

**Kristina Tomić**

Filozofski fakultet u Nišu

## **MEĐUJEZIČNA FORENZIČKA KOMPARACIJA GLASOVA – DUGOROČNE FREKVENCIJE FORMANATA**

**Sažetak:** Međujezična forenzička komparacija glasova (FKG) je jedan od velikih izazova u oblasti forenzičke fonetike. Cilj ovog istraživanja je da se utvrdi stabilnost dugoročnih frekvencija formanata (DFF) kao parametra u FKG kada je sporni uzorak na stranom a nesporni na maternjem jeziku. Deset izvornih govornica srpskog snimljeno je u spontanom govoru na srpskom i engleskom jeziku. Rezultati su pokazali da postoje značajne razlike u DFF između govornica kako u govoru na maternjem tako i u govoru na stranom jeziku. Dugoročni F4 i F3 se ističu kao dobri diskriminanti jer je za njih interspikerska varijacija značajno veća od intraspikerske varijacije u oba jezika.

**Ključne reči:** forenzička fonetika, forenzička akustika, forenzička komparacija glasova, forenzička identifikacija govornika, dugoročne frekvencije formanata

### **1. Uvod i cilj**

Forenzička komparacija glasova ili forenzička komparacija govornika (FKG)<sup>1</sup> deo je oblasti forenzičke fonetike i podrazumeva analizu poznatog (nespornog) i nepoznatog (spornog) uzorka glasa radi utvrđivanja da li ih je izgovorila ista osoba (Nolan 2007; Morrison 2009; French 2017). Uzorak glasa se kvantifikuje uz pomoć različitih auditornih i/ili

<sup>1</sup> U ranijoj literaturi se umesto ova dva termina sreće termin *forenzička identifikacija govornika* (eng. forensic speaker identification) kako bi se označio isti proces (Rose 2002, Kašić i Đorđević 2009). Međutim stručnjaci ukazuju na to da je upotreba ovog termina neprikladna iz razloga što forenzičari lingviste zapravo ne vrše proveru identiteta govornika, već samo upoređuju uzorke glasova i daju informacije sudu ili nadležnim organima u svrhe identifikacije (v. Nolan 2007; Morrison 2009). Dok neki autori koriste termin *forenzička komparacija glasova* (eng. forensic voice comparison), a drugi forenzička komparacija govornika (eng. forensic speaker comparison), u ovom radu nećemo praviti razliku između ova dva termina i koristićemo istu skraćenicu FKG.

akustičkih obeležja tzv. forenzičkih parametara (Rose 2002). Validan zaključak o sličnosti glasova može se doneti tek sa odgovarajućom kombinacijom obeležja, a nikako samo na osnovu jednog parametra (Rose 2002; Nolan 2007). Neki od parametara koji se najčešće analiziraju u FKG su kvalitet glasa, kvalitet artikulacije, specifičnost artikulacije pojedinačnih govornih segmenata, uključujući vokale i konsonante, visina osnovnog tona, prozodijska obeležja govora, nejezička, idiosinkretička obeležja i dr. (Kašić i Đorđević 2009: 7-9; French 2017: 4-5). Izbor forenzičkih parametara u velikoj meri zavisi od konteksta i od samog slučaja s obzirom da naš glas značajno varira pod različitim okolnostima, kao što su, na primer, društveni kontekst, emotivno stanje govornika, ali takođe i uređaji putem kojih je glas snimljen (Nolan 1997; 2007; Rose 2002; Hollien 2012; Kašić i Đorđević 2009, Jessen 2010).

Do pre desetak godina, forenzička komparacija glasova kada su uzorci koji se porede na različitim jezicima nije smatrana metodološki ispravnom, prevashodno jer se jezici mogu razlikovati u potencijalno važnim forenzičkim parametrima (Rose 2002). Međutim, neke forenzičke laboratorije zabeležile su da rade na ovakvim slučajevima (Künzel 2013; Milne et al. 2019). Dakle, s obzirom da se u praksi već radi međujezična komparacija glasova, jasno je da postoji potreba za istraživanjima koja će utvrditi koji su to jezički parametri koji se mogu upoređivati pod ovim okolnostima.

S obzirom da fonetski parametri na koje se stručnjaci najčešće oslanjaju pri komparaciji (npr. kvalitet izgovora pojedinačnih segmenata, intonacijski opseg ili čak tip artikulacije) mogu biti nekomparabilni u dva različita jezika (Laver 1994; Gordon and Ladefoged 2001; Rose 2002), Koster i saradnici (Köster et al. 2007: 1845) ukazuju na to da bi istraživanja u ovoj oblasti trebalo da se bave ispitivanjem parajezičkih i nejezičkih obeležja, kao na primer kvalitet glasa (v. Heeren et al. 2014; Meuwly et al. 2015; Krebs and Braun 2015), tempo govora (v. Tomić 2017) itd. Za razliku od frekvencija formantata izmerenih u pojedinačnim vokalima, koje su pre svega nosioci jezičkih informacija o kvalitetu vokala, smatra se da su dugoročne frekvencije formantata, izmerene u svim vokalima u nekom dužem uzorku, uslovljene anatomijom našeg govornog trakta, te se vezuju za kvalitet glasa pojedinca (Nolan 2007).

S tim na umu, cilj našeg istraživanja je da se utvrdi da li i u kojoj meri dugoročne frekvencije formantata (DFF1, DFF2, DFF3, DFF4) ostaju iste za pojedinačnog govornika kada govori na maternjem, srpskom, i na stranom, engleskom, jeziku.

## 2. Teorijski uvod

### 2.1. Kvalitet glasa i dugoročne frekvencije formanata.

Kvalitet glasa, u širem smislu, obuhvata larinksne i supralarinksne karakteristike ljudskog govora koje čine naš glas jedinstvenim i prepoznatljivim u odnosu na druge (Laver 1980). Larinksne karakteristike glasa odnose se na tip glasa, na primer, da li je u pitanju tipičan muški tj. ženski glas ili je on zadihan, hrapav, škripav, dok se supralarinksne karakteristike glasa odnose se na poziciju artikulatora, tj. na poziciju grkljana, usana, zuba, jezika i drugih delova vokalnog trakta dok govorimo (Laver 1980; 1991). Dakle, na primer, kada neko govori sa izbačenim usnama, kažemo da je njegov govor labijalizovan, ako je zadnje nepce spušteno i vazduh delimično struji kroz nos, onda je u pitanju nazalizovan govor. Pojedinačne supralarinksne karakteristike ogledaju se u promeni frekventnog opsega zvuka, tako da promene supralarinksnih karakteristika utiču na vrednosti formanata (Laver 1991).

