

Gordana Varošaneć-Škarić i Gabrijela Kišiček

gvarosan@ffzg.hr; gkisicek@ffzg.hr

Filozofski fakultet, Zagreb

Fonetsko forenzičko prepoznavanje i lingvistička analiza govornika

S dviju snimaka u ukupnom trajanju duljem od devet minuta na temelju službenog prisluškivanja razgovora preko GSM uređaja trebalo je provesti fonetsko prepoznavanje jedne muške osobe. Učinjena je fonetska transkripcija cjelokupnoga govora s obje snimke radi utvrđivanja regionalnog izgovora, dijalekta i idiolekta te montaža zvuka radi utvrđivanja postotka sličnosti – različitosti parova glasova u postupku slušnoga forenzičkog protokola (AP – SPID) te radi akustičke analize. Akustički parametri bile su prosječne vrijednosti F_0 (Hz), formantnih frekvencija F_1 (samo za vokal /a/), F_2 , F_3 za svih pet vokala, trajanja dugih naglašenih vokala (ms) (u programu Praat). Pomoćni postupci bili su dugotrajni prosječni spektar govora (LTASS u programu AS), na temelju kojih su izračunani indeksi sličnosti (R) i različitosti (SDDD) među parovima glasova. Zbog kraćeg trajanja snimki bila je važna lingvistička analiza radi utvrđivanja fonoloških, sintaktičkih i leksičkih razlika među govornicima.

Pomoću AP – SPID protokola utvrđeno je 96%-tno prepoznavanje tražene osobe u paru s obje snimke, što prema fonetskom forenzičkom postupku znači vrlo veliku vjerojatnost prepoznavanja. Prepoznati govornik pripada štakavskom ikavskom bosanskom govoru. Fonetski je opis glasa lagano povišen glas, glasan, uskog raspona čeljusti, blago dentaliziran, modalne fonacijske vrste. Prosječna vrijednost F_0 bila je 197 – 201 Hz, što su povišene vrijednosti, s rasponom od 76 do 317 Hz, ovisno o glasnoći. F_1 vokala /a/ prepoznatoga govornika značajno se razlikuje od standardne vrijednosti ($p = 0,0001$), a značajno se razlikuje F_2 za /a/ ($p = 0,0006$), /o/ i /u/ ($p < 0,0001$). F_3 je značajno niži samo za /i/ ($p < 0,00001$) zbog zatvorenog, prednjeg, zaokruženog izgovora u naglašenim pozicijama. Trajanje naglašenih vokala dulje je od trajanja u hrvatskome standardnom izgovoru. Indeks sličnosti između spektara s obje snimke tražene osobe veći je nego između tražene osobe i sugovornika ($R = 0,89 : 0,85$), a indeks različitosti manji je (SDDD = 2,10 : 3,75). Najmanji indeks sličnosti i najveći indeks različitosti pokazao se između tražene osobe i podmetnutog glasa.

U lingvističkoj analizi najzanimljivija je bila leksička razina izbora riječi, idiosinkratične riječi razlikovale su se među govornicima. Primjerice, traženi govornik rabi riječi poput *bolan*, *jaran*, koje se uopće nisu pojavile kod sugovornika, kojemu su idiosinkratične riječi bile *pajdo*, *kontaj*. Traženi govornik izgovorio je riječ *jaran* čak 69 puta, što je 170 puta više nego u upotrebi u hrvatskom jeziku, dok se riječ *pajdo* uopće ne pojavljuje u hrvatskome čestotnom rječniku.

Može se zaključiti da je traženi govornik pomoću slušnoga fonetskog protokola prepoznat s velikom vjerojatnošću, a da zvučni postupci mogu samo poduprijeti rezultat slušne analize. Pokazalo se da je u slučaju manje količine snimljenog materijala važna sinergija fonetskoga forenzičkog postupka i lingvističke analize.

1. UVOD

Rad se temelji na stvarnome forenzičkom slučaju prepoznavanja muškoga govornika. Na zahtjev Županijskoga suda u Zagrebu trebalo je identificirati jedan muški glas na temelju snimki službenog prislušivanja razgovora preko GSM uređaja. Sa snimke 1 u trajanju od 2,16 min (CD 1) trebalo je prepoznati muški glas na snimci 2 (CD 2) za vrijeme razgovora s drugom muškom osobom u trajanju od 7,39 min. Prepoznavanje govornika temeljilo se na forenzičkoj fonetskoj i lingvističkoj analizi.

Forenzička fonetika relativno je nova grana fonetike koja počinje dvadesetih godina 20. stoljeća tehničkim razvojem snimanja zvuka, a intenzivnije se razvija posljednjih nekoliko desetljeća. Surječje forenzičke fonetike uključuje područja prepoznavanja, identifikacije govornika, profiliranje govornika (regionalna pripadnost, socioekonomske i psihološke karakteristike govornika, patologija glasa, strani naglasak), *voice line-ups*, fonetsku transkripciju, tj. određivanje sadržaja izrečenoga (u slučajevima loše snimke), određivanje autentičnosti snimke (npr. jesu li postojale manipulacije sa snimkom itd.) (Rose, 2002).

Forenzička lingvistika može se definirati kao grana primijenjene lingvistike koja se primarno bavi vezom između jezika, zakona i kriminala. Naziv *forenzička lingvistika* prvi je upotrijebio Svartvik 1968. U posljednjih 15 godina primjetan je znatan razvoj istraživanja na tome području te sve češće korištenje forenzičkih lingvista kao stručnih svjedoka u sudskim procesima, o čemu svjedoče, primjerice, standardna djela iz toga područja Coultharda i Johnson (2007) i McMenamina (2002).

1.1. Značenje pojmova prepoznavanja, identifikacije i verifikacije u forenzičkoj fonetici

Govoreći o prepoznavanju govornika, možemo ga podijeliti na *identifikaciju* i *verifikaciju* govornika. Identifikacija govornika je proces utvrđivanja govora govornika od određenog broja mogućih govornika, a verifikacija govornika je proces prihvatanja ili odbijanja tvrdnje o identitetu, što znači da se podrazumijeva naknadno snimanje osumnjičenika u postupku fonetske verifikacije, obično kod nesigurnih dokazivanja krivnje ili nevinosti. Osnovna je razlika u broju mogućih odgovora. U procesu identifikacije broj mogućih rješenja jednak je broju ukupne populacije, dok je u procesu verifikacije moguće samo dvoje – prihvatiti ili odbaciti tvrdnju o identifikaciji, neovisno o veličini populacije. Za razliku od laičkog, odnosno naivnoga prepoznavanja pomoću slušnih svjedoka, ekspertno forenzičko prepoznavanje podrazumijeva slušnu procjenu sličnosti – različitosti parova glasova govornika pomoću izvježbanih eksperata (Varošaneć-Škarić, 2008:32) prema protokolu forenzičke fonetike (Hollien, 2002). Kako je osumnjičenik iz konkretnoga stvarnog slučaja razgovarao s fonetičarem, jer je sud tako zahtijevao, ali je tijekom intervjuja odbio naknadno snimanje pod nadziranom tehničkim standardnim uvjetima, nadalje ćemo postupak identifikacije na temelju snimki službenoga prislušivanja nazivati pre-

poznavanje prema određenim šiframa koje predstavljaju različite govornike, da ne bi došlo do pogrešnog razumijevanja razlikovanja termina identifikacije u pravnoj kriminalistici i mogućim značenjima u forenzičkoj fonetici. Forenzički fonetičari u svakome forenzičkom kontekstu uglavnom daju prednost nazivu identifikacije, primjerice Hollien (2002), Rose (2002), Künzel (2010), dok drugi, koji se u eksperimentalnom smislu bave forenzičkom fonetikom, daju prednost terminu prepoznavanje (engl. *speaker recognition*), primjerice Nolan (1983, 2009). S obzirom na forenzičko surjeđe ovoga rada, priklanjamo se terminu prepoznavanje.

