvatskoj forenzičnoj fonetici uobičajeno se akustičke metode rabe samo

kao dodatne slušnim percepcijskim postupcima u prepoznavanju govornika (Varošaneć-Škarić 2008).

Kako su računalni sustavi automatskog prepoznavanja vrlo skupi, time gotovo nedostupni, uz gore navedene ograde, znanstvenici se koriste i dostupnijim akustičkim programima, poput Praata (Boersma i Weenink 2009), koji se uostalom rabi i u vodećim fonetsko-forezičnim laboratorijima, poput onoga Petera Frencha u Yorku te statističko-akustičkim postupcima poput indeksa sličnosti (R) i različitosti (SDDD) autora Harmegnies i Landercyja (Harmegnies i Landercy 1985; Harmegnies 1995) budući da se programi računanja dugotrajnoga prosječnog spektra govora lako mogu prilagoditi navedenom akustičko-statističkom postupku, a pokazali su se zanimljivima u usporedbi istoga glasa u različitim okogovornim uvjetima (Varošaneć-Škarić 2008; Varošaneć-Škarić i Kišiček 2012), usporedbi različitih i istih glasova u filtriranim i nefiltriranim uvjetima (Varošaneć-Škarić i Bićanić 2007) te u istraživanjima umjetničkih glasova pri različitim postavljanjima (Varošaneć-Škarić 2013). Poznati forezični fonetičar Peter French (2013) koristi se automatskim sustavima za prepoznavanje govornika, primjerice ASR-om, što je program koji uspoređuje identificiranu snimku osobe, a dobivena obilježja uspoređuje sa statističkim modelom osumnjičenikova glasa da bi se odredila mjera sličnosti / različitosti između dva glasa. Iako takvi sustavi daju obećavajuće rezultate u prepoznavanju, French navodi da urođena obilježja pojedinca mogu biti izmijenjena pod utjecajem jezične ili dijalektalne zajednice, pa dolazi do sličnosti među njezinim pripadnicima, što je stvarno ograničenje u automatskim sustavima prepoznavanja. Stoga navodi da sustavi automatskog prepoznavanja govornika ne mogu zamijeniti uobičajenu fonetsku i akustičku analizu (French 2013). Künzel (2013) upozorava da su pri automatskom prepoznavanju govornika nužne veće međujezične baze, ali i baze s različitim vrstama forezičnih kanala različitih transmisija (GSM, Skype – VoIP transmisija). Također upozorava da bi se zbog različitih vrsta sustava automatskog prepoznavanja svaki sustav trebao eksperimentalno ispitati da bi se utvrdilo kolika je zapravo jezična neovisnost u prepoznavanju. S tvrdnjama obojice potonjih autora potpuno se slažemo. Stoga planiramo u bližoj budućnosti ispitati kakvo će biti prepoznavanje govornika u određenom sustavu automatskog prepoznavanja (ciljano BATVOX) kad uz govornike hrvatskog jezika budu uključeni govornici srpskog, bošnjačkog i crnogorskog jezika. Spomenuti sustav BATVOX tako je primjerice testiran na korpusu britanskih govornika (Harrison i French 2010). U ovome radu željeli smo ispitati akustičko-statističke metode indeksa sličnosti i različitosti kod hrvatskih govornika, unutargovorne, međugovorne te u uvjetima filtriranog i nefiltriranoga govora.

1.2. Svrha uporabe indeksa sličnosti i različitosti u forezičnoj fonetici

U ovome radu željelo se s pomoću akustičko-statističkih postupaka indeksa sličnosti i indeksa različitosti na temelju dugotrajnih prosječnih spektara govora (LTASS, program AS) ispitati na većem broju govornika hrvatskoga

jezika kolika je razlikovnost potonjeg postupka u različitim okogovornim uvjetima spontanoga govora i čitanja te u filtriranim i nefiltriranim uvjetima kod muških i ženskih govornika. Praktična je svrha istraživanja da podatci dobiveni na većem korpusu budu polazište za interpretaciju vrijednosti indeksa sličnosti i različitosti, naravno, kao pomoćna metoda u slušnoj fonetskoj procjeni sličnosti – razlike parova glasova i uobičajenim fonetskim akustičkim mjerenjima.

1.3. Pretpostavke o sličnostima i razlikama unutar govornika (unutargovorno) i između govornika (međugovorno)

Postavljene su ove pretpostavke:

H1: očekuju se statistički značajno veće međugovorne nego unutargovorne razlike kod ispitanika

H2: očekuju se veće razlike u spontanom govoru nego u čitanju

H3: očekuju se veće razlike u skupini ženskih govornika (pretpostavka je da žene imaju razvedeniju intonaciju).

2. POSTUPAK

2.1. Ispitanici

U istraživanju se analizirao govor 86 govornika hrvatskoga jezika iz osam većih hrvatskih gradova. Gradovi zastupljeni u istraživanju prezentiranom u ovom radu jesu: Čakovec, Zagreb, Pula, Rijeka, Zadar, Šibenik, Split i Dubrovnik. Ciljano su zastupljeni govori svih triju hrvatskih narječja. Područje kajkavskog narječja predstavljali su govori Čakovca i Zagreba, dok su urbana središta štokavskoga narječja predstavljena gradovima srednje i južne Dalmacije: Šibenikom, Zadrom i Dubrovnikom. Naposljetku, čakavsko je narječje zastupljeno Rijekom i Pulum, a grad Split uzet je kao primjer područja dodira štokavskog i čakavskog narječja.