Dok su prva dva formanta, F1 i F2, prevashodno odgovorna za kvalitet vokala (Ladefoged and Johnson 2010), možemo reći da treći i četvrti formant, F3 i F4, nisu direktno nosioci jezičkih obeležja. Četvrti formant je u sprezi sa veličinom vokalnog trakta, tj. zavisi od veličine glave govornika (Ladefoged and Johnson 2010). Generalno posmatrano, vrednosti formanata su relativne u odnosu na oblik i dužinu vokalnog trakta (Ladefoged 2001), tako da frekventne vrednosti viših formanata mogu značajno da se razlikuju od govornika do govornika, na primer, govornici sa većom glavom u proseku imaju niže vrednosti. Na kraju, iako ne možemo reći da su vrednosti formanata jedinstvene za svakog govornika, one nam svakako mogu dati sliku o nečijem kvalitetu glasa (Ladefoged and Johnson 2010).

Dugoročna frekvencija formanata (DFF) je globalna reprezentacija vrednosti formanata u svim vokalima u dužem uzorku glasa (Nolan and Grigoras 2005). Za potrebe sociolingvističkih i dijalektoloških istraživanja, vrednosti formanata se uzimaju u centralnom delu segmentisanih vokala, jer se smatra da su upravo ti delovi nosioci jezičkih informacija. Međutim, dugoročne frekvencije formanata ne zavise puno od jezičkih informacija, tj. od toga šta se govori (Nolan and Grigoras 2005; Jessen 2010), te se zbog toga smatraju korisnim parametrom u forenzičkoj fonetici (Nolan and Grigoras 2005; Gold 2014). Nolan (2007: 111) objašnjava da analiza glasa uz pomoć DFF može da nam pruži informacije o suprala-

rinksnim karakteristikama glasa. Merenje dugoročnih frekvencija formantata je takođe pogodno za forenzičku praksu jer ne iziskuje segmentaciju vokala iz vezanog govora, koja može biti dugotrajan i naporan proces.

## **2.2. Prethodna istraživanja dugoročnih frekvencija formantata.**

Nolan i Grigoras (2005: 143) merili su DFF1 i DFF2 kako bi uporedili anonimne snimke telefonskih razgovora sa glasom osumnjičenog i dokazali da osumnjičeni ima vidno viši DFF2 nego glas na snimcima. U svom kasnijem radu, na primeru istog slučaja, Nolan (2007) pojašnjava da su više frekvencije za F2 i F3 kod osumnjičenog pokazatelj da je njegov glas zvonkiji od glasa počinioca, što je se moglo percipirati i auditornom analizom (Nolan 2007: 122). Mus (2010) je analizirala DFF1, DFF2 i DFF3 u čitanom tekstu i spontanom govoru kod 71 govornika nemačkog jezika u snimcima razgovora preko mobilnog telefona. Po njoj, dugoročne vrednosti trećeg formanta su najkorisnije u FKG zbog toga što imaju najnižu intraspikersku varijaciju (relativno mala SD za individualne govornike) (Moos 2010: 19-20), pritom se najmanje razlikuju u čitanom tekstu i spontanom govoru za najveći broj govornika (Moos 2010: 15). Goldova i saradnici (2013) istraživali su dugoročne frekvencije prva četiri formanta kod 100 izvornih govornika engleskog uz pomoć Bajesove formule verovatnoće. Prema rezultatima tog istraživanja, DFF3 se pokazao kao najbolji diskriminant sa najmanjim stepenom greške, a najbolji rezultati su dobijeni kombinacijom sva četiri dugoročna formanta (Gold et al. 2013).

Jesen i Beker (2010) pokazali su da se nemački, ruski i albanski jezik ne razlikuju značajno po pitanju dugoročnog vokalnog prostora koji zauzimaju. Oni zaključuju da ako su uzorci snimljeni pod istim okolnostima, ne bi trebalo da postoje prepreke za međujezičnu komparaciju. Hiren i saradnici (2014) analizirali su DFF2 i DFF3 kod 12 muškaraca koji govore holandski i turski jezik u snimcima telefonskih razgovora koji su nastali tokom policijske istrage. Prema njihovim rezultatima, prosečne vrednosti u okviru govornika se nisu značajno razlikovale u dva jezika. Oni takođe pokazuju da je razlika unutar jednog govornika a između jezika manja od razlike između govornika a u okviru jednog jezika, pri čemu zaključuju da su dugoročne vrednosti formantata komparabilne u različitim jezicima (Heeren et al. 2014). Mjuli i saradnici testirali su DFF2 i DFF3 za jednog bilingvalnog govornika holandskog i turskog jezika u okviru teorijskog okvira verovatnoće koristeći LR wizard, softver koji je razvijen na Holandskom forenzičkom institutu (Meuwly et al. 2015). Rezultati su dali

potvrdu identiteta kada su poređeni uzorci na istom jeziku, bilo da je to holandski ili turski, međutim, kada su upoređivani uzorci na različitim jezicima, rezultati nisu potvrdili identitet govornika (Meuwly et al. 2015). Oni zaključuju da DFF2 i DFF3 ipak zavise od jezika kojim se govori i da se varijacija u parametrima povećava kada se doda još jedan jezik u jedinačinu, te se zbog toga smanjuje snaga dokaza. Krebs i Braun su snimili 16 bilingvalnih govornika francuskog i nemačkog u spontanom govoru i tokom čitanja i to pod studijskim okolnostima i u razgovoru preko mobilnog telefona (Krebs and Braun 2015). Oni su primetili male ali značajne razlike u DFF u dva jezika kao i jaku korelaciju u DFF između jezika. Oni takođe potvrđuju da DFF3 pokazuje veću interspikersku varijaciju od intraspikerske i smatraju da je ovaj parametar svojstven govornicima (Krebs and Braun 2015).

### 2.3. O srpskom i engleskom vokalnom sistemu

Srpski jezik obuhvata pet vokala, visoko prednje /i/, visoko zadnje /u/, centralno prednje /e/, centralno zadnje /o/ i nisko srednje /a/ (Stanojčić i Popović 1989: 27–28; Simić i Ostojić 1996: 178–179; Subotić et al. 2012: 44). U nekim dijalektima srpskog, i u standardnoj ortoepskoj normi, pod uticajem tonskog akcenta, vokali u dugim i kratkim slogovima mogu se razlikovati po dužini (Ivić i Lehiste 2002; Sredojević 2017). Svi vokali u srpskom jeziku su monoftonzi, ali se diftongizacija javlja kao dijalekatsko obeležje u nekim vojvođanskim govorima kod dugih vokala (Subotić et al. 2012: 45).