2. POSTUPAK

2.1. Preliminarni postupci u utvrđivanju prepoznavanja glasa govornika

Učinjena je fonetska transkripcija cjelokupnoga govora s obje snimke radi utvrđivanja regionalnoga izgovora, dijalekta i ideolekta te montaža zvuka radi akustičke analize.

Slušnom analizom opisan je govor tražene osobe, odnosno govornika sa snimke 1 i snimke 2 te su se transkribirali specifični glasnici (prvenstveno vokali) i uspoređivali sa standardnim glasnicima hrvatskoga jezika.

2.2. Slušno prepoznavanje govornika

U forenzičkoj fonetici slušna analiza eksperta – uvježbanog fonetičara vrednuje se kao najpouzdaniji postupak u prepoznavanju i verifikaciji govornika. Kombinacija slušne i akustičke analize najčešće se primjenjuje, dok je postupak automatskog prepoznavanja još u fazi usavršavanja i provjere, budući da su još moguće pogrešne pozitivne identifikacije. U potonjem smislu mogu se spomenuti novije metode automatskog prepoznavanja kao što su BATVOX (Harrison i French, 2010) te Automatic SPID na bazi višejezičnoga govornog materijala (Künzel, 2010).

U našem istraživanju postupak slušnoga prepoznavanja provodio se prema priznatom protokolu forenzičke fonetike (AP – SPID: Aural-perceptual Approach to Speaker Identification, Hollien 2002:80), koji je razvijen u Sjedinjenim Američkim Državama, Velikoj Britaniji i Njemačkoj, a koji Varošaneć-Škarić prilagođen primjenjuje u Hrvatskoj od 2003. u stvarnim forenzičkim slučajevima te u eksperimentalne svrhe. Kategorije su AP – SPID protokola redom ton, kvaliteta glasa, intenzitet, dijalekt, artikulacija, prozodija, ostalo (eventualni poremećaji glatkoće govora i ostali patološki poremećaji govora).

U postupku slušnoga prepoznavanja parove glasova govornika procjenjivala su dva ekspertna uvježbana fonetičara od kojih je jedan prvi autor rada, a redosljed i kombinaciju parova pripremio je fonetičar, drugi autor rada. Procjenjivalo se ukupno 11 parova označenih šiframa u različitim kombinacijama. Uz glas osumnjičenika i njegovoga sugovornika ubacili su se i tzv. *mućci* (podmetnuti glasovi, odnosno tragovi metaforički se u forenzičkom metaje-

ziku nazivaju mućci, engl. *foils*) poznati samo fonetičaru koji nije sudjelovao u prepoznavanju. Govornici *mućci* snimljeni su u fonetskom studiju kako izgovaraju tekst sličan onome u telefonskom razgovoru. Te su studijske snimke potom filtrirane u rasponu od 300 do 3000 Hz i montirana je buka, kojom se maskira studijsko snimanje (Varošaneć-Škarić i Bićanić, 2007). Na taj se način dobiju snimke koje su akustički poput stvarnoga telefonskog razgovora. Kombinacije su napravljene tako da su se uzela tri različita uzorka povezanoga govora traženoga govornika, zatim tri različita uzorka sugovornika sa snimke 1 i snimke 2, dva uzorka mućka 1 (pri čemu on u jednom uzorku govori isti tekst kao sugovornik te jedan drugačiji tekst, ali sličnog konteksta) i tri uzorka govora mućka 2 (jedan tekst isti kao tražena osoba, jedan kao sugovornik i jedan drugačiji, ali sličnoga konteksta). Tako je dobiveno 11 uzoraka govora raspoređenih u 11 parova. U postupku slušnoga prepoznavanja (AP – SPID) svaki se par puštao onoliko puta koliko je bilo potrebno da se procijeni svih sedam kategorija i njihovih varijabli, a najmanje sedam puta.

Za svaki par izračunani su postotci i rasponi prepoznavanja. U forenzičkoj fonetici ukupan rezultat pozitivnog prepoznavanja za sve dimenzije mora iznositi najmanje 90% podudaranja, a najmanji ukupni postotak točnosti utvrđivanja identiteta u paru mora iznositi 80%. Tek kad je ukupno prepoznavanje veće od 90% može se tvrditi da postoji vjerojatnost da se radi o istoj osobi, a prosječno prepoznavanje od 94% i više nesporna je identifikacija ako je u uzorku bio i glas osumnjičenika snimljen u verifikacijskom postupku.

2.3. Akustička analiza

Akustička analiza temeljila se na izračunavanju prosječnih vrijednosti F_0 (u Hz) na temelju govora, formantnih frekvencija (F_1 , F_2 , F_3 u Hz) pet vokala, trajanja vokala (u ms) u programu Praat (Boersma i Weenink, 2009); kao pomoćne akustičke metode usporedbi timbra korišteni su postupci dugotrajnoga prosječnog spektra govora govornika (LTASS) u programu AS (prilagodili Škarić i Stamenković 1990, prema Škarić 1993) te na temelju LTASS, akustičko-statističke metode indeksa sličnosti (R) i različitosti (SDDD) glasova govornika (Harmegnies, 1985, 1995).

Fundamentalna frekvencija (F_0) traženoga govornika mjerena je na temelju snimki 1 i 2. Fundamentalna frekvencija smatra se jednim od najsigurnijih akustičkih parametara u forenzičkoj identifikaciji govornika kao zvučni korelat brzine vibracija glasnica. Određena je anatomijom larinksa pa se pretpostavlja da će se na njoj u najmanjoj mjeri manifestirati promjene unutar samoga govornika. Osim toga, taj je parametar otporan na telefonsku transmisiju, što je posebice važno u forenzičkoj fonetici (Braun, 1995).

Izmjerena su prva tri formanta svih vokala sa snimke 1 i snimke 2, posebno za oba govornika. Izračunani su prosjeci formantnih frekvencija iz 10 riječi za svaki vokal te su se uspoređivali s frekvencijama formantata standardnoga hrvatskoga govora. Statističkim postupkom (z-test) nastojala se utvrditi značajnost razlike između formantata analiziranih govornika i standardnih vrijednosti.

Prosječno trajanje vokala traženoga prepoznatoga govornika izračunano je na temelju 10 riječi za svaki vokal, a ti su se rezultati usporedili s prosječnim trajanjem vokala standardnoga govora na temelju Bakranovih podataka (1996). Analiza izgovora vokala važna je u forenzičkoj fonetici stoga što je izgovor vokala tipičan za određeno geografsko podrijetlo. Iako su i suglasnici razlikovni s obzirom na geografsko podrijetlo i uzimaju se u obzir u tehničkome forenzičkom slušnom prepoznavanju govornika, samoglasnici su pogodniji za akustičku analizu.

Naveden pomoćni parametar akustičke analize, tj. dugotrajni prosječni spektri glasa tijekom govora (LTASS) obavještavaju o spektralnoj raspodjeli govornoga signala u nekom vremenu. Kako su vremenski parametri promjenljivi, npr. ista riječ ili izraz mijenjaju se u trajanju za istoga govornika, tehnička verifikacija govornika daje prednost dugotrajnomu prosječnom spektru govora pred kratkotrajnim spektralnim analizama, jer LTASS daje suprasegmentalnu sastavnicu govora, obavještava o glasovoj kvaliteti, odnosno timbru. Ovdje se uspoređivao LTASS i traženog govornika i sugovornika sa snimke 1 i snimke 2.

Uz to, akustičko–statističkim postupcima izračunani su indeksi razlika (SDDD) i indeksi sličnosti (R) za traženoga govornika sa snimke 1 i snimke 2 te uspoređujući traženoga govornika s nekoliko drugih muških glasova, koji su se koristili kao *mućci* (engl. *foils*) u postupku slušnog prepoznavanja. Indeks razlika, tj. indeks standardne devijacije razlika distribucije (SDDD) i indeks sličnosti (R) koriste se u usporedbi glasova u forenzičkoj fonetici, a formule za izračunavanje prilagodene su prema Harmegniesu (1995) i Harmegniesu i Landercyju (1985). U teorijskom smislu SDDD bi bio 0 kada bi LTAS crte bile gotovo paralelne, a najveća standardna devijacija bila bi kad ne bi bilo nijedne točke podudarnosti. Indeks sličnosti (R) obavještava nas o količini kovariranja dvaju spektara, a koeficijent međukorelacije bio bi 1 kada bi sličnost bila potpuna.