Ispitanici su se međusobno razlikovali po spolu, dobi te stupnju i vrsti obrazbe. Tako se korpus sastojao od 58 govornika ženskog spola i 28 ispitanika muškog spola. Prema kriteriju dobi govornici se mogu podijeliti na mlade i starije. Mlađi su govornici bili srednjoškolci ili osobe do 20. godine. Pritom valja naglasiti da se pri izboru mlađih govornika obraćala pozornost i na podrijetlo njihovih roditelja, s obzirom na to da govor u roditeljskom domu uvelike utječe na govor ispitivane osobe. Prema tome, u istraživanje su uključeni samo oni govornici koji su rođeni i odrasli u istom gradu kao i njihovi roditelji. Izabrani su ispitanici iz urbanih sredina, jer se pretpostavlja da razlike u vokalskom spektru nisu tako velike kao u ruralnim područjima. Time su se željele izbjeći veće razlike u tome dijelu spektra, a da budu zastupljeni tipovi općega hrvatskoga govora. Stariji su govornici bile osobe starije od 50 godina. Za starije se govornike pretpostavilo da je eventualan utjecaj grada u kojemu su studirali ili

privremeno živjeli smanjen, pa su se kriteriji odabira starijih ispitanika sastojali u rođenju i provođenju najvećeg dijela svoga života u istom gradu. U skupini starijih govornika zastupljene su sve razine stručne spreme, niža, srednja i visoka, raznolikih vrsta naobrazbe.

2.2. Tehnički uvjeti snimanja i korpus (materijal) govornih uzoraka istraživanja

Svi govornici snimani su istom opremom (Maranz PMD 660 i mikrofon AKGC 391) u sličnim uvjetima: u prostoriji sa smanjenom bukom (radi akustičke obrade snimljenog materijala) i mikrofonom udaljenim 10 cm pod kutom od 45°. Dio govornoga materijala (spontani govor) za analizu prikupljen je prema metodologiji opisanoj u sociolingvističkim studijama pionira sociolingvističkih kvantitativnih studija Labova (1972, 2006), ali i onima Wolframa i Fasholda (1997), prema kojima je temelj sociolingvističkih i sociofonetskih istraživanja spontani govor. Prema tome, govornici su se trebali predstaviti prema zadanim smjernicama, opisati što rade u slobodno vrijeme, kakvo im je bilo djetinjstvo i kakve su igre igrali. Potom su opisivali svoj grad, njegove dobre i loše strane te uspoređivali kakav je bio nekada, a kakav je sada. Naposljetku, govorili su o svojim interesima. Na taj način, tematikom bliskom svakom govorniku poticao se spontani govor, koji je trajao od 10 do 15 minuta. Osim spontanoga govora korpus istraživanja sadržavao je i zvučni materijal čitanja istoga teksta u trajanju od 60 do 70 s ovisno o tempu čitanja.

2.3. Metodologija i statistička obrada rezultata

U provedenom istraživanju analiziran je korpus koji se sastojao od govora ispitanika u spontanom izričaju i za vrijeme čitanja teksta. Učinjeni su LTASS za svakoga ispitanika zasebno u spontanom govoru i čitanju. Obje su kategorije zvučnih materijala frekvencijski filtrirane od 800 do 4000 Hz. Ako uzmemo u obzir kriterij spola (ženski i muški), vrste zvučnog materijala (spontani govor i tekst) te frekvencijske karakteristike (filtrirano: 800 – 4000 Hz i nefiltrirano: 0 – 10000 Hz), određene su ove kategorije u analizi:

- ženski filtrirani nefrikativni test
- ženski filtrirani spontani govor
- ženski nefiltrirani nefrikativni test
- ženski nefiltrirani spontani govor
- muški filtrirani nefrikativni test
- muški filtrirani spontani govor
- muški nefiltrirani nefrikativni test
- muški nefiltrirani spontani test.

Dakle, za četiri uvjeta (vrsta govora i vrsta snimki: filtriranost / nefiltriranost) za 86 govornika dobile su se ukupno 344 zvučne snimke govora.

U svim gore spomenutim kategorijama promatrane su vrijednosti indeksa sličnosti (R) i indeksa različitosti (SDDD), na temelju kojih su određene njihove minimalne i maksimalne vrijednosti, mjere centralne tendencije, odnosno prosječne vrijednosti (M) te mjere varijabilnosti (S.D.). Između ostaloga, govornici su uspoređivani s drugim ispitanicima istoga spola, s drugim ispitanicima suprotnoga spola te sami sa sobom. U potonjoj kategoriji iste su kategorije uspoređivane na dvama različitim uzorcima iz iste vrste zvučnog materijala.

Tako se primjerice 60 sekundi spontanoga govora govornika XY uspoređivalo s drugih 60 sekundi spontanoga govora istoga govornika. Drugim riječima, žene i muškarci analizirani su unutar svoje skupine (Ž : Ž i M : M), unutar samoga govornika ($XY_1 : XY_2, YZ_1 : YZ_2$) te između dviju skupina (Ž : M).

Naposljetku, statističkim postupkom *t*-testa utvrđivalo se je li razlika između uspoređivanih kategorija statistički značajna i prema tome moguća u široj populaciji ili je pak slučajna. Uspoređivane su vrijednosti indeksa sličnosti i različitosti u čitaćem tekstu i spontanom govoru u skupinama muškaraca i žena, prema kriteriju filtriranosti odnosno nefiltriranosti i u okviru svakoga govornika posebno i unutar skupina. Također, uspoređivane su vrijednosti indeksa sličnosti i različitosti u spontanom govoru i tekstu, filtrirano i nefiltrirano prema kriteriju spola.