U standardnom britanskom govoru postoji dvanaest monoftonga, dok ih u standardnom američkom ima jedanaest, mada njihov izgovor može biti drugačiji u različitim delovima zemlje (Kreidler 2004: 45; Ladefoged 2001: 27–28; Ladefoged and Johnson, 2010: 39). U standardnom britanskom govoru, vokali prednjeg reda su visoko /i:/, /ɪ/, srednje /e/ i nisko ka srednjem /æ/. U srednjem redu su srednje /ɜ:/ i /ə/, koje se pojavljuje samo u nenaglašenim slogovima, kao i nisko ka srednjem /ʌ/, dok se u zadnjem redu nalaze visoko /u:/, /ʊ/, visoko ka srednjem /ɔ:/, nisko ka srednjem /ɒ/ i nisko /ɑ:/ (Roach 1991: 14–22; Ladefoged 2001: 27–28; Ladefoged and Johnson 2010: 39). Standardni američki govor razlikuje se po tome što su vokali praćeni glasom /r/ „obojeni” u izgovoru. Centralni vokal /ɜ:/ se u nekoj literaturi transkribuje kao /ɝ/ kako bi se naznačila njegova potpuna „obojenost” glasom /r/. Najveći broj govornika ovog varijeteta ne pravi distinkciju između vokala /ɒ/ i /ɑ:/, već ih, u zavisnosti od područja, izgo-

varaju kao /ɑ:/ ili /ɔ:/ (Roach 1991: 240; Kreidler 2004: 55; Ladefoged and Johnson, 2010: 41–42). U engleskom takođe postoji šest diftonga /aɪ/, /aʊ/, /eɪ/, /oʊ/, /ɔɪ/ i /ju/ (Ladefoged and Johnson 2010: 92-93).

U okviru metodike nastave izgovora engleskog kao stranog jezika nastao je određeni broj istraživanja koja se bave usvajanjem engleskih vokala od strane izvornih govornika srpskog jezika. Rezultati ovih istraživanja uglavnom pokazuju da postoji određeno preklapanje u vokalnom prostoru pri govoru na maternjem i stranom jeziku ali da govornici ipak ne asimiluju sve vokale stranog jezika u vokalske kategorije koje postoje u maternjem (v. Paunović 2011; Marković 2012; Marković i Jakovljević 2016; Bjelaković 2018; Tomić i Milenković 2019). Međutim, nije poznato kako se te razlike ogledaju u dugoročnoj distribuciji formanata.

### 3. Metodologija i procedura

#### 3.1. Ispitanici i korpus

Korpus za ovo istraživanje čine snimci spontanog govora 10 izvornih govornica srpskog jezika iz Niša, koje su u periodu snimanja imale između 20 i 23 godine<sup>2</sup> (prosek 21,5; SD 1,35). Sve govornice su bile studenti u oblasti humanističkih ili društvenih nauka. Počele su da uče engleski kao strani jezik između 7 i 11 godina starosti i slušale su ovaj predmet na svim nivoima obrazovanja (prosečan period učenja engleskog 12,4 godine, SD 1,51).

Govornice su snimljene na srpskom i engleskom jeziku u vidu intervjua koji je održan po ugledu na usmeni deo ispita IELTS (Cullen et al. 2014). Pitanja koja su im postavljana su se ticala prvo svakodnevnih, poznatih tema, a zatim postajala sve apstraktnija. Ovakav tip intervjua dao je priliku svakoj govornici da se služi engleskim jezikom u onoj meri u kojoj su njene mogućnosti. Snimanje je obavljeno u zvučno izolovanoj kabini<sup>3</sup> uz pomoć besplatnog softvera za obradu zvuka Audacity i kardiodnog kondenzatorskog mikrofona sa pop filterom, *Rode NT-USB*.

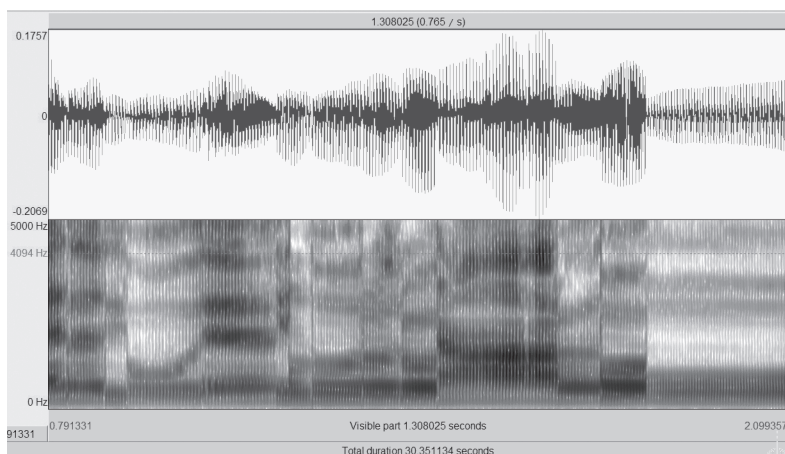
---

<sup>2</sup> Govornice u ovom istraživanju rođene su između 1994. i 1997. godine.

<sup>3</sup> Za potrebe eksperimenta korišćena je zvučno izolovana kabina XXX, kojima se autor/i ovom prilikom zahvaljuju na saradnji.

### 3.2. Akustička i statistička analiza

Koristeći spektrogramski prikaz u programu *Praat* (Boersma and Weenink 2018), modifikovali smo svaki snimak tako da sadrži samo vokale sa jasno vidljivom formantskom strukturom (v. Jessen and Becker 2010) kao na Slici 1. Nakon obrade, dobili smo uzorke u prosečnom trajanju od 46,8 s (SD 8,24) za govor na srpskom i 46,2 s (SD 8,33) za govor na engleskom jeziku. Frekvencije formantata su izmerene na 10 milisekundi uz pomoć Burgovog algoritma baziranog na linearnoj predikciji, zatim smo svakih sto mera udružili u takozvane „pakete” i izračunali srednju vrednost svakog „paketa”, te je u nastavku ta srednja vrednost posmatrana kao jedno merenje (v. Gold et al. 2013). Na ovaj način, dobili smo vrednosti takozvanih „lokalnih“ dugoročnih frekvencija formantata.



Slika 1: Deo snimka koji je obrađen da sadrži samo vokale sa jasno vidljivom formantskom strukturom

Statistička analiza podataka urađena je uz pomoć besplatnog softvera *RStudio* (RStudio Team 2016). Jednofaktorskom analizom varijanse testirali smo intraspikersku i interspikersku varijabilnost DFF-a u govoru na srpskom, a zatim i u govoru na engleskom jeziku. Za upoređivanje srednjih vrednosti DFF u dva jezika koristili smo Velčov t-test (Welch 1947), i to za svaku govornicu pojedinačno a zatim za sve govornice zajedno. Ovaj statistički test je odabran zbog toga što je pogodan za skupove podataka koji se razlikuju po obimu i imaju nejednaku varijansu (Ruxton 2006; Derrick et al. 2016).