2.4. Lingvistička analiza

Budući da akustički materijal nije bio obilan jer se radilo o manje od 10 minuta snimljenoga materijala, napravljena je i lingvistička analiza. Sinergija fonetske i lingvističke analize daje pouzdanije rezultate jer se migovi kvalitete glasa, dobiveni slušnom analizom nadopunjuju migovima lingvističke analize (korištenje idiosinkratičkih riječi, regionalizama i dijalektalizama). Prvi dio analize uključivao je usporedbu fonetske transkripcije govora s dobivenim službenim zapisnikom. Lingvistička analiza podupire mapiranje podrijetla govornika (na leksičkoj, ali i fonološkoj razini), a ima i pragmatičku korist u utvrđivanju točnosti i pouzdanosti transkripcije nastale u istražnom postupku.

Ukupan broj idiosinkratičkih riječi u odnosu prema ukupnom broju riječi u govoru traženoga govornika uspoređen je s čestotom tih riječi u hrvatskome jeziku prema podacima Mogaša, Bratanić i Tadića (1999).

3. REZULTATI I RASPRAVA

3.1. Rezultati slušne fonetske analize

3.1.1. Fonetski opis izgovora i glasa traženoga govornika

Utvrđeno je da traženi govornik pripada štakavskom ikavskom bosanskom govoru: sintaktički i morfofonološki njegov je govor istočna, bosanska varijanta. Fonetski, na suprasegmentalnoj razini izgovora naglasaka govornik ostvaruje markirane naglaske istočne varijante, što znači da dulji dugouzlazni naglasak i ostvaruje dulje zanaglasne duljine u usporedbi s hrvatskim standardnim tražanjem. Uz to, ostvaruje istočne oblike genitiva množine s naglašenim drugim slogom i dugim zadnjim slogom [ʒivã:tsa:], redovito prebacuje silazne naglaske sa značenjskih riječi na proklitiku, neoslabljeno i oslabljeno, što odudara od verificiranog općeprihvatljivog i prihvaćenoga govora hrvatskoga (Škarić, 2009) koji je po određenju alokalan, a prebacivanja se događaju i poželjna su samo u nekim slučajevima. Tražena osoba redovito prebacuje sve silazne naglaske kao u neoslabljenim primjerima *u koliko /ùkolko/*, i oslabljenim *bez veze /bèzveze/*.

Traženi govornik na segmentalnoj razini ima karakterističan dijalektalan izgovor samoglasnika. Vokali su predniji i zatvoreniji od kardinalnoga izgovora vokala u hrvatskome standardu. Na fonološkoj razini redovito je prisutna zanaglasna redukcija, odnosno sinkopiranje samoglasnika, poput *ljud'ma* [lǔ:dmɐ], *proš'o*, *sračun'o*, *rek'o*, *uplat'mi*, *doš'o*, *kup'li*, *naš'o*, *vrač'o*. U nekim riječima dolazi do obezvučavanja zvučnog suglasnika, primjerice [d] u *Vod'cama* [wõdʑɐmɐ]: zbog ispadanja samoglasnika došlo je i do suviše asimilacije po napetosti, jer se [d] našlo uz /c/ [ts], što se u govoru ostvarilo kao [d]. Vokal /a/ često se, za razliku od hrvatskoga kardinalnog izgovora, ostvaruje zanaglasno i kao središnjiji i nešto zatvoreniji [ɐ]. Koji put ispred sinkopiranog vokala dolazi i do ispadanja suglasnika. Pri sinkopiranoj redukciji, sažimanju dvaju samoglasnika u jedan u finalnoj poziciji riječi, tj. [ao] u [o] kao u primjerima *prošao* [prõʃõ:], *došao* [dõʃo:], *sračunao* [srãtʃuno:], *rekao* [rêko:], *psovao* [psõwo:], *jebao* [jêbo:], sažimanje se u fonološkom smislu očituje i znatnijim duljenjem vokala [o:]. Redovita je zanaglasna redukcija do ispadanja vokala [i]: *koliko* /kol'ko/, *kupili* /kup'li/, *govornice* /govorn'ce/, *dobiti* /dob'ti/, *ugasilo* /ugas'lo/. U naglašenoj poziciji [i] je zatvoreniji od standardnog izgovora, a nakon zanaglasnih redukcija, završni je katkad i labaviji, tj. [i], kao u primjerima *kupili* [kũ:plɪ], *dobiti* [dõbʲtɪ]; pritom se zvučni suglasnici obezvučuju zbog dodatne asimilacije zbog ispadanja i slijeda bezvučnog suglasnika kao u potonjem primjeru. Što se tiče izgovora suglasnika /č/[tʃ], traženi govornik izgovara ga prema umekšanom, palataliziranom.

Prema deskriptivnom fonetskom protokolu glas traženoga govornika je lagano povišen, glasan, u suprasegmentalnom postavljanju uskog raspona čeljusti, katkad dentaliziran, u laringalnom postavljanju modalne fonacijske vrste. Drugi glas, tj. glas sugovornika s druge snimke, kojega nije trebalo prepoznati, prosječne je visine, bez zamjetnih modifikacija vokalnog trakta. Pripada hrvatskomu južnjačkom miješanom ikavsko-jekavskom dijalektu.

3.1.2. Rezultati na temelju AP – SPID protokola

Rečeno je da se slušno prepoznavanje govornika provodilo prema fonetskom protokolu u forenzičkoj fonetici, tj. prema AP – SPID. Dvoje ekspertnih fonetičara uspoređivalo je 11 parova govornika. Kombinacija i redosljed bili su poznati samo fonetičaru, drugom autoru koji je za potrebe prepoznavanja montirao parove glasova slučajnim redosljedom. Parovi su se uspoređivali u različitim parametrima od tona, kvalitete glasa, preko intenziteta, prozodije, procjene dijalektalnog podrijetla, artikulacije vokala i konsonanata i u njihovim poddimenzijama. Rezultati se uobičajeno navode u prosjeku i rasponima prepoznavanja. Najveće prepoznavanje bilo je 96%–tno za par X1–M3, što je par tražene osobe sa snimke 1 i snimke 2 (tablica 1). Visok postotak prepoznavanja (94%) imao je i par M2–A8, što je bio mućak 2 sa samim sobom te par D5–M1 (91%), što je bila kombinacija sugovornika sa snimke 1 i snimke 2. U forenzičkoj fonetici tek se prepoznavanje više od 90% uzima u obzir pri donošenju mišljenja da bi se moglo raditi o istoj osobi. Prepoznavanje je pak bilo najmanje (14%) u paru M3–N3 koji je kombinacija traženoga govornika s mućkom 1 (tablica 2). Rezultati prosjeka postotaka prepoznavanja s rasponima za sve parove prikazani su na tablici 3.

Slušnom prepoznavanju traženoga govornika uvelike je pridonijela karakteristična vokalska boja koja obavještava o njegovu dijalektalnom i regionalnom podrijetlu, a koje je procijenjeno kao strano, tj. bosansko. U postupku prepoznavanja tražene osobe, tj. osumnjičenika, znakovita je bila i visina tona te opća kvaliteta glasa, glas je bio blistave kvalitete i blago nazaliziran.

