

Metodologija akustičkih istraživanja

Bašić, Iva; Biočina, Zdravka

Source / Izvornik: **Zbornik radova Filozofskog fakulteta u Splitu, 2020, 19 - 40**

Journal article, Published version

Rad u časopisu, Objavljena verzija rada (izdavačev PDF)

<https://doi.org/10.38003/zrffs.13.10>

Permanent link / Trajna poveznica: <https://um.nsk.hr/um:nbn:hr:172:738244>

Rights / Prava: [In copyright](#)/[Zaštićeno autorskim pravom.](#)

Download date / Datum preuzimanja: **2024-12-26**

Repository / Repozitorij:

[Repository of Faculty of humanities and social sciences](#)



UNIVERSITY OF SPLIT



DIGITALNI AKADEMSKI ARHIVI I REPOZITORIJI

UDK 001.8:534
Izvorni znanstveni rad
Primljeno 4. studenoga 2019.
DOI: 10.38003/zrffs.13.10

Iva Bašić
Filozofski fakultet Sveučilišta u Zagrebu
HR-10000 Zagreb, Ivana Lučića 3
ibasic@ffzg.hr

Zdravka Biočina
Zagrebačka škola ekonomije i managementa
HR-10000 Zagreb, Jordanovac ul. 110
zbiocina@windowslive.com

METODOLOGIJA AKUSTIČKIH SNIMANJA

Sažetak

S obzirom na to da različiti ciljevi istraživanja zahtijevaju i različite metodološke postavke, izuzetno je važno ispravno procijeniti koji su postupci nužni i dovoljni kako bi se valjano ispitala postavljena istraživačka pitanja. U ovome se radu istražio utjecaj uvjeta akustičkih snimanja na procjenu vrijednosti akustičkih parametara: formantske frekvencije vokala i osnovnoga tona glasa. Za potrebe istraživanja snimljeni su spontani govor i čitanje dvaju govornika (jedan muški i jedan ženski glas) u različitim uvjetima te su statistički testirane razlike izmjerenih akustičkih vrijednosti u odnosu na različite uvjete snimanja. Od mjera fundamentalne frekvencije najrobusnijima pokazale su se bazična vrijednost i medijan, a od formantskih frekvencija treći formant. Najznačajnije su razlike u vrijednostima u neoptimalnim udaljenostima govornika od mikrofona.

Ključne riječi: sociofonetika, akustička snimanja, metodologija, fundamentalna frekvencija, formantske frekvencije

1. Uvod

Sociofonetske studije koje proučavaju i bilježe fonetske varijacije i promjene u stvarnom vremenu oslanjaju se na akustičke analize zvučnih zapisa snimljenih u različito vrijeme, često koristeći različitu opremu i postupke snimanja. S druge strane, njihov je cilj smanjiti vanjske utjecaje na snimljeni govor kako bi otkrile postupne, potencijalno neznatne, fonetske promjene tijekom desetljeća (Rathcke i sur. 2017: 25). No, do danas se još uvijek malo zna o izvorima, vrstama i veličini tehničkih utjecaja na varijabilnost različitih akustičkih mjera.

S druge strane, u stvarnim forenzičnim slučajevima fonetičari pri analizi često raspolažu zvučnim materijalima upitne kvalitete, o čemu pišu brojni autori.¹

1 Baldwin i French 1990; Hollien 1990; Fraser 2014; 2017. itd.

U manjem broju akustičkih studija ispitani su i utjecaji različitih vrsta govornih stilova, upotrebe različite opreme za snimanje, položaja glave i sl., na vrijednosti različitih akustičkih parametara u engleskome jeziku, nizozemskome, srpskome itd. Unatoč slaboj istraženosti, Tillery i Bailey (2003) naglašavaju kako je u sociolingvistici velika potreba za ovakvim metodološkim istraživanjima. Upravo se doprinos ovoga rada ogleda u ispitivanju utjecaja uvjeta snimanja na procjenu vrijednosti formantskih frekvencija i fundamentalne frekvencije.