## 4. Rezultati istraživanja

### 4.1. Interspikerska i intraspikerska varijacija

Tabela 2 sadrži prosečne vrednosti DFF u srpskom jeziku, sa standardnom devijacijom i koeficijentom varijacije za svaku govornicu kao i za sve govornice zajedno, dok Tabela 3 prikazuje jednofaktorsku analizu varijanse.

Tabela 2: Dugoročne frekvencije formanta (DFF) za govor na srpskom, prosečna vrednost (PV) u Hercima (Hz), standardna devijacija (SD) i koeficijent varijacije (KV) u procentima (%)

Gov.	DFF1			DFF2			DFF3			DFF4		
	PV	SD	KV	PV	SD	KV	PV	SD	KV	PV	SD	KV
G1	604,0	48,59	8,04	1652	188,0	11,38	2737	96,44	3,52	3907	92,80	2,38
G2	684,4	59,87	8,75	1763	148,8	8,44	2971	77,77	2,62	3977	76,54	1,92
G3	565,0	39,54	7,00	1606	138,8	8,64	2706	59,63	2,20	3851	53,00	1,38
G4	602,7	56,31	9,34	1592	172,7	10,85	2887	48,11	1,67	3956	68,65	1,74
G5	620,8	57,90	9,33	1658	163,6	9,87	2922	66,78	2,29	3868	77,12	1,99
G6	592,3	50,11	8,46	1621	171,4	10,57	2737	79,84	2,92	3811	71,32	1,87
G7	617,1	58,97	9,56	1586	184,9	11,66	2749	69,19	2,52	3784	60,71	1,60
G8	653,1	48,52	7,43	1516	123,0	8,11	2698	63,38	2,35	3625	45,35	1,25
G9	609,0	50,19	8,24	1567	169,9	10,84	2643	58,04	2,20	3812	75,05	1,97
G10	624,4	62,83	10,06	1664	170,7	10,26	2856	88,21	3,09	4031	93,57	2,32
pro-sek	617,28	53,28	8,62	1623	163,18	10,06	2791	70,74	2,54	3862	71,41	1,84
st. dev	32,77	7,08	0,97	67,1	20,45	1,26	109,7	14,78	0,53	115	15,5	0,36
KV	5,31	13,28	11,3	4,14	12,53	12,53	3,93	20,9	20,8	2,99	21,7	19,8

U Tabeli 2 možemo da vidimo da je u proseku koeficijent varijacije za pojedinačne govornice najniži za četvrti formant (1,84%), što ga čini izuzetno dobrim parametrom u komparaciji glasova. Nešto veći koeficijent zapaža se za DFF3 (2,54%), dok DFF1 i DFF2 imaju vidno veću varijaciju u okviru govornica (8,62% i 10,06%). Kada je u pitanju interspikerska varijacija, varijacija između govornica, najveće razlike mogu se zapaziti za DFF1 (5,31%), zatim za DFF2 (4,14%), DFF3 (3,93%), dok DFF4 ima najniži koeficijent varijacije između govornica (2,99%). Međutim, uprkos niskoj interspikerskoj varijaciji, DFF3 i DFF4 se i dalje izdvajaju kao najbolji forenzički parametri zbog toga što je njihova interspikerska varijacija veća od intraspikerske.



Tabela 3: Jednofaktorska analiza varijanse (ANOVA)  
za DFF za govor na srpskom; značajno za  $p < 0,05$

Param.	Između grupa			Unutar grupa			F	p-vrednost
	SS	df	MS	SS	df	MS		
DFF1	398789	9	44309,9	1303771	458	2846,66	15,566	3,04e-22
DFF2	1698550	9	188727,7	12456972	458	27198,6	6,9389	2,10e-09
DFF3	4510926	9	501214	2423154	458	5290,7	94,734	9,9e-99
DFF4	5807072	9	645230	2411103	458	5264,4	122,564	5,3e-116

Ovo je potvrdila i jednofaktorska analiza varijanse, koja je pokazala da postoje statistički značajne razlika između govornica za svaki DFF. Dugoročni F3 i F4 se ističu zbog toga što imaju izuzetno visoke F vrednosti, a interspikerska varijacija (između grupa) je veća od intraspikerske (unutar grupa). S obzirom na to da su i u srpskom i u engleskom prevashodno prva dva formanta nosioci jezičkih informacija i vezani su za kvalitet vokala, očekivano je da DFF1 i DFF2 imaju veću intraspikersku varijabilnost. Međutim, ovo ne znači da DFF1 i DFF2 ne treba razmatrati kao parametre u forenzičkoj komparaciji glasova, naročito kada su u pitanju govornici čije vrednosti frekvencija značajno odstupaju od proseka, što ćemo videti u narednom odeljku.

Tabele 4 i 5 sadrže vrednosti DFF i rezultate statističke analize podataka za korpus na engleskom jeziku.

Tabela 4: Dugoročne frekvencije formanata (DFF) za govor na engleskom, prosečna vrednost (PV) u Hercima (Hz), standardna devijacija (SD) i koeficijent varijacije (KV) u procentima (%)

Gov.	DFF1			DFF2			DFF3			DFF4		
	PV	SD	KV	PV	SD	KV	PV	SD	KV	PV	SD	KV
G1	639,2	56,66	8,86	1745	168,3	9,64	2760	103,14	3,74	3933	112,43	2,86
G2	736,4	82,98	11,27	1818	182,8	10,06	3058	70,41	2,30	3978	72,17	1,81
G3	610,4	50,16	8,22	1711	144,7	8,46	2730	83,96	3,08	3909	82,42	2,11
G4	570,3	41,21	7,23	1530	209,3	13,68	2809	82,33	2,93	3901	107,08	2,74
G5	609,3	41,42	6,80	1525	194,8	12,77	2842	102,59	3,61	3749	104,33	2,78
G6	596,4	36,95	6,20	1662	149,8	9,01	1700	50,37	2,96	3779	103,30	2,73
G7	628,6	40,71	6,48	1580	177,2	11,22	2664	146,55	5,50	3774	97,85	2,59
G8	615,1	62,11	10,1	1566	188,7	12,05	2713	107,56	3,96	3629	70,99	1,96
G9	619,2	44,13	7,13	1625	151,4	9,32	2708	114,96	4,25	3841	150,80	3,93
G10	641,3	61,50	9,59	1683	219,4	13,04	2882	138,64	4,81	4082	153,65	3,76
pro-sek	626,6	51,78	8,19	1645	178,6	10,93	2687	100,05	3,71	3858	105,5	2,73
st. dev	43,8	14,22	1,71	96,9	25,42	1,87	364,8	29,56	0,96	130	28,52	0,7
KV	6,99	27,47	20,92	5,89	14,23	17,13	13,58	29,55	25,9	3,38	27,03	25,5

I u govoru na engleskom, kao i u govoru na srpskom jeziku, koeficijent varijacije za pojedinačne govornice najniži je za četvrti formant

(2,73%), zatim za treći (2,54%), a opet znatno viši za DFF1 i DFF2 (8,19% i 10,93%). Mada je intraspikerska varijacija za DFF3 i DFF4 izraženija u stranom jeziku ( $t(\text{DFF4}) = 5,103648$ ,  $p = 0,00064$ ,  $t(\text{DFF3}) = 3,706415$ ,  $p = 0,00487$ ), ova dva formanta i dalje imaju veću interspikersku varijaciju, kao i u govoru na maternjem jeziku. Po svojoj interspikerskoj varijaciji u govoru na engleskom, naročito se ističe DFF3, koji sa 13,58% posto ima veću varijaciju i od DFF1 i DFF2.