3. REZULTATI

Rezultati unutar govornika (unutargovorno) pokazuju da se prosječne vrijednosti indeksa R u nefiltriranim uvjetima kreću od 0,94 kod muških glasova pri neutralnom čitanju do 0,98 kod žena u oba uvjeta. Rezultati međugovorno pokazuju niže prosječne vrijednosti indeksa R u istim uvjetima: od 0,86 u spontanom govoru do 0,94 u čitanju teksta kod žena (tablice 1. i 2.). Prosječne vrijednosti indeksa R u filtriranim uvjetima (unutargovorno) kreću se od 0,83 kod žena i muškaraca u spontanom govoru do 0,95 kod žena i muškaraca pri čitanju (tablica 3.). Rezultati (međugovorno) prosječnih vrijednosti indeksa sličnosti znatno su nižih prosječnih vrijednosti, od 0,57 u muškome spontanom govoru do 0,79 kod žena u čitanju (tablica 4.).

Tablice 1. i 2. Statističke vrijednosti indeksa R u unutargovornim i međugovornim skupinama, u nefiltriranim uvjetima

NF – unutargovorno	Čt		Sg	
	Indeks R			
Žene	min.	0,95	min.	0,93
	maks.	0,99	maks.	0,99
	M	0,98	M	0,98
	S. D.	0,01	S. D.	0,01
	Indeks R			
Muškarci	min.	0,8	min.	0,88
	maks.	0,98	maks.	0,99
	M	0,94	M	0,95
	S. D.	0,04	S. D.	0,02

NF – međugovorno	Čt		Sg	
	Indeks R			
Žene	min.	0,82	min.	0,45
	maks.	0,99	maks.	0,97
	M	0,94	M	0,862
	S. D.	0,02	S. D.	0,063
	Indeks R			
Muškarci	min.	0,79	min.	0,69
	maks.	0,98	maks.	0,96
	M	0,93	M	0,87
	S. D.	0,03	S. D.	0,056

Čt – čitaći tekst
 Sg – spontani govor
 NF – nefiltrirani uvjeti
 F – filtrirani uvjeti
 M – prosječna vrijednost

Tablice 3. i 4. Statističke vrijednosti indeksa R u unutargovornim (lijevo) i međugovornim (desno) skupinama, u filtriranim uvjetima

NF – unutargovorno	Čt		Sg	
	Indeks R			
Žene	min.	0,64	min.	0,3
	maks.	0,99	maks.	0,98
	M	0,95	M	0,83
	S. D.	0,06	S. D.	0,12
	Indeks R			
Muškarci	min.	0,82	min.	0,24
	maks.	0,99	maks.	0,97
	M	0,95	M	0,83
	S. D.	0,04	S. D.	0,02

NF – međugovorno	Čt		Sg	
	Indeks R			
Žene	min.	0,03	min.	0,004
	maks.	0,97	maks.	0,96
	M	0,79	M	0,68
	S. D.	0,02	S. D.	0,15
	Indeks R			
Muškarci	min.	0,003	min.	0,03
	maks.	0,94	maks.	0,9
	M	0,73	M	0,57
	S. D.	0,15	S. D.	0,11

Prosječne vrijednosti indeksa SDDD u nefiltriranim uvjetima (unutargovorno) kreću se unutar nižih vrijednosti, od 2,27 kod žena do 3 kod muškaraca pri čitanju (tablica 5.). Rezultati indeksa SDDD (međugovorno) pokazuju prosječno više vrijednosti, od 4,75 u ženskome čitaćem govoru i muškome spontanom govoru do 5,12 u muškome čitaćem govoru (tablica 6.). Rezultati u filtriranim uvjetima indeksa SDDD (unutargovorno) kreću se u rasponu nižih vrijednosti, od 2,14 kod muškaraca pri čitanju do 3,01 kod žena u spontanom govoru (tablica 7.). Očekivano, rezultati međugovorno pokazuju prosječno više vrijednosti indeksa SDDD, od 3,06 za ženski do 4,71 za muški čitaći govor (tablica 8.). Standardna devijacija prosječnih rezultata općenito je malena i kod muških i ženskih govornika, očekivano nešto veća u spontanom govoru.

Tablice 5. i 6. Statističke vrijednosti indeksa SDDD u unutargovornim (lijevo) i međugovornim (desno) skupinama, u nefiltriranim uvjetima

NF – unutargovorno	Čt		Sg	
	Indeks SDDD			
Žene	min.	1,52	min.	1,48
	maks.	3,81	maks.	4,22
	M	2,27	M	2,39
	S. D.	0,55	S. D.	0,69
	Indeks SDDD			
Muškarci	min.	1,9	min.	1,85
	maks.	5,04	maks.	4,47
	M	3	M	2,72
	S. D.	0,61	S. D.	0,68

NF – međugovorno	Čt		Sg	
	Indeks SDDD			
Žene	min.	2,19	min.	2,6
	maks.	9,83	maks.	11,53
	M	4,75	M	4,76
	S. D.	1,058	S. D.	1,116
	Indeks SDDD			
Muškarci	min.	2,51	min.	2,52
	maks.	9,57	maks.	8,7
	M	5,12	M	4,75
	S. D.	1,15	S. D.	1,12