1.1. Metodologija sociolingvističkih, sociofonetskih i forenzičnih fonetskih istraživanja

U brojnim radovima naglašava se važnost postizanja optimalnih uvjeta snimanja kojima se pridonosi spontanosti govornika, posebice sa svrhom prikupljanja audiomaterijala organskih idioma. Optimalni uvjeti nužni su i za osiguravanje kvalitetnih zvučnih zapisa koji se kasnije mogu koristiti za daljnje akustičke analize i prepoznavanje govornika. Naime, osim što su kvalitetne snimke važne za akustičku analizu,² potrebne su i za uspješno istraživanje percepcije i prepoznavanje.³

1.1.1. Govorni materijal i stil

Prije samoga snimanja potrebno je odabrati vrstu diskursa, odnosno materijal za snimanje te odabrati govornike s obzirom na vrstu i ciljeve istraživanja. Prema Labovu (1972) temelj je sociolingvističkih istraživanja intervju, kojim se potiče spontani govor. U tradicionalnim dijalektalnim istraživanjima koriste se dijalektološki upitnici, a u novijim sociolingvističkim i sociofonetskim istraživanjima dijalekata u Hrvatskoj (Bašić 2012; Biočina 2019 itd.) primjenjuje se intervju čiji je cilj spontani govor na dijalektu.⁴ U akustičkim istraživanjima osim spontanoga govora, snima se i čitači govorni stil na listama riječi i različitim tekstovima. Primjerice, čitanje nefrikativnoga teksta uobičajeni je govorni stil za potrebe procjene glasovne kvalitete na temelju dugotrajnoga prosječnoga spektra (eng. *long term average sound spectre* – LTASS). Za određivanje formantskih frekvencija Van Heuven i Cortés (2017) preporučuju govor pod nadzorom (čitanje određenoga teksta) zbog utjecaja lingvističkoga sadržaja na vrijednosti formantata što se često pojavljuje prilikom čitanja liste riječi ili rečenica zbog nabiranja. U zadacima čitanja važno je da se materijal čita neutralno, bez uzlazne rečenične intonacije. Ako se tijekom snimanja govornici počnu koristiti spomenutom vrstom uzlazne intonacije, može im se pomoći pokretom – silaznim spuštanjem dlana, kojim se oponaša silazna rečenična intonacija. Nadalje, govornicima valja napomenuti da materijal čitaju svojom prirodnom govornom brzinom jer su istraživanja pokazala da vremenski pritisak može dovesti do snažnih koartikulacijskih utjecaja na akustičke parametre (Farnetani 2013). Za fonetsko forenzična istraživanja, uz formu intervju koja je karakteristična za stvarne forenzične slučajeve, poželjno je imati i zvučne baze govora snimljene u različitim uvjetima, poput govora u stresnim situacijama, govora

2 Fraser 2003; 2014; 2017; Rathcke i sur. 2017. itd.

3 Kösteret i sur. 1995; Köster i Schiller 1997; Rose 2002; Hollien 2002; Jessen 2008.

4 Zanimljivo je da je i Šimunović (2011: 31), premda dijalektolog, također smatrao da je korisno proučavati spontane realizacije dijalektalnih govora, jer iako spontani govor u svakodnevnoj komunikaciji pokazuje manje arhaizama, važan je za istraživanja jezičnih dodira, među dijalektima i između dijalekta i standarda.

jačega intenziteta, prikrivenim glasom itd.

Vermeulen i Cambier-Langeveld (2017) napominju da je isto govorno ponašanje odnosno govorni stil optimalan za usporedbu govornika i akustičkih vrijednosti. Naime, Boyd i sur. (2015) usporedili su podatke dobivene klasičnim sociolingvističkim metodama intervjua s podacima dobivenim samo-snimanjem (eng. *self-recordings*) te podatke dobivene različitim metodama poticanja spontanoga govora, koje se uobičajeno koriste u lingvističkim laboratorijima. Više fonetskih promjena uočeno je u govoru bez prisutnosti intervjuišta, a utvrđeno je i nekoliko značajnih razlika između govora dobivenog intervjuištom i zadacima u laboratoriju (Boyd i sur. 2015). Stoga autori savjetuju da se sociolingvistika i posebice sociofonetika trebaju češće koristiti samo-snimanjem kao načinom prikupljanja podataka jer su rezultati pokazali da tako dobiven govor često rezultira izvornijom produkcijom od onoga dobivenog intervjuištom. Iako je to pogodno za neke lingvističke analize, za akustička mjerenja vrlo je važno da se snimanje provodi: u tihoj sobi, da se kontrolira položaj glave u odnosu na mikrofoni i sl., zbog utjecaja na akustičke mjere. Nadalje, i sami su autori svjesni da takva vrsta snimanja predstavlja značajan metodološki izazov zato što zahtijeva dodatnu pomoć sudionika koji moraju biti i uvježbani za snimanje zvuka. U radu problematiziraju i zadatke čitanja te zaključuju da često nisu idealan izbor na terenskim istraživanjima jer neki govornici mogu biti nepismeni, polupismeni ili slabijega vida (Boyd i sur. 2015).