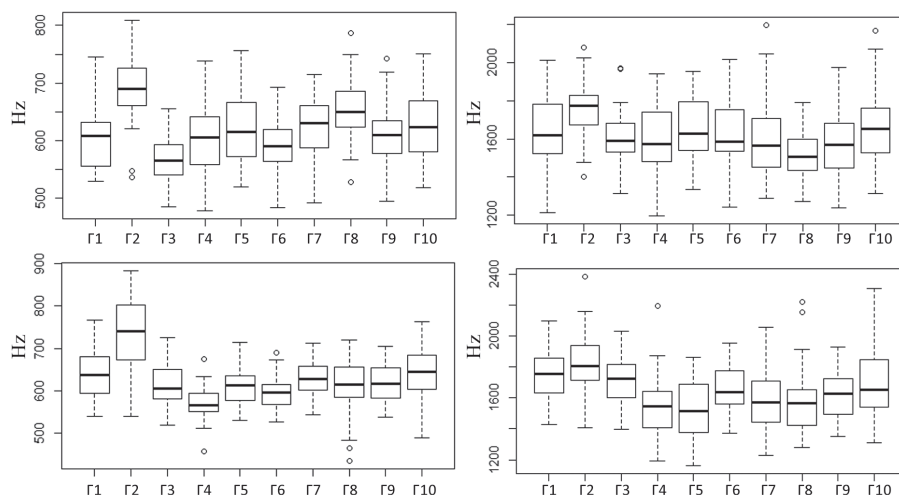
Tabela 5: Jednofaktorska analiza varijanse (ANOVA) za DFF za govor na engleskom; značajno za  $p < 0,05$

Param..	Između grupa			Unutar grupa			F	p-vrednost
	SS	df	MS	SS	df	MS		
DFF1	706509	9	78501	1313229	452	2905,37	27,019	2,25e-37
DFF2	3704175	9	411575	14440514	452	31948	12,883	2,45e-18
DFF3	5618251	9	624250	5086013	452	11277,2	55,5	2,16e-67
DFF4	7290592	9	810066	5307070	452	11741,3	68,993	3,27e-79

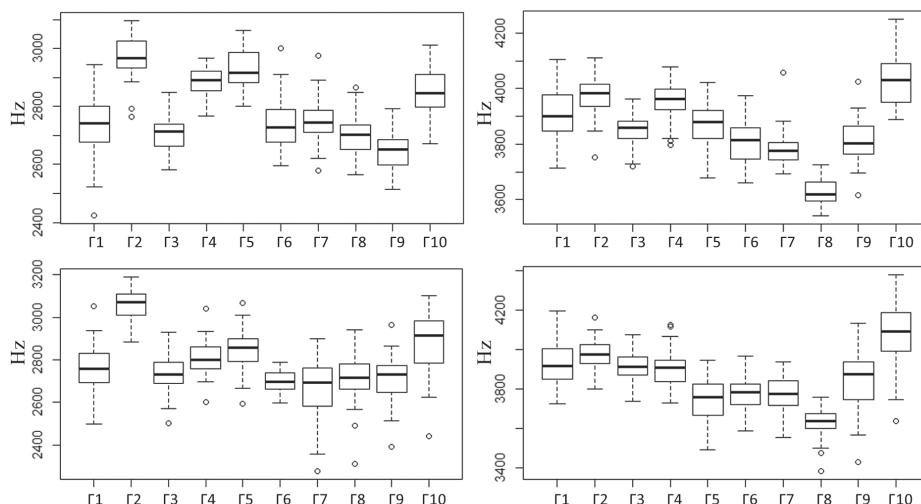
Rezultati pokazuju da, kao i u maternjem jeziku, postoji značajna razlika između govornica u DFF i kada one govore stranim jezikom. Međutim, F vrednosti za DFF3 i DFF4 su приметно niže za korpus na engleskom, s obzirom na to da postoji veća intraspikerska varijabilnost kada govornice pričaju na engleskom jeziku nego kada pričaju na srpskom.

Slike 2 i 3 su boks dijagrami na kojima možemo očitati razlike u DFF između govornica u oba jezika. Detaljnije komentare grafikona daćemo u odeljku 5.

Slika 2: Boks dijagram za DFF1 (levo) i DFF2 (desno) za govor na srpskom (gore) i govor na engleskom jeziku (dole)



Slika 3: Boks dijagram za DFF3 (levo) i DFF4 (desno) za govor na srpskom (gore) i govor na engleskom jeziku (dole)



## 4.2. Međujezična komparacija glasova

Kako bismo uporedili dugoročne frekvencije formanta u dva jezika, uradili smo Velčov t-test za svaku govornicu pojedinačno kao i za celokupne skupove podataka. Kako je prikazano u Tabeli 6, statistički test je pokazao da su DFF1 i DFF2 u govoru na engleskom u proseku viši nego u srpskom dok za DFF3 i DFF4 ne postoje značajne razlike u dva jezika.