Tablice 7. i 8. Statističke vrijednosti indeksa SDDD u unutargovornim (lijevo) i međugovornim (desno) skupinama, u filtriranim uvjetima

NF – unutargovorno	Čt		Sg	
	Indeks SDDD			
Žene	min.	1,35	min.	1,85
	maks.	4,22	maks.	5,75
	M	2,15	M	3,01
	S. D.	0,57	S. D.	0,74
	Indeks SDDD			
Muškarci	min.	1,3	min.	1,59
	maks.	3,99	maks.	4,04
	M	2,14	M	2,48
	S. D.	0,69	S. D.	0,58

NF – međugovorno	Čt		Sg	
	Indeks SDDD			
Žene	min.	1,68	min.	1,88
	maks.	9,55	maks.	9,11
	M	3,06	M	3,78
	S. D.	0,32	S. D.	0,12
	Indeks SDDD			
Muškarci	min.	2,04	min.	2
	maks.	10,32	maks.	9,08
	M	4,71	M	3,73
	S. D.	1,25	S. D.	1,16

Statistički značajne razlike (t-test) indeksa sličnosti unutargovorno zabilježene su kod žena ($p < 0,0001$; $p = 0,000005$; $t = 5,006$) i kod muškaraca ($p < 0,05$; $p = 0,0002$; $t = 4,235$; tablica 9.) u spontanom govoru te u čitaćem govoru ($p < 0,0001$; $p = 0,0000001$; $t = 6,148$ za ženske glasove i $p < 0,05$; $p = 0,003$; $t = 3,272$ za muške glasove; tablica 9.). U istim kategorijama unutargovorno, razlike u vrijednostima indeksa različitosti nisu se pokazale značajnima ($p > 0,05$; tablica 10.). Međutim, rezultati među skupinama (međugovorno) pokazali su da se indeks sličnosti statistički značajno razlikuje kod muških ($p < 0,0001$; $p = 0,753 \times 10^{-88}$; $t = 22,625$ pri čitanju i $p < 0,0001$; $p = 0,239 \times 10^{-89}$; $t = 22,880$ u spontanom govoru; tablica 11.) i ženskih govornika ($p < 0,0001$; $p = 0,9001 \times 10^{-245}$; $t = 118,015$ pri čitanju i $p < 0,0001$; $p = 0,813 \times 10^{-228}$; $t = 101,924$ u spontanom govoru; tablica 11.). Nešto niže vrijednosti indeksa različitosti, ali svejedno statistički značajne prisutne su i kod muških ($p < 0,0001$; $p = 0,0000027$; $t = 4,727$ pri čitanju i $p < 0,0001$; $p = 0,00000046$; $t = 5,086$) i kod ženskih govornika ($p < 0,0001$; $p = 0,220 \times 10^{-51}$; $t = 15,184$ pri čitanju i $p < 0,0001$; $p = 0,991 \times 10^{-41}$; $t = 13,368$).

Tablica 9. Rezultati t-testa za indekse R i SDDD u skupini unutargovorno, u nefiltriranim i filtriranim uvjetima, za žene i muškarce

ŽENE			MUŠKARCI		
	R	SDDD		R	SDDD
Čt	$p < 0,0001$ ($p = 0,000005$)	$p > 0,05$ ($p = 0,278$)	Čt	$p < 0,05$ ($p = 0,0002$)	$p > 0,05$ ($p = 0,190$)
	$t = 5,006$	$t = 1,095$		$t = 4,235$	$t = 1,342$
Sg	$p < 0,0001$ ($p = 0,0000001$)	$p > 0,05$ ($p = 0,956$)	Sg	$p < 0,05$ ($p = 0,003$)	$p > 0,05$ ($p = 0,176$)
	$t = 6,148$	$t = 0,056$		$t = 3,272$	$t = 1,389$

Tablica 10. Rezultati t-testa za indekse R i SDDD u skupini unutargovorno, između muškaraca i žena, za nefiltrirane i filtrirane uvjete

NEFILTRIRANO			FILTRIRANO		
	R	SDDD		R	SDDD
Čt	$p > 0,05$ ($p = 0,372$)	$p > 0,05$ ($p = 0,415$)	Čt	$p < 0,05$ ($p = 0,959$)	$p > 0,05$ ($p = 0,948$)
	$t = 0,898$	$t = 0,819$		$t = 0,052$	$t = 0,065$
Sg	$p < 0,05$ ($p = 0,014$)	$p < 0,05$ ($p = 0,075$)	Sg	$p < 0,05$ ($p = 0,950$)	$p < 0,05$ ($p = 0,00079$)
	$t = 2,510$	$t = 1,804$		$t = 0,063$	$t = 3,484$

Tablica 11. Rezultati t–testa za indekse R i SDDD u skupini međugovorno u nefiltriranim i filtriranim uvjetima, za žene i muškarce

ŽENE		
	R	SDDD
NF test	p<0,0001 (p = 0,90063 × 10 ⁻²⁴⁵)	p<0,0001 (p = 0,219519 × 10 ⁻⁵¹)
	t=118,015	t=15,184
SP test	p<0,0001 (p = 0,81335 × 10 ⁻²²⁸)	p<0,0001 (p = 0,989542 × 10 ⁻⁴¹)
	t=101,924	t=13,368
MUŠKARCI		
	R	SDDD
NF test	p<0,0001 (p = 0,753279 × 10 ⁻⁸⁸)	p<0,0001 (p = 0,0000027)
	t=22,625	t=4,727
SP test	p<0,0001 (p = 0,239054 × 10 ⁻⁸⁹)	p<0,0001 (p = 0,00000046)
	t=22,880	t=5,086