Vrsta govornoga stila važna je i stoga što vrijednosti nekih akustičkih mjera variraju u različitim diskursima. Traunmüller i Eriksson (1995) utvrđuju da se standardna devijacija (S. D.) F_0 smanjuje s povećavanjem živosti diskursa, a u većini življih diskursa (kao npr. u glumačkome govoru) žene imaju veću S. D. F_0 od muškaraca, dok je u manje živim tipovima diskursa niža nego kod muškaraca (kao npr. u čitanju tekstova za klinička ispitivanja glasa). I istraživanja s češkim (Skarnitzl i Vaňková 2016; 2017) te hrvatskim i srpskim muškim govornicima (Varošaneć-Škarić i sur. 2016; Biočina i sur. 2016) pokazuju niže vrijednosti prosječne F_0 za spontani govor nego čitanje. Navedene razlike treba posebice uzeti u obzir kad se uspoređuju vrijednosti F_0 dobivene u istraživanjima s različitim metodološkim postavkama.

1.1.2. Izbor govornika

Pri odabiru govornika važan je kriterij izvornosti govornika u odnosu na idiom koji se istražuje te se on smatra jednim od temeljnih kriterija za odabir govornika (Ladefoged 2003), posebice u sociolingvistici. S druge strane, tradicionalna dijalektološka istraživanja temelje se na idealnim govornicima, pobliže, starijim starijima koji nisu geografski mobilni. Vrlo rijetko, a i tek odnedavno, u dijalektološke opise u Hrvatskoj uključuju se i mlađi govornici (vidi npr. Galović 2012; Šprljan 2015). Još jedna razlika među navedenim područjima odnosi se na broj ispitanika. Sociolingvistički opis uobičajeno se temelji na većem broju govornika ($N = 20 - 30$), dok se dijalektološki opisi temelje na tek nekoliko govornika. U akustičkim fonetskim ispitivanjima uglavnom je važno da govornici imaju zdrav glas i uredan govorno-jezični status, posebice kad se mjere referentne vrijednosti.⁵ Podatke o govornicima (dob, spol, obrazovanje, podrijetlo roditelja,

⁵ Osim u ispitivanjima patoloških glasova.

mjesto rođenja, mjesto najdužeg boravka te datum snimanja) poželjno je zabilježiti sociodemografskim upitnikom. U dijalektološkim ispitivanjima treba pripaziti ako je upitnik na standardnome jeziku da ga se ponudi govornicima na ispunjavanje nakon snimanja ili prije ako je preveden na dijalekt.

1.1.3. Govorni i okogovorni uvjeti snimanja

Nakon izbora materijala i govornika, može se započeti sa snimanjem. Labov (1984) nudi dvije upute za uspješan intervju: 1. ne treba inzistirati na zadanom redoslijedu tema i 2. ispitivač treba aktivno sudjelovati u intervjuu, to jest treba biti zainteresiran za sadržaj (Starčević 2016). Naime, Labov (1984) ističe kako sociolingvistički intervju nije uspješan ako govornik samo odgovara na zadana pitanja i ne skreće s teme jer upravo skretanja s teme mogu biti najvažniji dijelovi za istraživanje. Nadalje, tijekom intervjuu poželjno je razgovarati na varijetetu govornika jer se tako smanjuje formalnost situacije i time ostvaruje spontanija komunikacija s govornicima (Hraste 1960; Bašić 2012; Starčević 2016). Sociolingvistička snimanja poželjno je napraviti u domovima govornika jer se smatra da je važno istraživanja provoditi u poznatoj i svakodnevnoj okolini za govornika (Bower 2010: 341 prema Starčević 2016: 9), a prije početka intervjuu može se uputiti govornike da govore kao da su u društvu obitelji ili prijatelja. Snimanja je važno provoditi u tihim prostorijama bez zamjetne izvanjske buke, posebice u akustičkim snimanjima. Za neke akustičke mjere snimanja je potrebno provoditi u visoko kontroliranim uvjetima, primjerice u akustičkom studiju. Visokokvalitetna oprema i prostorije snižene razine buke nužne su kako bi se osigurale kvalitetne snimke koje će se koristiti za akustičku i fonetsku analizu te prepoznavanje. No, u stvarnosti su audiosnimke kojima raspolaže forenzični fonetičar vrlo često u nekome dijelu signala maskirane pozadinskom bukom, različitim vrstama distorzije signala, interferirajućim zvukovima itd. (Maher 2009). Često se kao glavni problem osiguravanja kvalitetne snimke navodi eliminiranje pozadinske buke (Ladefoged 2003). Nolan (2005) navodi da je uglavnom riječ o buci javnoga prometa, govoru drugih sudionika razgovora ili pojedinaca koji se nalaze u blizini ciljanoga govornika. Osim spomenutih izvora buke u snimačkom radu autorica zamijećeni su još neki poput buke govornika i različitih uređaja. Pod potonjom podrazumijevamo buku koju proizvode klimatizacijski uređaji, ventilatori, mobilni uređaji, telefoni, računala te javni promet. Buka koju proizvode govornici može se umanjiti tako da im se savjetuje da ne odjenu šuštavu odjeću i nakit, da im se za vrijeme snimanja oduzmu kemijske olovke i papiri i sl. Istraživanja pokazuju da su u snimkama s puno buke povišene vrijednosti prvoga formanta svih vokala, a učinak je najznačajniji na otvorenim vokalima (Rathcke i sur. 2017).