Tablica 1. Rezultati prepoznavanja govornika (%) prema AP – SPID protokolu forenzičke fonetike (usporedba traženoga govornika sa snimke 1 i snimke 2 – najveća sličnost)

	REZULTAT / SCORE	RASPON / RANGE
TON / PITCH		
Visina tona / Level	10	10 – 10
Promjenljivost / Variability	10	10 – 10
Obrasci pravilnosti / Regularity patterns	9	9 – 9
KVALITETA GLASA / VOICE QUALITY		
Opća kvaliteta / General quality	10	10 – 10
Fonacijski tip / Phonation type	9,5	9 – 10
Ostalo / Other		
INTENZITET / INTENSITY		
Promjenljivost / Variability	10	10 – 10

DIJALEKT / DIALECT		
Regionalan / Regional	10	10 – 10
Strani / Foreign	10	10 – 10
Idiolekt / Idiolect	9,5	9 – 10
ARTIKULACIJA / ARTICULATION		
Vokali / Vowels	10	10 – 10
Konsonanti / Consonants	9,5	9 – 10
Pogrešan izgovor / Misarticulations		
Nazalnost / Nasality	8,5	8 – 9
PROZODIJA / PROSODY		
Brzina / Rate	10	10 – 10
Prekidi govora / Speech Bursts	9,5	9 – 10
Staccato		
OSTALO / OTHER		
Govorni poremećaji/ Speech disorders		
Srednja vrijednost / Mean	96%	94 – 98%

Tablica 2. Najslabiji rezultati prepoznavanja govornika (%) prema AP – SPID protokolu forenzičke fonetike (par M3 – N3: usporedba traženoga govornika i mučka 1 – najmanja sličnost)

	REZULTAT / SCORE	RASPON / RANGE
TON / PITCH		
Visina tona / Level	0	0 – 0
Promjenljivost / Variability	0,5	0 – 1
Obrasci pravilnosti / Regularity patterns	0,5	0 – 1
KVALITETA GLASA / VOICE QUALITY		
Opća kvaliteta / General quality	1,5	1 – 2
Fonacijski tip / Phonation type	1,5	0 – 3
Ostalo / Other		
INTENZITET / INTENSITY		
Promjenljivost / Variability	0	0 – 0
DIJALEKT / DIALECT		
Regionalan / Regional	0	0 – 0

Strani / Foreign	0	0 - 0
Idiolekt / Idiolect	0	0 - 0
ARTIKULACIJA / ARTICULATION		
Vokali / Vowels	0,5	0 - 1
Konsonanti / Consonants	1	1 - 1
Pogrešan izgovor / Misarticulations		
Nazalnost / Nasality	2	0 - 4
PROZODIJA / PROSODY		
Brzina / Rate	1	0 - 2
Prekidi govora / Speech Bursts	1	0 - 2
Staccato	0,5	0 - 1
OSTALO / OTHER		
Govorni poremećaji/ Speech disorders		
Srednja vrijednost / Mean	14%	1 - 12%

Tablica 3. Rezultati slušne procjene (%) za sve parove

PAROVI	RASPON (%)	PROSJEK (%)
A5+D5	30-46	44
A8+A1	35-49	50
X1+M3	94-98	96
M2+A1	27-39	40
D5+M1	80-87	91
M3+N3	1-12	14
X1+M2	12-18	23
A5+Q1	20-29	31
M2+A8	91-95	94
X1+M2	12-18	23
D5-N3	3-15	16

3.2. Rezultati akustičke analize

3.2.1. Rezultati mjerenja fundamentalne frekvencije

Fonetski opis traženoga govornika utvrdio je da se radi o muškoj osobi lagano povišenoga glasa, a to je potvrdilo i mjerenje fundamentalne frekvencije koja je u slučaju traženoga govornika viša od prosječnoga muškoga glasa. Prema Fantu (1970), Varošaneć-Škarić (2005) raspon fundamentalne frekvencije

za ljudski glas iznosi 60–240 Hz, za muškarce iznosi prosječno 120 Hz, a za žene 200 Hz. Prosječna F_0 traženoga govornika sa snimke 1 bila je 201 Hz, a na drugoj snimci 197 Hz, s gotovo jednakim minimalnim i maksimalnim rasponima od 76 i 316 Hz i 76 i 317 Hz, a medijani su također viši kod traženoga glasa, tj. iznose 200 i 195 Hz prema 175 Hz kod sugovornika (tablica 4). Sugovornik je imao nižu prosječnu F_0 od traženoga govornika, tj. 177 Hz. Kako je razgovor bio obojen emocionalno, tj. svadali su se, povišenijem glasom traženoga govornika bila je u ljutnji i vrlo glasnom govoru znatno viša F_0 . Primjerice, utjecaj glasnoće i tona u forenzičkom surječju istraživao je French (1998). Zbog transmisije i presnimavanja zvuka ipak dolazi i do laganih izobličenja zvuka, a mjerenja F_0 vrlo glasnoga govora i u studijskim uvjetima snimanja utječu na mjerenje, jer programi uzimaju u obzir najjače amplitude, koje u vrlo glasnom govoru mogu biti i harmonici F_0 , pogotovo kad se radi o maksimalnim vrijednostima. Srednja glasnoća i neutralni govor daju pouzdanije rezultate u akustičkim mjerenjima, ali u forenzičkim slučajevima takav je govor moguće dobiti samo pri postupku verifikacije u intervjuu s fonetičarem. U ovome slučaju tražena osoba po savjetu branitelja naknadno je odbila verifikacijski postupak, koji je zatražio istražni sudac.

Tablica 4. Vrijednosti fundamentalne frekvencije za traženoga govornika i sugovornika

Izmjerene vrijednosti (u Hz)	Traženi govornik X1	Traženi govornik M3	Sugovornik M1
F_0	201	197	177
Standardna devijacija	49	46	52
Minimum	76	76	75
Maksimum	316	317	317
Medijan	200	195	175

3.2.2. Rezultati mjerenja formantnih frekvencija

Izgovor vokala, u užem smislu vokalska boja, važan je pokazatelj dijalektalnog, regionalnog pa i društvenog podrijetla govornika, stoga je analiza vokala u postupku prepoznavanja govornika vrlo važna. Osim slušne analize, provode se i akustička mjerenja kao što su mjerenje vrijednosti formantata, mjerenje trajanja vokala te njihovo uspoređivanje sa standardnim vrijednostima. Formanti su rezonancijske frekvencije vokalnog trakta, odnosno frekvencije istaknutih dijelova spektra, a u suodnosu su s veličinom i oblikom vokalnog trakta. U užem smislu vokalski su formanti F1 i F2, iako se utjecaj izgovora vokala reflektira i na F3, pa se i on u zadnje vrijeme uzima u obzir pri analizi vokalskog izgovora. Što se tiče mjerenja vokalnog trakta, uzima se viši formant F4, koji je stoga vokalni formant (a ne vokalski) jer na njega više ne utječe izgovor vokala, već je rezonantna frekvencija cjelokupnoga vokalnog trakta. Općenito, u akustičkome mjerenju analize formantnih frekvencija treba biti oprezan pri prepoznavanju, ono treba slijediti tek nakon pomne slušne analize da ne bi

došlo do tzv. pozitivnih te negativnih pogrešaka (upućujemo na Rose, 2002). Budući da su u slušnoj analizi uočene razlike u izgovoru vokala tražene osobe i sugovornika, u programu *Praat* mjerene su vrijednosti vokalskih formanata F1, F2 i F3. Pri tome treba napomenuti da je F1 relevantan samo za vokal /a/, jer na njega zbog prosječnih viših vrijednosti ne utječe GSM transmisija. Uspoređivani su drugi i treći formanti svih vokala jer ni na njihove frekvencije ne utječe GSM transmisija, tj. ne izobličavaju se vrijednosti. Napravljen je prosjek od 10 riječi sa svake snimke te se uspoređivao sa standardnim vrijednostima hrvatskoga jezika (prema Škarić, 1991). Statističkim izračunom (z-test) provjeravala se značajnost razlike. Budući da se na obje snimke radilo o razgovoru preko GSM-a čiji je frekvencijski raspon bio od 300 do 3400 Hz, u analizi prvog formanta, kao što je rečeno, uzet je u obzir samo vokal [a] čija je prosječna vrijednost 700 Hz. Upravo je prvi formant vokala [a] pokazao statistički značajnu razliku u odnosu prema standardnoj vrijednosti ($p = 0,0001$; tablica 5).