4. RASPRAVA

Prva pretpostavka potvrđena je, jer se pokazalo da su međugovorne razlike veće od unutargovornih. Rezultati ovoga istraživanja potvrdili su dosadašnje tvrdnje da se tek statističko podudaranje od 94% može uzeti u obzir kad govorimo o prepoznavanju govornika. S obzirom na to da takav rezultat znači veliki stupanj prepoznavanja u AP – SPID postupku (Hollien 2002), prosječni rezultati indeksa sličnosti koji su dobiveni u ovom istraživanju na temelju akustičko–statističkih metoda i na temelju LTASS unutar govornika u nefiltriranim uvjetima od 0,94 kod muških glasova pri neutralnom čitanju teksta do 0,98 kod žena u oba zvučna uvjeta, tj. u nefiltriranim i filtriranim uvjetima doista mogu biti korisni u forenzičnim prepoznavanjima kao dodatna metoda. Pri tome se ne smiju zanemariti različiti okogovorni uvjeti (buka, glasnoća, udaljenost od slušalice, odnosno mobitela do usta, prostorni uvjeti itd.), ali prije svega i različiti zvučni uvjeti samih prijenosnih aparata. Dakako, pojedinačni rezultati mogu biti ispod prosječnih vrijednosti, ali je bilo važno utvrditi koje su to vrijednosti da bismo mogli dati veću težinu prosječnim vrijednostima u prepoznavanju, budući da se prosječne vrijednosti unutar govornika (unutargovorno) kreću od 0,83 kod oba spola u spontanom govoru do 0,95 kod žena i muškaraca u čitanju. Važan je rezultat koji pokazuje značajno niže prosječne vrijednosti indeksa sličnosti između govornika (međugovorno) koji se kreću od

0,57 u muškome spontanom govoru do 0,9 kod žena pri čitanju. Ovi rezultati djelomično ruše pretpostavku 3 o većim razlikama među ženama. Kako razlika u intonaciji u spontanom govoru i čitanju nije bila u žarištu ovoga istraživanja, nećemo dalje raspravljati o tim rezultatima, ali zamjetno je da se možda radi o stereotipu da žene imaju razvedeniju intonaciju. Nagovještavamo samo da je to indeksikalni marker »ženskosti« koji je više vezan uz afektiranje nego uz stvarni govor žena, budući da je govorni uzorak sadržavao spontani govor u kojem se ženama i muškarcima nisu davale upute kako treba govoriti, tj. mogli su se koristiti paralingvističkim slojem glasa u emocionalnoj izražajnosti kako im je bilo prirodno.

Jednako je važno da se pokazala značajna razlika između rezultata prosječne vrijednosti indeksa različitosti (SDDD) u nefiltriranim uvjetima unutar govornika i među govornicima, tj. da rezultati unutargovorno pokazuju prosječno značajno niže vrijednosti od 2,27 kod žena do 3 kod muškaraca u čitanju, u odnosu na rezultate međugovorno od 4,75 u čitaćem govoru kod žena i muškome spontanom govoru do prosječno 5,12 u muškome čitaćem govoru, što je još jedna potvrda o tome da ženski govor nije toliko razlikovniji intonacijski koliko se pretpostavlja. Unutargovorna i međugovorna razlika zadržava se slijedno i u filtriranim uvjetima, što potvrđuje svrhu da se rezultati indeksa sličnosti i različitosti mogu iskoristiti u forenzične svrhe kao pomoćna metoda. Uz tradicionalne akustičke mjere koje se rabe u vještačenju snimki govora preko GSM uređaja (F0, F1 – F3), rezultati s pomoću statističko–akustičkih postupaka R i SDDD pokazali su se pouzdanima u istraživanju Varošaneć-Škarić i Bićanić (2007). U tom su istraživanju učinjeni LTASS za područje od 800 do 3500 Hz za 43 para istih i različitih glasova u stvarnim forenzičnim slučajevima u različitim govornim i okogovornim kontekstima. Očekivano su se skupine značajno razlikovale, najviše skupina muških govornika snimljena u kontroliranim studijskim uvjetima u nefiltriranom zvuku, potom u filtriranim uvjetima (800 – 3500 Hz) te glasovi snimljeni službenim prisluškivanjem. U stvarnim slučajevima u uvjetima različitih govornih konteksta i GSM transmisije pokazala se značajna razlika SDDD indeksa ($p < 0,01$) i indeksa sličnosti (R; $p = 0,01$) unutargovorno i međugovorno. U potonjem istraživanju pokazalo se da je moguć i veći raspon indeksa unutar istih glasova, od onih koji se manje razlikuju unutar sebe (R = 0,98; SDDD = 1,06) do onih koji se u paralingvističkom smislu vrlo razlikuju u različitim govornim uvjetima (R = 0,81; SDDD = 5,3).