Tabela 6: Velčov t-test za DFF u govoru na engleskom i srpskom jeziku; značajno za  $p < 0,05$

Говор.	ДДФ1		ДДФ2		ДДФ3		ДДФ4	
	t-тест	p-вред.	t-тест	p-вред.	t-тест	p-вред.	t-тест	p-вред.
Г1	3,4628	0,00078	2,7047	0,00799	1,214	0,22744	1,2735	0,20568
Г2	3,0314	0,00337	1,4421	0,15359	5,016	4,6e-06	0,1021	0,91900
Г3	5,0627	1,9e-06	3,6671	0,00041	1,6446	0,10326	4,2163	5,6e-05
Г4	-2,7122	0,00859	-1,305	0,19701	-4,6208	2,9e-05	-2,4383	0,01838
Г5	-1,0623	0,29125	-3,4002	0,00107	-4,2019	8,2e-05	-5,9225	1,0e-07
Г6	0,4547	0,6504	1,2242	0,22416	-2,7296	0,00761	-1,6647	0,10094
Г7	1,1964	0,23451	-0,1641	0,86999	-3,8992	0,0002	-0,6681	0,5058
Г8	-3,4774	0,00077	1,6003	0,11335	0,8893	0,37661	0,3321	0,7407
Г9	1,0265	0,30745	1,726	0,08782	3,3668	0,0013	1,154	0,2529
Г10	1,348	0,18102	0,47052	0,63914	1,0817	0,28261	1,9573	0,0539
све	2,9507	0,00325	2,7789	0,00557	0,2283	0,81947	0,86664	0,38638

Primitićemo da vrednosti DFF3 i DFF4 nisu iste u oba jezika za svaku pojedinačnu govornicu, kao i da vrednosti DFF1 i DFF2 nisu uvek značajno različite. U Tabeli 6, zasenčane ćelije označavaju da ne postoji razlika u prosečnim vrednostima u govoru na srpskom i engleskom za dati DFF za datu govornicu. Zapravo, tek četiri od deset govornica imaju iste DFF3 u srpskom i engleskom, dok 50% govornica ima slične DFF1. Parametar koji se pokazuje kao najpovoljniji je DFF4, s obzirom da sedam od deset govornica ima slične vrednosti u oba jezika, zatim sledi DFF2, koji ima iste vrednosti u maternjem i stranom jeziku za šest govornica.

## 5. Diskusija

Rezultati našeg istraživanja na prvi pogled deluju dvosmisleno. Naime, iako možemo da primetimo sličnosti u dugoročnim distribucijama frekvencija formanta u govoru na srpskom kao maternjem i engleskom kao stranom jeziku, ne možemo doneti neke generalne zaključke. Broj DFF koji za datu govornicu mogu biti slični u dva jezika varira od jednog (za G3, G4 i G5) pa sve do četiri (za G10), pritom, čini se da ne postoji pravilo koji DFF će uspešno identifikovati koju govornicu. Dakle, potrebno je posmatrati širu sliku.

Gledajući boks dijagrame na slikama 2 i 3, primitićemo da su relativne vrednosti dugoročnih distribucija frekvencija formanta približno iste i u srpskom i u engleskom. Na primer, na Slici 3, vidimo da je DFF3 za G2 u govoru na srpskom ispod 3000 Hz, dok je u govoru na engleskom iznad ove vrednosti ( $t = 5,016$ ,  $p = 4,6e-06$ )<sup>4</sup>, s tim što je u oba jezika, ova vrednost znatno viša u odnosu na DFF3 kod ostalih govornica. Dakle, za identifikaciju govornika kada je uzorak na stranom jeziku, nije dovoljno samo uzeti u obzir sličnosti i razlike između dva uzorka na različitim jezicima, već i njihove relativne vrednosti u referentnoj populaciji.

Ovaj koncept nije stran u forenzičkim naukama, zbog čega i postoji tendencija ka upotrebi Bajesove statistike za testiranje efikasnosti parametara (Rose 2002). Kada je u pitanju komparacija glasova sa uzorcima na istom jeziku, Bajesova formula verovatnoće se pokazala kao značajna upravo zbog toga što uzima u obzir sličnosti između uzoraka kao i njihovu tipičnost u referentnoj populaciji (Rose 2002). Sledeći korak u našem

<sup>4</sup> Rezultat Velčovog t-testa iz Tabele 6.

istraživanju međujzčne komparacije glasova je upravo testiranje dugoročnih frekvencija formanata Bajesovom formulom za verovatnoću.

Ako pogledamo rezultate kroz prizmu Lejverove teorije o kvalitetu glasa, možemo reći da govornice zadržavaju svoj kvalitet glasa kada govore stranim jezikom. Na primer, G2 u proseku ima više vrednosti formanata od ostalih govornica i kada govori na maternjem i kada govori na stranom jeziku, što nam ukazuje na to da je njen glas palatalizovan, tj zvonak, ili zvonkiji u odnosu na ostale glasove (Nolan 2007). Slično, veoma niske vrednosti za DFF4 i malo povišeni DFF1 specifičan za G8 i u srpskom i u engleskom ukazuje nam na njen škripav glas (Laukkanen et al. 2004).

## 6. Zaključak

Cilj ovog istraživanja bio je da se uporede dugoročne frekvencije formanata (DFF) kod izvornih govornica srpskog kada govore na svom maternjem i engleskom kao stranom jeziku u svrhe međujezične forenzičke komparacije glasova. Statistička analiza (ANOVA) je pokazala da u oba jezika postoji značajna razlika između govornica kod svih izmerenih formanta (DFF1, DFF2, DFF3, DFF4), pri čemu DFF3 i DFF4 imaju značajno veću interspikersku varijaciju od intraspikerske i u govoru na engleskom i u govoru na srpskom.

Sveukupno gledajući, u proseku, DFF1 i DFF2 se značajno razlikuju u govoru na srpskom i engleskom jeziku, dok DFF3 i DFF4 imaju slične vrednosti u dva korpusa. Naime, imajući u vidu činjenicu da kvalitet vokala u oba jezika pretežno zavisi od prva dva formanta, možemo reći da naši rezultati pokazuju da se vokalni prostor srpskog i engleskog kao stranog jezika značajno razlikuju. Izvorni govornici srpskog prave određena odstupanja od svog maternjeg vokalnog prostora kako bi svoj izgovor približili izgovoru stranog jezika, te se usled toga vrednosti za DFF1 i DFF2 razlikuju. Ovim smo pokazali da dugoročne frekvencije formanata, iako u nekoj meri oslikavaju supralarinksne karakteristike glasa, nisu sasvim nezavisne od jezika kojim se govori, kako je to rečeno u nekim prethodnim istraživanjima (v. Jessen and Becker 2010; Heeren et al. 2014).

Međutim, ako posmatramo vrednosti kod pojedinačnih govornica, rezultati nisu tako jasno definisani. Naime, DFF1 i DFF2 se ne razlikuju značajno u dva jezika za svaku govornicu, kao što ni DFF3 i DFF4 nisu

uvek isti. Dakle, kada se forenzičari lingviste oslanjaju na dugoročne distribucije formanta kao parametar u međujezičnoj komparaciji glasova, trebalo bi da budu naročito pažljivi i da se ne oslanjaju samo na sličnosti između uzoraka već i na njihove relativne vrednosti u referentnoj populaciji. Naime, velika je verovatnoća da će govornici koji odstupaju od proseka u govoru na maternjem jeziku na isti način odstupati od proseka i u govoru na stranom jeziku, te se, kao i za svaki drugi forenzički parametar, upravo oni mogu najpreciznije identifikovati.