Drugi formant pokazuje statistički značajnu razliku u odnosu prema standardnim vrijednostima za vokale [a] ($p = 0,0006$), [o] ($p = 0,00002$), [u] ($p = 0,00002$), F_2 je značajno niža za [a] i [o] te viša za [u]. Treći formant pokazuje statistički značajnu razliku samo za vokal [i], tj. značajno nižu vrijednost ($p = 0,000000006263$), što se može objasniti prednjom zaokruženošću bosanskog izgovora vokala. Na formantne vrijednosti traženoga govornika, tj. na spuštanje vrijednosti viših formanata utjecalo je i dentalizirano postavljanje glasa, što pokazuju i druga istraživanja (Varošaneć-Škarić, 2005). U usporedbi sa sugovornikom traženi govornik ima viši F_2 vokala [i] ($p = 0,008$), jer mu je [i] prednjiji i zatvoreniji od sugovornika te niži F_3 vokala [o] ($p < 0,001$) zbog utjecaja dentaliziranog postavljanja glasa. Iako je četvrti formant vrlo važan u identifikaciji govornika jer daje informacije o kvaliteti glasa, prosječne vrijednosti više su od 3000 Hz pa zbog deformacije zvuka uslijed telefonske transmisije, vrijednosti ne bi bile pouzdane, stoga se nije mjerio. Dakle, formanti vokala [a], [i], [o], [u] tražene osobe statistički se značajno razlikuju od vrijednosti formanata standardnoga hrvatskoga govora. Autori su svjesni da ove tvrdnje valja potkrijepiti daljnjim istraživanjima, jer prosječne formantne vrijednosti za standardan izgovor hrvatskih vokala kod Škarića (1991) i Bakrana (1996) nisu dobivene na temelju sto i više govornika općega hrvatskog posebno za muške i ženske glasove, koliko bi bilo potrebno za čvršću usporedbu odnaka idiosinkratičnog izgovora u odnosu prema standardnom izgovoru vokala.

Tablica 5. Rezultati usporedbe i statističke značajnosti F1 (Hz) između traženoga govornika i standardnih vrijednosti (prema Škarić, 2009)

VOKAL	Traženi govornik	Standardni
[i]	423	360
[e]	535	500
[a]	737 $p=0.0001$	700
[o]	550	450
[u]	501	380

Tablica 6. Rezultati usporedbe i statističke značajnosti F2 između traženoga govornika i standardnih vrijednosti (prema Škarić, 2009)

VOKAL	Traženi govornik	Standardni
[i]	2124	2200
[e]	1838	1800
[a]	1378 p=0.0006	1400
[o]	1006 p=0.00002	1150
[u]	1040 p=0.00002	750

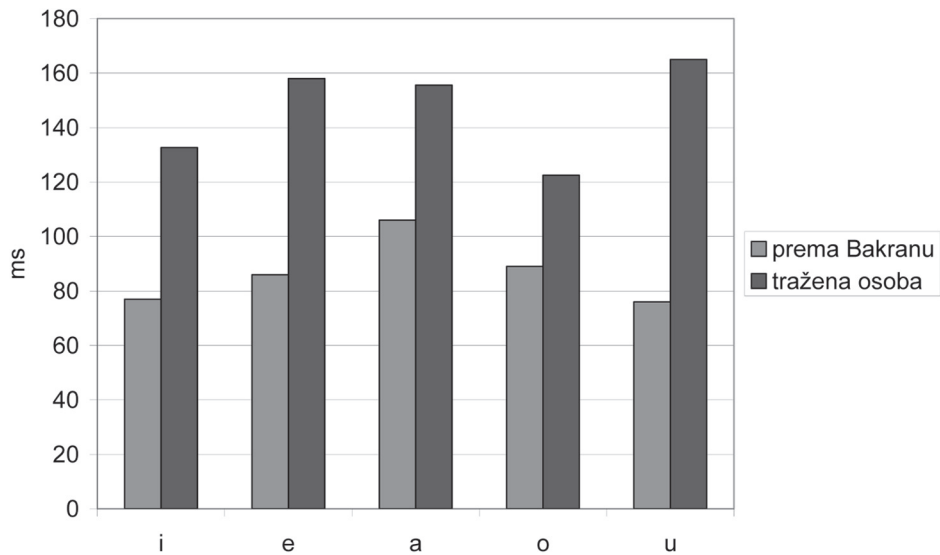
Tablica 7. Rezultati usporedbe i statističke značajnosti F3 između traženoga govornika i standardnih vrijednosti (prema Škarić, 2009)

VOKAL	Traženi govornik	Standardni
[i]	2570 p= 0, 000000006263	2850
[e]	2375	2500
[a]	2261	2250
[o]	2381	2300
[u]	2452	2450

3.2.3. Rezultati trajanja vokala

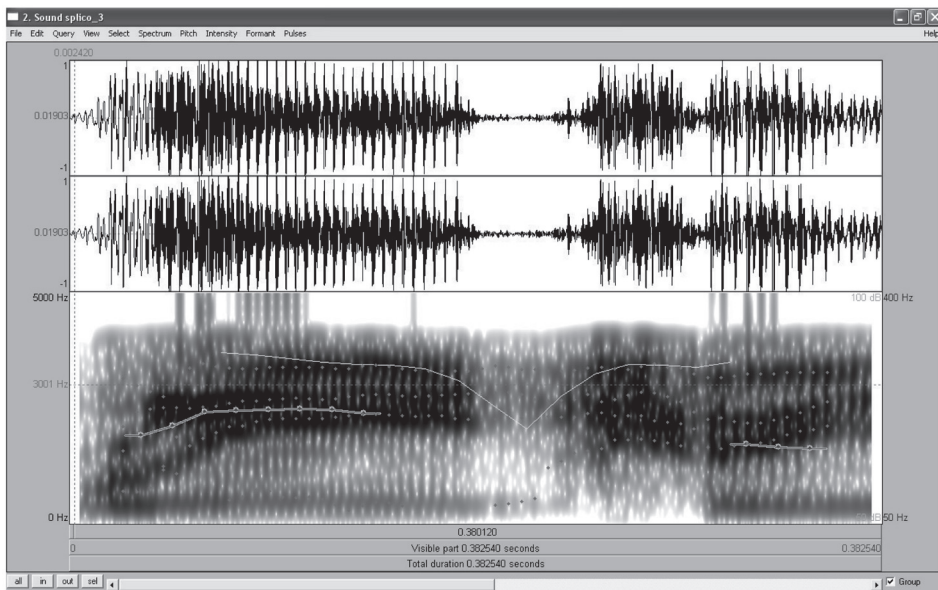
Slušnom analizom utvrđeno je da je trajanje naglašanih vokala dugih naglasaka (dugouzlaznog i dugosilaznog) traženoga govornika duže nego u hrvatskome općeprihvatljivom izgovoru. Budući da za hrvatski jezik postoje izmjerenja trajanja vokala u različitim akcenatskim uvjetima u radu Bakrana (1996), mogla se napraviti usporedba. Naime, i Bakran uzima zajedničku prosječnu vrijednost za duge akcente s obzirom na to da kretanje tona ne utječe znatno na trajanje (Bakran, 1996:258). Izmjerena su trajanja naglašanih vokala u povezanome govoru, jer su oni i slušnom analizom prepoznati kao snažan marker govora traženoga govornika. To ne znači da ne postoji razlika i u izgovoru kratkih naglasaka, ali to je slušno teže utvrditi, pa se njihovo vrijeme trajanja nije ni akustički provjeravalo. Za deset riječi za svaki dugi naglašeni vokal sa snimke 1 i deset riječi sa snimke 2 izmjereno je prosječno trajanje (ms), primjerice u riječima *ljúdi, lúđ, drúg, kúpío, ugásío, vrátío, dvâ, plân, znâm, nísu, fíno, príče, pět, iz početka (ispočetka), dôći, stô* itd.).

Usporedba je pokazala da su svi dugi naglašeni vokali traženoga govornika duži od standardnih vrijednosti (slika 1), a najviše za vokale [i] i [e]. Slika 2a pokazuje trajanje vokala [i] (oscilogram i sonogram široke analize) tražene osobe pri izgovoru idiosinkratične riječi *Splíćo* [splī:tɕo], a slika 2b u standardnom izgovoru.