U verifikaciji govornika bilingvalnog albansko–hrvatskoga govornika Varošaneć-Škarić (2008) dobila je prosječne rezultate indeksa sličnosti (R) za istoga govornika u različitim okogovornim uvjetima 0,9 i indeks različitosti (SDDD) od 2,8, što su dobri rezultati budući da su dobiveni na temelju službenog prisluškivanja telefona te filtriranih snimki govornika u razgovoru s fonetičarem na oba jezika. U sličnim okogovornim uvjetima u telefonskim razgovorima na oba jezika dobiven je najbolji rezultat indeksa sličnosti od 0,94 do čak 0,98, što se može interpretirati kao identifikacija uz ostale tradicionalne forenzično–fonetske metode (Varošaneć-Škarić 2008). Zamjetno je da se potonji rezultati podudaraju s rezultatima ovog istraživanja na velikom uzorku govornika, ali i navodi na oprez u interpretiranju. Ovo je istraživanje važno i stoga što se moglo doći do vrijednosti R i SDDD među najslabijim i najrazličitijim

glasovima. Prethodno navedeni rezultati pokazali su da su najbližnja dva para glasova (posebno muški i ženski) imala vrijednosti R od 0,94 do 0,97 (uzimajući u obzir da se muški i ženski glasovi uvijek uspoređuju zasebno), a vrijednosti SDDD od 1,75 do 2,04.

Indeksi sličnosti i različitosti u istraživanju Varošaneć-Škarić i Kišiček (2012) pokazali su unutar i između govornika rezultate slične onima u ovom istraživanju na manjem uzorku, tj. indeks sličnosti na dvije različite snimke iste osobe bio je manji unutar tražene osobe (0,89) nego između tražene osobe i sugovornika (R = 0,85), a indeks različitosti bio je znatno manji između različitih snimki tražene osobe (SDDD = 2,1) nego između tražene osobe i sugovornika (SDDD = 3,75).

Slijedno prethodnim spomenutim istraživanjima i rezultatima ovoga istraživanja proizlazi da se nužno moraju rabiti i slušni protokoli koji će uzeti u obzir i sve druge dimenzije govora te lingvističke markere, jer sam spektar neće o njima obavijestiti jednoznačno, a neke ne može uopće pokazati. Dakako, u akustičko–statističkim postupcima nije problem s najrazličitijim glasovima. I ovo istraživanje pokazuje da su u tim primjerima indeksi sličnosti očekivano manji od 0,66 do 0,79, a indeksi različitosti iznad 7: od 7,15 do 7,93. Dakle, nije problem kad se akustičkim i akustičko–statističkim postupcima uspoređuju glasovi različiti prema kvaliteti glasa i ostalim dimenzijama, nego kad se u forenzičnom uzorku nađu prosječni glasovi slični prema svim dimenzijama tona, kvalitete glasa, glasnoće, regionalnog podrijetla, naobrazbe itd. Tu mogu biti odlučujući neki lingvistički markeri, idiosinkratičnost govornika i segmentalan izgovor.

Ovim istraživanjem dobili smo i okvirne prosječne vrijednosti te raspone indeksa koji mogu biti polazište u interpretaciji rezultata u različitim uvjetima. Tu tvrdnju temeljimo na statistički značajnim rezultatima navedenim u prethodnom poglavlju, uzimajući u obzir da su tzv. »čisti« podaci vrlo rijetki u forenzičnome fonetskom vještačenju. Rodman i dr. (2002) podupiru metodologiju koja se bazira na prosječnim ukupnim spektralnim oblicima vokalnog trakta, budući da je njihovo istraživanje pokazalo da uprosječivanje spektralnih oblika umanjuje degradirajući učinak transmisijskog medija. Također upozoravaju da je lakše govoriti o identifikaciji kad se radi o tzv. »čistim« podacima, a u forenzičnoj fonetici moramo biti sposobni raditi s vrlo kratkim obrascima, pa treba rabiti akustičke metode da bi se moglo uspoređivati slično sa sličnim, ali i znati kako izolirati ili ublažiti kad imamo siromašan odnos signala i šuma. Na kraju, prepoznavanje ili identifikacija s pomoću akustičkih metoda mora biti neovisna o tekstu, a naše je istraživanje upravo potvrdilo tu tvrdnju u usporedbi unutargovornog i međugovornog na istim i različitim tekstovima. Raspravljajući o identifikaciji govornika u forenzičnoj fonetici, Nolan i Grigoras (2005) i Nolan (2007) upozoravaju na teškoće pri usporedbi glasova, kao što su nekonzistentnost uzoraka i unutar istoga govornika u akustičkom, stilističkom, sociolingvističkom smislu. U forenzičnoj praksi zbog takvih okolnosti vrlo je teško utvrditi identitet govornika, pa je obvezno koristiti se i tradicionalnim fonetskim tehnikama kao što su slušna i računalna akustička analiza. Nolan nadalje upozorava na ograničenja ovisna o vrsti transmisije, što ekspert mora uzeti u obzir. Stoga smo u ovom radu u filtriranju rabili pomalo neobičnu metodu filtriranja govornog signala od 800 do 4000 Hz, jer je to koji put u

forenzici približnije stvarnomu zvučnomu materijalu nego same karakteristike telefona i mobitela koje sežu u donjem dijelu od 350 do 400 Hz. Treba poznavati vrstu izobličenja do kojih dolazi i u policijskom snimanju prijenosnika, a da ne govorimo o samim prislušnim uređajima koji kadšto uopće ne mogu kontrolirati udaljenost prisluškivanih osoba ili pak o mogućim interferencijama.

5. ZAKLJUČAK

Zajednički rezultati za žene i muškarce potvrdili su pretpostavke da su međugovorne razlike statistički značajno veće od unutargovornih te da su veće razlike u spontanom govoru nego pri čitanju. Očekivano su veće vrijednosti indeksa sličnosti u nefiltriranim uvjetima nego u filtriranim, a indeksi različitosti veći su u nefiltriranim uvjetima međugovorno, za žene i unutargovorno.