Nadalje, dobro je poznato da intenzitetske varijacije utječu na artikulacijsku i glotalnu konfiguraciju (Huber i sur. 1999). Kod glasnoga govora dolazi do kompenzacijskih pokreta (ponajprije se pokretima jezika kompenziraju veliki pokreti čeljusti, čime se postiže prikladna visina u izgovoru pojedinog vokala) (Schulman 1989), koji mogu utjecati na vrijednosti formanta (Huber i sur. 1999), a katkad i na fundamentalnu frekvenciju (Jessen i sur. 2003). Stoga valja govornike uputiti da budu umjereno glasni za vrijeme snimanja. Ako su govornici tiši ili glasnjiji od razine umjereno glasnoga govora i nakon uputa, na snimaču se može prilagoditi dinamički/intenzitetski raspon prema karakteristikama svakoga govornika. Široki intenzitetski

rasponi uglavnom su češći za vrijeme spontanoga govora nego na primjer čitanja. Bolja kvaliteta akustičke opreme preporučuje se za sva akustička istraživanja, posebice u području forenzične fonetike, zbog toga što će snimka u većoj mjeri odgovarati originalnom govoru (Hayward 2000). Švec i Granqvist (2010) navode preporuke o poželjnim karakteristikama mikrofona, među kojima naglašavaju da frekvencijski raspon mikrofona treba biti ravan (varijacije manje od 2 dB) te unutar ispitivanoga frekvencijskoga područja (najniže i najviše očekivane osnovne frekvencije glasa). Nadalje, ističu kako omnidirekionalni mikrofoni obično snime više pozadinske buke nego oni direkcionalno usmjereni, a verzije s malim vrhom snimaju posebno bučne snimke (Švec i Granqvist 2010). Osim vrste mikrofona, i različita udaljenost od mikrofona utječe na akustičke mjere pri analizi vokala te je tijekom snimanja potrebno paziti na to da govornik ne mijenja često položaj glave i udaljenost u odnosu na mikrofona (Vermeulen 2009). U istraživanju Jovanovića i Jovičića (2011) u kojem su analizirani formanti srpskoga jezika u trima govornim situacijama: šaptu, prilikom držanja mobitela između obraza i ramena te za vrijeme prikriivanja uređaja rukom, zabilježene su značajne promjene u vrijednostima prvih dvaju formanata.

U nekim istraživanjima korpus je prikupljen u različitim uvjetima te su govornici snimani različitom opremom i u različitim studijskim uvjetima. Upravo je zbog te metodološke heterogenosti bilo otežano odrediti jesu li rezultati odraz lingvističke univerzalnosti ili su pak uzrokovane različitim metodološkim postavkama (Byrne i sur. 1994). Preporučuje se zadržavanje iste opreme i uvjeta snimanja tijekom čitavoga istraživanja te fotografiranje postavljene opreme za snimanje (Rathcke i sur. 2017). Iako je većina suvremenih sociolingvističkih istraživanja postavljena na metodološkim postulatima pionira sociolingvistike, u istraživanjima metodologije akustičkih snimanja postoje brojna neslaganja, a i broj radova te tematike vrlo je ograničen. Posljedično su izravni utjecaji uvjeta snimanja govornika na rezultate akustičkih istraživanja nedovoljno ispitani i opisani (npr. studijski i nestudijski uvjeti snimanja, korištenje različite opreme za audiosnimanje, otpornost različitih akustičkih parametara i mjera itd.). Također, valja napomenuti da utjecaji nekih uvjeta snimanja do sada nisu ispitani, primjerice, snimanje govornika u različito doba dana.