Rezultati ovog istraživanja razlikuju se od prethodnih po tome što je u ranijim istraživanjima treći formant predstavljen kao najstabilniji, a samim tim i najpouzdaniji parametar (Moos 2010; Gold et al. 2013; Krebs and Braun 2015). Prema našim rezultatima, četvrti formant se pokazao kao najstabilniji zbog toga što ima slične vrednosti u dva jezika za 7 od 10 govornica i najmanji koeficijent varijanse i za govor na maternjem i za govor na stranom jeziku. U fonetskim istraživanjima komparacije glasova, četvrti formant se često ne uzima u obzor zbog toga što je uglavnom forenzički materijal u praksi snimljen putem telefona, a široko je prihvaćeno da u telefonskim razgovorima zvuk ispod 300 Hz i iznad 3500 Hz može biti odsečen i blizu ovih frekvencija mogu se javiti distorzije u zvučnom spektru (Nolan 2007: 119). Ipak, imajući u vidu razvoj tehnologije i poboljšanja kvaliteta zvuka koji se prenosi putem telefonskog signala kao i sve veću težnju ka komunikaciji putem interneta, ne treba zanemariti značaj četvrtog formanta.

Istraživanje koje smo ovde predstavili je pilot za veću studiju sa većim brojem ispitanika koja će testirati dugoročne frekvencije formanta uz pomoć Bajesove formule verovatnoće. Buduća istraživanja se takođe mogu usredsrediti na testiranje DFF u međujezičnoj komparaciji glasova snimljenih putem mobilnih telefona ili u komunikaciji preko interneta, kao i na testiranje DFF u različitim parovima jezika. Osim toga što su nosioci nejezičkih informacija i mogu da oslikavaju kvalitet glasa, dugoročne frekvencije formanta se relativno lako mere jer ne zahtevaju segmentaciju i to ih čini veoma korisnim u forenzičkoj praksi. Međutim, kao što je Nolan rekao (2007), jedan parametar je samo delić u velikoj slagalici i svaki parametar nam pomaže da jasnije sagledamo celokupnu sliku te tako preciznije identifikujemo snimljenog govornika. Zbog toga je važno naglasiti da se, u praksi, forenzičkoj komparaciji glasova ne bi smelo pristupiti samo na osnovu ovde testiranih parametara. Dugoročne frekvencije formanta mogu se koristiti isključivo u kombinaciji sa drugim parametrima, čija se efektivnost u međujezičnoj komparaciji glasova tek treba utvrditi.

## Literatura

- Baldwin, J., and French, P. (1990). *Forensic Phonetics*. London: Pinter.
- Bjelaković, A. (2018). *Vokali savremenog standardnog britanskog izgovora i njihovo usvajanje kod izvornih govornika srpskog jezika*. Doktorska disertacija, Univerzitet u Beogradu
- Boersma, P., and Weenink, D. (2018). Praat: Doing Phonetics by Computer (Version 6.0.42) [Computer program}. Preuzeto 16. avgusta 2018. sa [http://www.fon.hum.uva.nl/praat/download\\_win.html](http://www.fon.hum.uva.nl/praat/download_win.html)
- Cullen, P., French, A., and Jakeman, V. (2014). *The Official Cambridge Guide to IELTS for Academic and General Training*. Cambridge : Cambridge University Press.
- Derrick, B., Toher, D., and White, P. (2016). Why Welch's test is Type I error robust. *The Quantitative Methods for Psychology*, 12(1), 30-38. doi:10.20982/tqmp.12.1.p030
- French, P. (2017). A developmental history of forensic speaker comparison in the UK. In *English Phonetics* (pp. 271-286). Preuzeto 5. oktobra 2019. sa <http://eprints.whiterose.ac.uk/117763/>
- Gold, E. A. (2014). *Calculating likelihood ratios for forensic speaker comparisons using phonetic and linguistic parameters*. PhD Thesis. The University of York, Department of Language and Linguistic Science. Preuzeto 15. januara 2020 sa <http://etheses.whiterose.ac.uk/id/eprint/6166>
- Gold, E., French, P., and Harrison, P. (2013). Examining long-term formant distributions as a discriminant in forensic speaker comparisons under a likelihood ratio framework. *The Journal of the Acoustical Society of America*, 133, 3294. doi:10.1121/1.4800285
- Gordon, M., and Ladefoged, P. (2001). Phonation types: a cross-linguistic overview. *Journal of Phonetics*, 29, 383-406. doi:10.1006/jpho.2001.0147
- Heeren, W., van der Vloed, D., and Vermeulen, J. (2014). Exploring long-term formants in bilingual speakers. *Proceedings of the International Association for Forensic Phonetics and Acoustics conference, Zurich, Switzerland*, (pp. 39-40)
- Hollien, H. (2012). About Forensic Phonetics. *Linguistica*. 52, 27-53. Ljubljana: Tiskana. doi:10.4312/linguistica.52.1.27-53
- Ivić, P., and Lehiste, I. (2002). *O srpskohrvatskim akcentima*. (D. Petrović, ur.) Novi Sad: Izdavačka knjižarnica Zorana Stojanovića.
- Jessen, M. (2010). The forensic phonetician: Forensic speaker identification by experts. In M. Coulthard, & A. Johnson (Eds.), *The Routledge Handbook of Forensic Linguistics* (pp. 378-394). Abingdon and New York: Routledge.

Jessen, M., and Becker, T. (2010). Long-Term Formant Term Formant Distribution as a Forensic Phonetic Feature. *ASA - 2<sup>nd</sup> Pan-American / Iberian Meeting on Acoustics*. Cancún, Mexico, Nov 15-19. doi:10.1121/1.3508452

Kašić, Z., and Đorđević, J. P. (2009). Zašto je lingvistika postala forenzička veština. U D. Radovanović (Ur.), *Istraživanja u specijalnoj pedagogiji* (str. 469-482). Beograd: Univerzitet u Beogradu - Fakultet za specijalnu edukaciju i rehabilitaciju.

Köster, O., Jessen, M., Khairi, F., and Eckert, H. (2007). Auditory-perceptual identification of voice quality by expert and non-expert listeners. *Proceedings of the 16<sup>th</sup> International Congress of Phonetic Sciences*, (pp. 1845–1848). Saarbrücken.

Krebs, P., and Braun, A. (2015). Long Term Formant measurements in bilingual speakers. *Proceedings of the IAFPA Annual Conference, 8 - 10 July 2015*. Leiden.

Kreidler, C. W. (2004). *The Pronunciation of English: A Course Book* (2<sup>nd</sup> ed.). Oxford: Blackwell Publishing.