Prosječne vrijednosti u nefiltriranim uvjetima indeksa sličnosti u rasponu su od 0,94 kod muškaraca pri neutralnom čitanju do 0,98 kod žena u filtriranim i nefiltriranim uvjetima.

Prosječne vrijednosti indeksa različitosti kreću se unutar nižih vrijednosti (unutargovorno) u nefiltriranim uvjetima od 2,27 kod žena do 3 kod muškaraca pri čitanju. Rezultati indeksa različitosti međugovorno prosječno su viših vrijednosti u rasponu od 4,75 kod žena u čitanju i muškome spontanom govoru do 5,12 kod muškaraca pri čitanju. Filtrirani uvjeti zbog manjeg broja spektralnih točaka usporedbe nešto su manje razlikovni, s tim da su rezultati međugovorno prosječno viših vrijednosti od 3,06 za žene do 4,71 za muškarce pri čitanju.

Pokazalo se da se ženski i muški govornici statistički značajno razlikuju u spontanom i čitaćem govoru i gledajući sveukupne rezultate, ne može se sa sigurnošću potvrditi pretpostavka da je ženski govor ukupno razlikovniji u govoru. Može se reći da se u unutargovornim razlikama kod žena u filtriranim i nefiltriranim uvjetima i u spontanom govoru i pri čitanju statistički više razlikuje indeks sličnosti ($p < 0,0001$) nego kod muških glasova ($p < 0,05$), ali indeks različitosti usporedbom unutargovornih razlika žena i muškaraca u filtriranome spontanom govoru podjednake je statističke značajnosti ($p < 0,05$). Međugovorne razlike statistički se značajno razlikuju kod žena i muškaraca, i to u usporedbi nefiltriranih i filtriranih uvjeta u spontanom govoru i čitanju (R: $p < 0,0001$; SDDD: $p < 0,0001$).

Kad je to moguće, trebalo bi u stvarnim forenzičnim slučajevima provesti i verifikacijski postupak snimanja osumnjičenika u kontroliranim uvjetima da bi se filtriranjem rezultati mogli usporediti sa snimkama službenog prisluškivanja koji mogu biti različitih transmisijskih obilježja. To je važno i zbog etičkog kodeksa fonetičara, koji se moraju oprezno odnositi u zaključcima prepoznavanja / neprepoznavanja govornika i koji bi trebali biti neovisni u zaključivanju o identifikaciji / neidentifikaciji govornika. Pomoćne akustičke i akustičko–statističke metode mogu biti korisne pri usporedbi s rezultatima slušnoga fonetskog prepoznavanja, a rezultati pokazuju da su indeksi sličnosti i različitosti dobra metoda te da se ne smiju zanemariti okogovorni uvjeti.

Zahvala

Ovaj je rad financijski pomoglo Ministarstvo znanosti, obrazovanja i sporta Republike Hrvatske (MZOS RH) u okviru projekta: Forenzična fonetika: slušno prepoznavanje i zvučna analiza glasova; broj 1–300–2679 (130–0000000–0786).

Referencije

- Baldwin, John R., Peter French (1990). *Forensic Phonetics*. London and New York: Pinter Publishers.
- Boersma, Paul, David Weenink (2009). Praat: doing phonetics by computer, version 5.1.20. Dostupno na: www.fon.hum.uva.nl/praat/. Datum pristupa: 1. prosinca 2013.
- French, Peter (2013). Forensic speaker comparison: man and machine – Forenzična usporedba govornika: čovjek i stroj. U: Vlašić Duić, J. i Varošanec-Škarić, G. (ur.) *Knjiga sažetaka – Istraživanja govora*, Zagreb: Filozofski fakultet Sveučilišta u Zagrebu, Hrvatsko filološko društvo. Osmi znanstveni skup s međunarodnim sudjelovanjem, od 5. do 7. prosinca 2013., str. 25–26.
- Harmegnies, Bernard (1995). Contribution à la caractérisation acoustique des sigmatismes – Etude de deux indices acoustico-statistiques. U: A. Braun i J.-P. Köster (ur.) *Studies in Forensic Phonetics*, 56–66. Trier: Wissenschaftlicher Verlag Trier.
- Harmegnies, Bernard, Albert Landercy (1985). Language Features in the Long-Term Average Spectrum. *Revue de Phonétique Appliquée* 73–74–75, 69–79.
- Harrison, Philip, Peter French (2010). Assessing the suitability of BATVOX for UK Casework or Evaluation of the BATVOX automatic speaker recognition system for use in UK based forensic speaker comparison casework Part II. U *Abstracts for the 19th Annual Conference of the International Association for Forensic Phonetics and Acoustics*, Trier, Germany, Department of Phonetics, University of Trier, str. 13.
- Hollien, Harry (2002). *Forensic Voice Identification*. San Diego: Academic Press.
- Künzel, Hermann J. (2010). Automatic Speaker Identification with Multilingual Speech Material. U *Abstracts for the 19th Annual Conference of the International Association for Forensic Phonetics and Acoustics*, Trier, Germany, Department of Phonetics, University of Trier, str. 20.
- Künzel, Hermann J. (2013). Automatic speaker recognition with cross-language speech material. *Journal of Speech, Language and the Law* 20.1, 21–44.
- Labov, William (1972). *Sociolinguistic Patterns*. Philadelphia: Philadelphia University of Pennsylvania Press.
- Labov, William (2006). A sociolinguistic perspective on sociophonetic research. *Journal of Phonetics* 34, 500–515.
- Nolan, Francis (1983, digitally printed version 2009). *The phonetic bases of speaker recognition*. Cambridge: Cambridge University Press.
- Nolan, Francis (2007). Voice quality and forensic speaker identification. *Govor* XXIV, 2, 111–128.
- Nolan, Francis, Catalin Grigoras (2005). A case for formant analysis in forensic speaker identification. *Speech, Language and the Law* 12, 2, 143–173.
- Rodman, Robert, David F. McAllister, Donald L. Bitzer, Luis F. Cepeda, Pamela Abbitt (2002). Forensic speaker identification based on spectral moments. *Forensic Linguistics* 9, 1, 22–43.
- Rose, Phil (2002). *Forensic Speaker Identification*. London, New York: Taylor and Francis.
- Varošanec-Škarić, Gordana (2008). Verifikacija govornika u forenzičnoj fonetici. *Govor – časopis za fonetiku* XXV, 1, 31–44.
- Varošanec-Škarić, Gordana, Jordan Bičanić (2007). A comparison of indices of difference and similarity based on voices in real forensic case and in controlled conditions. *Proceedings*. Dostupno na: www.icphs20007.de, 16th International Congress of Phonetic Sciences (Eds. Jürgen Trouvain i William J. Barry), pp 2085–2088. Datum pristupa: 3. prosinca 2013.