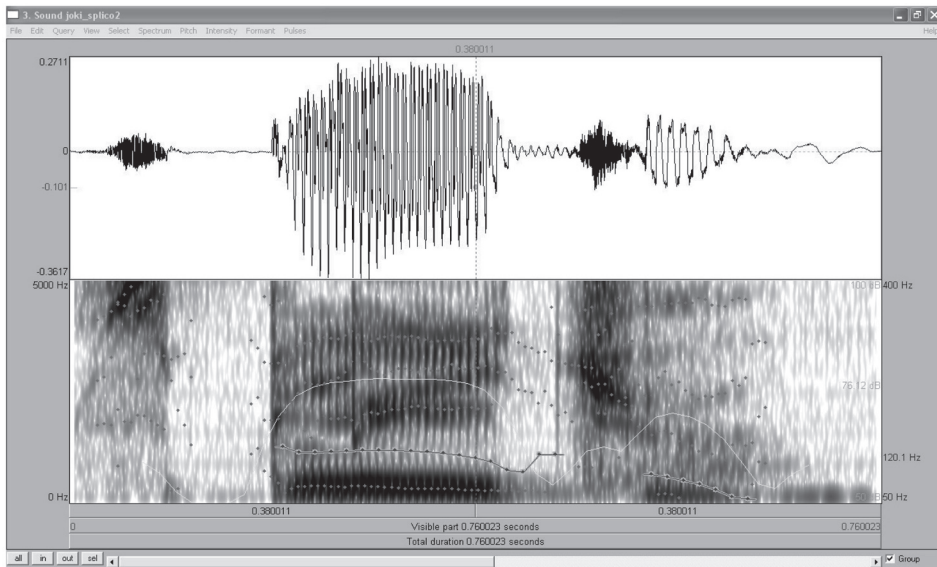


Slika 1. Odnos trajanja dugih naglašenih vokala traženoga govornika i standardnih vrijednosti (ms)

a



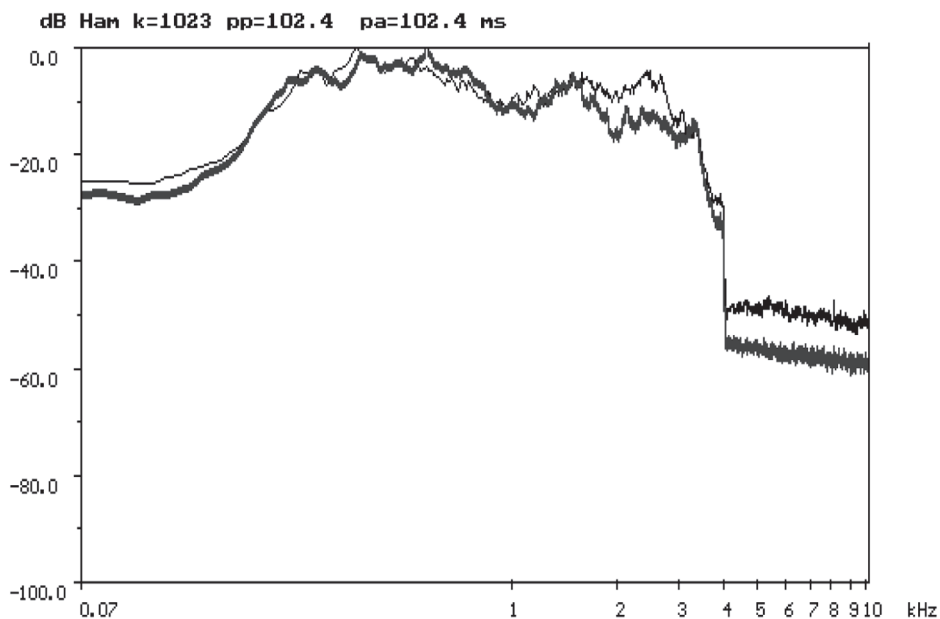
b



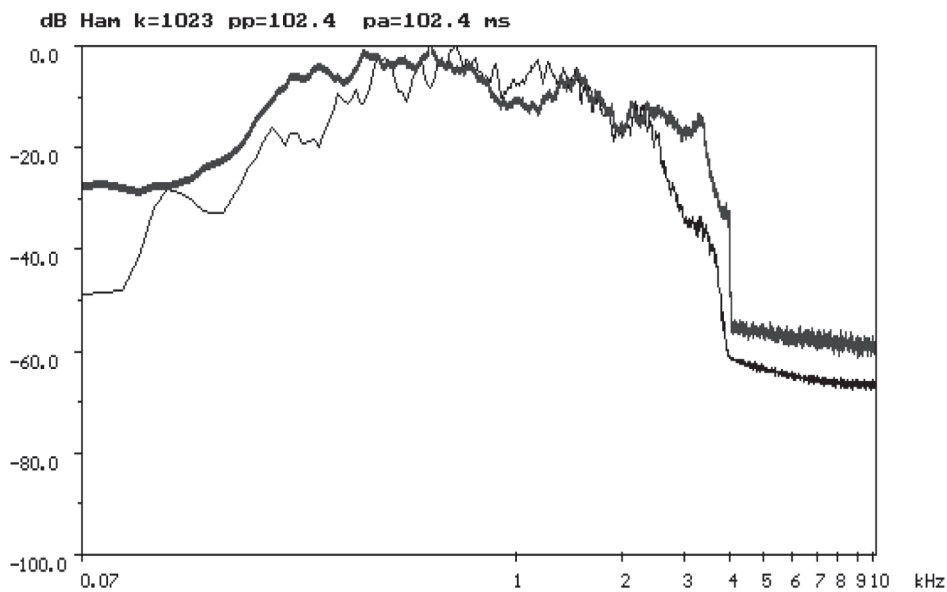
Slika 2. Oscillogram i sonogram široke analize riječi *Splićo* u izgovoru tražene osobe (a) i u standardnom izgovoru (b)

3.2.4. Rezultati usporedbe dugotrajnih prosječnih spektara (LTASS) i akustičko–statističkih postupaka

Na slikama dugotrajnih prosječnih spektara (slika 3. i 4) zamjetna su tipična izobličenja u nižim i višim dijelovima spektra do kojih dolazi zbog transmisije preko GSM uređaja, što znači da se vrijednosti u nižem dijelu spektra ispod 400 Hz povisuju te da se one ispod 3400 Hz snižavaju. Stoga se u statističko–akustičkom postupku izračunavanja indeksa sličnosti između spektara (R) i indeksa različitosti (SDDD) uzima manji broj točaka negoli bi se uzeo na temelju studijskih snimaka. Dugotrajni spektri tražene osobe sa snimke 1 (podebljana crta) i snimke 2 (tanka crta) na slici 3. pokazuju veću sličnost nego spektri između traženoga govornika (deblja crta) i sugovornika (tanja crta) (slika 4), što se zamjetno vidi u višem području blistavosti. Indeks sličnosti između spektara tražene osobe veći je nego između tražene osobe i sugovornika ($R = 0,89 : 0,85$), a indeks različitosti manji je za traženoga govornika nego između traženoga govornika i sugovornika ($SDDD = 2,10 : 3,75$) (tablica 8). Takav rezultat upućuje da u slučaju tražene osobe postoji vjerojatnost da se radi o istoj osobi. Najmanji indeks sličnosti i najveći indeks različitosti pokazao se između tražene osobe i mučka ($R = 0,74$; $SDDD = 12,85$), što znači neprepoznavanje. Naravno je da se znatno bolji rezultati u zvučnom prepoznavanju mogu dobiti u slučaju studijskih snimaka (na tablici 8: mućak 2), stoga u etičkom smislu statističko–akustički postupci utvrđivanja indeksa sličnosti i različitosti trebaju biti samo dodatni. Stoga se jedino slušnom procjenom na temelju AP – SPID protokola mogu utvrditi potpuna prepoznavanja odnosno neprepoznavanja osoba, a akustički se može utvrditi samo vjerojatnost da se radi o prepoznavanju odnosno neprepoznavanju.