Künzel, H. (2013). Automatic speaker recognition with crosslanguage speech material. *International Journal of Speech Language and the Law*, 20(1), 21-44. doi:10.1558/ijssl.v20i1.21

Ladefoged, P. (2001). *Vowels and Consonants: An Introduction to the Sounds of Language*. Massachusetts, Oxford : Blackwell Publishers.

Ladefoged, P., and Johnson, K. (2010). *A Course in Phonetics* (6<sup>th</sup> ed.). Wadsworth: Cengage Learning.

Laukkanen, A.-M., Sundberg, J., and Björkner, E. (2004). Acoustic study of throaty voice quality. *TMH-QPSR*, 46(1), 13-24. Preuzeto 1. februara 2019. sa <http://www.speech.kth.se/qpsr/>

Laver, J. (1980). *The Phonetic Description of Voice Quality*. Cambridge: Cambridge University Press.

Laver, J. (1991). *The gift of speech: papers in the analysis of speech and voice*. Edinburgh: Edinburgh University Press.

Laver, J. (1994). *Principles of Phonetics*. Cambridge: Cambridge University Press.

Marković, M. (2012). *Usporedna proučavanja vokala engleskog i srpskog jezika: između univerzalnog i specifičnog*. Novi Sad: Filozofski fakultet u Novom Sadu.

Marković, M., and Jakovljević, B. (2016). Phonetic cue interpretation in the acquisition of a non-native vocalic contrast. *Godišnjak Filozofskog fakulteta u Novom Sadu*, 41(1), 215-227.

Meuwly, D., Heeren, W., and Bolck, A. (2015). Exploring the strength of evidence of long-term formants in bilingual speakers. *Proceedings of*



*Annual Conference of the International Association for Forensic Phonetics and Acoustics*, (pp. 75-76).

Milne, P., Cavanagh, C., van der Vloed, D., and Dellwo, V. (2019). A survey of voice-related cases in three forensic speech laboratories. *A paper presented at the 28th Annual Conference of the International Association for Forensic Phonetics and Acoustics*. Istanbul, 14th -17th July 2019.

Moos, A. (2010). Long-term formant distribution as a measure of speaker characteristics in read and spontaneous speech. *The Phonetician*, 101, 7-24. Preuzeto 15. januara 2019. sa [http://www.isphs.org/Phonetician/Phonetician\\_101.pdf](http://www.isphs.org/Phonetician/Phonetician_101.pdf)

Morrison, G. S. (2009). Forensic voice comparison and the paradigm shift. *Science and Justice*, 49, 298–308. doi: 10.1016/j.scijus.2009.09.002

Nolan, F. (1997). Speaker Recognition and Forensic Phonetics. In W. J. Hardcastle, & J. Laver (Eds.), *The Handbook of Phonetic Sciences* (pp. 744-767). Oxford: Blackwell.

Nolan, F. (2007). Voice Quality and Forensic Speaker Identification. *Govor*, 24(2), 111-128. Preuzeto 15. januara 2019. sa <https://hrcak.srce.hr/173611>

Nolan, F., and Grigoras, C. (2005). A case for formant analysis in forensic speaker identification. *International Journal of Speech, Language and the Law*, 12(2), 143-173. doi:10.1558/sll.2005.12.2.143

Paunović, T. (2011). Sounds Serbian? Acoustic properties of Serbian EFL students' speech. U E. Kitis, N. Lavidas, N. Topintzi, & T. Tsangalidis (Eds.), *Selected Papers from the 19<sup>th</sup> International Symposium on Theoretical and Applied Linguistics (ISTAL19)* (pp. 357-369). Thessaloniki: Aristotle University of Thessaloniky, School of English, Department of Theoretical & Applied Linguistics.

Roach, P. (1991). *English Phonetics and Phonology: A Practical Course* (2<sup>nd</sup> ed.). Cambridge: Cambridge University Press.

Rose, P. (2002). *Forensic Speaker Identification*. London and New York: Taylor & Francis.

RStudio Team. (2016). RStudio: Integrated Development for R. Boston, MA. Preuzeto 15. novembra 2016. sa <http://www.rstudio.com>

Ruxton, G. D. (2006). The unequal variance t-test is an underused alternative to Student's t-test and the Mann–Whitney U test. *Behavioral Ecology*, 17(4), 688-690. doi:10.1093/beheco/ark016

Simić, R., and Ostojić, B. (1996). *Osnovi fonologije srpskog književnog jezika*. Beograd: Univerzitet u Beogradu.

Sredojević, D. (2017). *Fonetsko-fonološki opis akcenata u standardnom srpskom jeziku: od specifičnog ka opštem*. Novi Sad: Sajnos.

Stanojčić, Ž., and Popović, L. (1986). *Gramatika srpskoga jezika za gimnazije i srednje škole*. Beograd: Zavod za udžbenike.

Subotić, L., Sredojević, D., and Bjelaković, I. (2012). *Fonetika i fonologija: ortoepska i ortografska norma standardnog srpskog jezika*. Novi Sad: Filozofski fakultet Novi Sad.

Tomić, K. (2017). Temporal Parameters of Spontaneous Speech in Forensic Speaker Identification in Case of Language Mismatch: Serbian as L1 and English as L2. *Comparative Legilinguistics*, 32, 117-143. doi:10.14746/cl.2017.32.5

Tomić, K., and Milenković, K. (2019). Forenzičko profilisanje govornika iz uzorka na engleskom kao stranom jeziku - analiza kvaliteta vokala. *Zbornik Matice srpske za filologiju i lingvistiku*, 62(1), 151-170

Welch, B. L. (1947). The generalization of Student's' problem when several different population variances are involved. *Biometrika*, 34(1), 28-35. doi:10.2307/2332510

**Kristina Tomić**

## **CROSS-LANGUAGE FORENSIC VOICE COMPARISON - LONG-TERM FORMANT FREQUENCIES**

**Summary:** Forensic Voice Comparison (FVC) has helped solve numerous forensic cases over the past fifty years; however, comparing voice samples in different languages remains one of the challenges of the field. The goal of the current research is to test the robustness of long-term formant frequencies (LTF) as a parameter in FVC with samples in different languages. Ten female native speakers of Serbian were recorded while speaking spontaneously in Serbian and English and LTFs of the first four formants were extracted. The results confirm that there are significant differences in LTFs between speakers both when talking in their mother tongue and in the foreign language. Long-term F4 and long-term F3 emerge as particularly important discriminants since between-speaker variation for these two formants is higher than within-speaker variation in both languages. In addition, the research has confirmed that speakers who diverge from the mean the most in their mother tongue, diverge from the mean in the same way in the foreign language. Therefore, when drawing conclusions in FVC, experts should not only take into account the similarity between samples across languages but also their relative values in the referent population.

**Key words:** forensic phonetics, forensic acoustics, forensic voice comparison, forensic speaker identification, long-term formant frequencies