- Varošaneć-Škarić, Gordana, Gabrijela Kišiček (2012). Fonetsko forenzičko prepoznavanje i lingvistička analiza govornika. *Suvremena lingvistika* 73, 89–108.
- Wolfram, Walt, Ralph W. Fashold (1997). Field methods in the study of social dialects. U: Coupland N., Jaworski A. (ur.) *A Sociolinguistics. A Reader and a Coursebook*. London: Macmillan, 89–115.

Similarity and dissimilarity indices for the speakers of the Croatian language in filtered and non-filtered conditions

The main goal of this research was to compare the acoustic-statistical measurements of similarity index (R) and dissimilarity index (SDDD) on the basis of long term average spectra (LTASS). The collected speech samples consisted of both spontaneous speech and reading of a standardized text which were recorded in controlled conditions. The recorded speakers (N = 86) originated from 8 larger and dialectically different Croatian towns. The cities chosen for this research represented Sthokavian dialect (Zadar, Šibenik, Dubrovnik), Chakavian dialect (Rijeka, Pula), Kaikavian dialect (Zagreb, Čakovec) and Chakavian-Sthokavian variety (Split). Speech samples (duration 60 s) were edited for each speaker in both conditions (filtered and non-filtered speech) and compared on the basis of LTASS – non filtered conditions (0 – 10 kHz) and filtered (0.8 – 4 kHz). Using index R and index SDDD intraspeaker variations and interspeaker variations were compared respectively for male and female speakers. Results of intraspeaker variations showed that average values of similarity index (R) in non-filtered conditions were between 0.94 for male speakers in reading texts to 0.98 for female speakers in reading text and spontaneous speech. Results of interspeaker variations showed lower values of index R in the non-filtered conditions: from 0.86 in spontaneous speech to 0.94 in reading text for female speakers. Average values of R in filtered conditions for intraspeaker variations were between 0.83 for both female and male speakers in spontaneous speech to 0.95 in reading texts. Average values of R index in filtered conditions for interspeaker variations were significantly lower; from 0.57 for male spontaneous speech to 0.9 for female reading texts. Average values of index SDDD in non-filtered conditions for intraspeaker variations were generally lower – from 2.27 for female speakers to 3 for male speakers in reading. SDDD index showed higher values in non-filtered conditions for interspeaker variations; from 4.75 in female reading speech and male spontaneous speech to 5.12 for male reading speech. In filtered conditions intraspeaker variations resulted with SDDD index between 2.14 for male reading speech to 3.01 for female spontaneous speech. As expected, results in filtered conditions for interspeaker variations showed higher values of SDDD index, from 3.06 for female to 4.71 for male reading speech. The differences between similarity index (R) in intraspeaker variations were statistically significant for female speakers ($p < 0.0001$) and for male speakers ($p < 0.05$) in both spontaneous speech and reading. Results of interspeaker variations showed statistically significant differences in similarity index (R) for male speakers ($p < 0.0001$ in reading and $p < 0.0001$ in spontaneous speech) and female speakers ($p < 0.0001$ in reading and $p < 0.0001$ in spontaneous speech) and statistically significant dissimilarity index (SDDD) differences for male speakers ($p < 0.0001$ in reading and $p < 0.0001$ in spontaneous speech) and female speakers ($p < 0.0001$ in reading and $p < 0.0001$ in spontaneous speech). Overall results of this research show that acoustic-statistical measurement of similarity and dissimilarity indices are a useful method in speaker recognition in forensic phonetic expertise. Further on, results show that speaking conditions should not be neglected in forensic phonetic cases.

Key words: speaker identification, similarity index (R), dissimilarity index (SDDD), long term average spectra (LTASS), filtered speech, non-filtered speech, spontaneous speech, reading speech

Ključne riječi: prepoznavanje govornika, indeks sličnosti (R), indeks različitosti (SDDD), dugotrajni prosječni spektar govora (LTASS), filtrirani govor, nefiltrirani govor, spontani govor, čitaći govor