Slika 3. LTASS traženoga govornika sa snimke 1 i snimke 2



Slika 4. LTASS traženoga govornika i sugovornika

Tablica 8. Prosječne vrijednosti indeksa sličnosti i razlika među spektrima izračunane su za svih 11 parova.

PAROVI	R	SDDD
A5+D5	0,85	3,75
A8+A1	0,69	11,62
X1+M3	0,89	2,10
M2+A1	0,74	12,85
D5+M1	0,95	2,28
M3+N3	0,85	6,95
X1+M2	0,74	12,85
A5+Q1	0,92	3,68
M2+A8	0,98	2,06
X1+M2	0,74	12,85
D5-N3	0,81	4,05

3.3. Rezultati lingvističke analize u govoru osumnjičenika

Lingvističkom analizom utvrđivali su se posebni fonološki, sintaktički i leksički čimbenici u govoru tražene osobe. Neke smo osobitosti, koje su prisutne u izgovoru na fonološkoj razini, primjerice redovite zanaglasne redukcije, odnosno sinkopiranje samoglasnika, naveli u poglavlju 3.1.1. Također je primijećeno kraćenje infinitiva, primjerice: *uzet'*, *nač'*, *jest'*, *pričat'*, *govorit'*, *pokazat'*, *završit'*, *spremit'*, *dat'*, *vidit'*, *kupit'*, *istuširat'*, *obuč'*, ikavski primjeri poput *donit'*, karakteristična skraćivanja za hrvatske dijalekte od glagola *vući* i *ići*, poput *vuč'* [wu:tɕ], *ić'* [i:tɕ], *otič'*, *odvuč'*. Traženi govornik skraćuje i imperativne oblike poput *rec'* umjesto *reci*. Za govor tražene osobe karakteristična je omisija inicijalnog [x], primjerice /'odo/, /'odaju/ [ð:daju:], /'oćeš/, /'ajd'/ umjesto /hodao/, /hodaju/, /hoćeš/, /hajde/, a u nekim slučajevima izostavljanje inicijalnoga vokala [o], primjerice /'vamo/. Također su primjetni specifični morfološki oblici poput: *ne merem* umjesto *ne mogu*, gramatičkih oblika, npr. *taksijom* umjesto *taksijem*, upotreba nominativa umjesto vokativa, npr. *drug* umjesto *družę*, sintaktičkih poput *imaš da slušaš* te upitna čestica *šta*, za razliku od drugoga govornika.

No, za analizu je najzanimljivija bila leksička razina. Uz oblike poput *milijon*, *kužiš*, *okej*, izraza poput *i to je to*, zanimljiv je izbor riječi koji, između ostaloga, prilično jasno otkriva govorno podrijetlo tražene osobe. Riječi poput *deder*, *prodavnica*, *bolan* i prije svega *jaran*. Riječ *jaran* u snimci razgovora od 7,39 minuta pojavila se 59 puta, a u razgovoru na snimci od 2,16 minuta pojavila se 10 puta kod prepoznatoga govornika. Dakle, tijekom 9,65 minuta snimaka, ukupan broj riječi koje je izgovorila tražena osoba bio je oko 4000, a riječ *jaran* izgovorena je 69 puta. Relativna čestota riječi *jaran* iznosi 0,017, što je u usporedbi s čestotom u hrvatskome jeziku (prema Moguš, Bratanić i

Tadić, 1999) koja iznosi 0,0001, 170 puta više. Relativna čestota riječi *bolan* u čestotnom rječniku iznosi 0,0013, apsolutna čestota 13, a u ovom znatno manjem korpusu traženi govornik izgovorio je riječ šest puta. U uzorku drugoga govornika te se idiosinkratične riječi uopće nisu pojavile, već neke druge poput *pajdo*, *razumiš ti*, *kontaj*. Zanimljivo je da se riječ *pajdo* u značenju prijatelja, sudruga ne pojavljuje u čestotnom rječniku, nego samo oblik *pajdaš*. Dakle, u južnjačkome štokavskom ikavsko-jekavskom govoru drugoga govornika pojavljuje se idiosinkratičan oblik *pajdo* četiri puta.

3.3.1. Usporedba službenoga zapisnika i transkripcije fonetičara

Drugi dio analize odnosio se na usporedbu službenog zapisnika u istražnom postupku i transkripta fonetičara. Usporedba je pokazala nedostatke prvog načina transkribiranja koje je katkada površno, a katkada i netočno, budući da službenici nisu uvježbani u slušanju i razumijevanju govora praćenog bukom i preklapanjima govora. U takvim situacijama javljaju se pogreške koje mogu utjecati na cjelokupnu istragu. Primjerice, jedan od govornika govori da sutra mora biti kod Matulja, a u policijskom zapisniku piše da je rekao da mora biti kod Macelja (što su mjesta u različitim dijelovima Hrvatske). Nadalje, ne bilježe se idiosinkratični uzvici i fatičke čestice (kao *ej!*, *a?*), a na nekim dijelovima dodaju se samo objašnjenja »nerazumljivo«, a fonetičar tu bilježi i nepravilne i reducirane oblike.

Takoder, službeni transkript ne bilježi karakteristike govora kao što su sinkopiranje, redukcije niti karakteristične prozodijske crte (poput mjesta ili vrste naglasaka, prebacivanja naglasaka na proklitiku), dugi jat bilježi se dvosložno, iako govornik uvijek ostvaruje jednosložan izgovor itd. Sve su te osobine govora važne u određivanju regionalnog, dijalektalnog i društvenog podrijetla govornika, što je nužno u postupku identifikacije, prepoznavanja i verifikacije u forenzičkoj fonetici i lingvistici (Varošaneć-Škarić i Kišiček, 2010)

Tablica 9. Usporedba dijela službenog zapisnika i fonetske transkripcije telefonskog razgovora

POLICIJSKI ZAPISNIK	TELEFONSKI RAZGOVOR
M (sugovornik) mu govori ma daj čovjeće da je bio gore cijelo vrijeme u Vodicama, murija ga je štopala...	D5: ma daj, deder, bio sam gore cijelo [cjê:lo] vrijeme [vrje:me], u Vodicama [wõditsama] me murija štopala, razumiš'!?
X (traženi govornik) pita da gdje u Vodicama...	X1: pa di u Vod'cama [wõdtsəmɐ] bolan?
M (sugovornik) govori da su ga vidjeli 100 posto...	D5: ma vidili ste me čovjeće sto posto
X (traženi govornik) pita da gdje ga je u Vodicama murija štopala...	X1: pa gdje te u Vod'cama [wõdtsəmɐ] murja štopala

4. ZAKLJUČAK

Stvarni slučaj prepoznavanja muškoga govornika uključivao je slušno prepoznavanje prema AP – SPID protokolu, akustička mjerenja fundamentalne frekvencije, formanta vokala, trajanja vokala, mjerenja dugotrajnih prosječnih spektara (LTASS), akustičko–statistička mjerenja indeksa sličnosti i različitosti te forenzičkolingvističku analizu. Postupkom AP – SPID protokola utvrđeno je 96%–tno prepoznavanje, a akustička i lingvistička analiza poduprla su vjerojatnost da se na snimci 1 i snimci 2 radi o istoj osobi ($R = 0,89$; $SDDD = 2,10$). Rezultati akustičke analize u programu Praat pokazali su da se F_1 i F_2 vokala [a], F_2 za [a], [o] i [u] tražene osobe statistički značajno razlikuju od standardnih vrijednosti u hrvatskom jeziku ($p < 0,001$) te F_3 za vokal [i] ($p < 0,00001$). Vokali tražene osobe akustički se razlikuju i od vokala sugovornika, što pokazuje statistički značajna razlika za F_2 vokala [i] i F_3 vokala [o]. Svi dugi naglašeni vokali traženoga govornika duži su od standardnih vrijednosti, a najviše vokali [i] i [e].

Lingvistička analiza pokazala je razliku na fonološkoj, sintaktičkoj i leksičkoj razini izgovora tražene osobe i sugovornika. Na leksičkoj razini osobito je idiosinkratičnost pokazala riječ *jaran*.

I na ovome se stvarnom slučaju potvrdila važnost ekspertnog slušanja fonetičara koji mogu prepoznati različit izgovor, naglasne oblike te idiosinkratične oblike u izgovoru govornika. Fonetičari trebaju transkribirati govor sa snimki neovisno o postojanju službenog zapisnika ili zapisnika druge vrste vještaka.

Pri akustičkoj usporedbi formantnih frekvencija vokala pokazala se potreba dodatnih istraživanja prosječnih vrijednosti standardnog izgovora radi što točnije usporedbe idiosinkratičnog izgovora govornika.

Sinergija postupaka forenzičke fonetike i lingvistike može pridonijeti pouzdanijim rezultatima prepoznavanja govornika u pojedinim stvarnim forenzičkim slučajevima, osobito u nedostatku veće količine snimljenoga prisluškivanoga govora.

Referencije

- Bakran, J. (1996). *Zvučna slika hrvatskoga govora*. Zagreb: IBIS grafika.
- Boersma, P., Weenink, D. (2009). Praat: doing phonetics by computer, version 5.1.20 www.fon.hum.uva.nl/praat/.
- Braun, A. (1995). Fundamental frequency – how speaker-specific is it?. U: Braun i Köster (ur.), *Studies in Forensic Phonetics*, Beiträge zur Phonetik und Linguistik 64, Trier: Wissenschaftlicher Verlag, str. 9–23.
- Coulthard, M., Johnson, A. (2007). *An Introduction to Forensic Linguistics: Language in Evidence*. London, New York: Routledge.
- French, P. (1998). Mr. Akbar's nearest ear vs. the Lombard reflex: a case study in forensic phonetics. *Forensic Linguistics* 5, 1: 58–68.
- Harrison, Ph., French, P. (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.

- Harmegnies, B., Landercy, A. (1985). Language Features in the Long-Term Average Spectrum. *Revue de Phonétique Appliquée* 73–74–75, 69–79.
- Harmegnies, B. (1995). Contribution à la caractérisation acoustique des sigmatismes – étude de deux indices acoustico-statistiques. U: A. Braun i J.-P. Köster (ur.) *Studies in Forensic Phonetics*, 56–66. Trier: Wissenschaftlicher Verlag Trier.
- Hollien, H. (2002). *Forensic Voice Identification*. San Diego: Academic Press.
- McMenamin, G. R. (2002). *Forensic Linguistics: Advances in Forensic Stylistics*. CRC Press.
- Künzel, H., 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.
- Moguš, M., Bratanić, M. i Tadić, M. (1999). *Hrvatski čestotni rječnik*. Zagreb: Zavod za lingvistiku Filozofskog fakulteta Sveučilišta u Zagrebu i Školska knjiga.
- Nolan, F. (1983 digitally printed version 2009) *The phonetic bases of speaker recognition*. Cambridge University Press.
- Rose, Ph. (2002). *Forensic Speaker Identification*. London, New York: Taylor and Francis.
- Škarić, I. (1991). Fonetika hrvatskoga jezika. U R. Katičić (ur.), *Povijesni pregled, glasovi i oblici hrvatskoga književnog jezika* (str. 61–376). Zagreb: HAZU i Globus.
- Škarić, I. (1993). Prosječni spektar govora kao slika boje glasa. U Likar, Kosmač (ur.), *Strokovno srečanje logopedov Slovenije: Multidisciplinarni pristop v logopediji*; Portorož, Maj 13–15, 1993, 202–205. Portorož: Aktiv logopedov Južne Primorske.
- Škarić, I. (2009). »Hrvatski izgovor«. Nakladni zavod Globus, Zagreb.
- Varošaneć-Škarić, G. (2005). *Timbar*. Zagreb: FF press.
- Varošaneć-Škarić, G. (2008). Verifikacija govornika u forenzičnoj fonetici. *Govor – časopis za fonetiku* XXV, 1, 31–44.
- Varošaneć-Škarić, G. i Bičanić, J. (2007). A comparison of indices of difference and similarity based on voices in real forensic case and in controlled conditions. *Proceedings*. www.icphs20007.de, 16th International Congress of Phonetic Sciences (Eds. Jürgen Trouvain i William J. Barry), pp 2085–2088.
- Varošaneć-Škarić, G. i Kišiček, G. (2010). Phonetic and linguistic forensic identification of male speaker (real case). U *Abstracts for 19th Annual Conference of the International Association for Forensic Phonetics and Acoustics*, 18–21 July 2010, Germany. Trier: Dpt. of Phonetics – University of Trier, str. 34.

Forensic Phonetic identification and linguistic analysis of the speaker

Upon the request of the County Court of Zagreb, on the basis of the recordings obtained by the police eavesdropping of the GSM device, the identification of a male voice was to be performed. The male voice from the recording No. 1 – duration of 2.16 min (CD 1) had to be compared with the male voice from the recording No. 2 – duration 7.39 min. The whole material was phonetically transcribed for the purpose of examining regional pronunciation, dialect and idiolect. For the purpose of sound analysis, dubbing was performed as well. Linguistic analysis has shown the difference in the use of idiosyncratic words, as well as the difference on the phonological level. For instance, recognized speaker used word *jaran* 69 times (which is 170 times more than in the usage of Croatian language). In the sample of two recordings with the recognized speaker, idiosyncratic meaningful words such as [järan; bôlan] show significantly more frequent absolute and relative appearance when compared with the appearance in Croatian language. Those words didn't appear in the sample of the second speaker, but the recordings had some other words that coincided such as [razümi:š; päjdo]. Although the phonetic and linguistic analyses differ, they are tightly connected. According to the forensic phonetics protocol (SPID, Hollien 2002, Varošaneć-Škarić, 2008), auditory perception analysis of speech show high possibility of male person identification, i.e. experts' recognition was 96% in range from

94% to 98%. Recognized speaker belongs to the group of *stakavian ikavian* Bosnian speeches: syntactically and morphologically that is eastern Bosnian variant (specific use of present tense, instead of negation plus infinitive [ně:tçe; xõ:dati] he uses the typical form [ně:tçe: da_õ:daju:], he is shortening infinitives, omitting initial [x], for example [õ:daju:], [õ:tçeš]) instead of RP Croatian [xõ:daju], [xõ:tçeš]. Phonetically, on the suprasegmental level of accent realization, the recognized speaker uses marked accents of the eastern *stakavian* variant (longer long-rising accent and after-stress length), he is shifting falling accents from the meaningful words to the proclitic, and on the segmental level his pronunciation of vowels is typical for *stokavian* dialect. Vowels are realized more forward and closed than those of cardinal vowels of Croatian Received Pronunciation, reduction and after-stress omitting of vowels appear. Phonetic description of voice shows slightly higher voice, narrow jaw aperture, dentalization, and modal phonation type. Furthermore, formant analysis of all vowels was made for both speakers (F2, F3), while vowel /a/ provided the analysis of F1 as well, because its average values are around 800 Hz and is not affected by the distortion of transmission. Formant analysis based on the stressed vowels has shown the difference (z-test) when compared with standard Croatian Pronunciation (for lower F1 /a/: $p = 0,001$, F2 higher for /a/, /e/, /o/, /u/) and in comparison with the other speaker F2 was higher for vowel /i/ ($p = 0,08$), F3 lower for vowel /o/ ($p < 0,001$). Long average spectra of speech that give information about voice timbre are different for those two speakers. Typical distortions in lower and higher parts of the spectrum during the transmission via GSM device were taken into consideration.

According to Harmegnies (1995) similarity index (R) was calculated for two spectra from different recordings of the same identified male speaker was 0.91 and dissimilarity index (SDDD) was 0.96. In pair with unknown speaker R was in range from 0.75 to 0.88 and SDDD in range from 3.32 to 4.50. According to total results it can be concluded that male speaker was identified with a high rate of identification probability.

Ključne riječi: prepoznavanje govornika, formantne frekvencije, protokol slušne procjene, forenzička fonetika, forenzička lingvistika

Key words: speaker identification, formant frequencies, speaker identification protocol, forensic phonetics, forensic linguistics