2. Tradicionalni i suvremeni pristupi u analizi fundamentalne frekvencije i formanata

Fundamentaln frekvencija (F_0) već se dugo koristi u forenzičnoj fonetici kao robusna mjera u različitim uvjetima snimanja⁶ te se pokazala kao najučinkovitiji kriterij za razlikovanje govornika (Nolan 1983: 129). F_0 ispunjava i tri glavna uvjeta koja su, prema Nolanu (1983), važna u forenzičnom prepoznavanju govornika, a to su robusnost, mjerljivost i dostupnost. Također, neosjetljiva je na lingvistički sadržaj zvučnoga odsječka (Kinoshita i sur. 2009). Iako je pitanje dužine zvučnoga zapisa vrlo važno u forenzičnoj fonetici, ne postoji dogovor među stručnjacima koliko je vremena potrebno da se F_0 stabilizira. Vrijednosti variraju od jedne (Nolan 1983) do dvije minute (Baldwin i French 1990), a neka novija istraživanja bilježe stabilizaciju i unutar 30 sekundi od početka snimke (Arantes i Eriksson 2014; Arantes i sur. 2017). Novija saznanja upućuju na brže vrijeme stabilizacije određenih mjera F_0 , neovisno

6 Nolan 1983; Braun 1995; Traunmüller i Eriksson 1995; Rose 2002; Lindh 2006 itd.

o vrsti diskursa (Arantes i Eriksson 2014; Arantes i sur. 2017). Prema potonjima, bazična se vrijednost (F_b) stabilizira u značajno kraćem vremenu (oko 5 s) od prosječne F_0 i medijana F_0 (oko 10 s), u svim istraživanim jezicima.

F_0 se najčešće računa računalnim programima napravljenim za tu svrhu, stoga je relativno jednostavna za mjerenje.⁷ Gold i French (2011) izračunali su učestalost korištenja pojedinih mjera F_0 u svijetu, a njihovi rezultati pokazuju kako se 94 % stručnjaka iz područja forenzične fonetike koristi prosječnom F_0 , a samo 25 % mjerom F_b . To iznenađuje s obzirom na to da je bazična vrijednost (F_b) robusnija mjera od F_0 te bolje predstavlja individualne karakteristike govornika, što su već potvrdila brojna istraživanja (Traunmüller 1994; Traunmüller i Eriksson 1995; Lindh 2006; Arantes i Eriksson 2014; Varošaneć-Škarić i sur. 2016; Biočina i sur. 2016; Varošaneć-Škarić i sur. 2017; da Silva i sur. 2016; Arantes i sur. 2017; Biočina 2019). Nadalje, F_b je otpornija mjera i s obzirom na čimbenike kao što su govornikove emocije i vokalni napor (Arantes i Eriksson 2014). Računa se iz prosječne F_0 ili medijana F_0 , a većina autora upozorava da je precizniji izračun iz medijana jer je to robusnija mjera na čije vrijednosti manje utječu oktavna skakanja za razliku od prosječne F_0 (Lindh 2006; Arantes i Eriksson 2014; da Silva i sur. 2016; Arantes i sur. 2017). Bazična vrijednost na temelju medijana pokazala se robusnijom i u mjerenjima u spontanome dijalektalnom govoru u odnosu ne neke tradicionalne mjere poput minimalne i maksimalne F_0 (Biočina 2019).

U današnje vrijeme postoje brojne metode za procjenu formantskih frekvencija, a potreban alat dostupan je i u nekoliko besplatnih softvera za akustičku analizu govora (Bašić 2018). Prema Harrisonu (2013) najfrekventnije se koristi Praat (Boersma i Weenink 2015), uz koji se istraživači često koriste i WaveSurferom (Sjölander i Beskow 2017). Prema Harrisonu (2013) tri su najčešće korištene metode u mjerenju formantskih frekvencija: frekvencijski spektar, spektrogrami te LPC metoda (eng. *linear predictive coding*). Svaka od spomenutih metoda ima svoje prednosti i nedostatke, a uloga istraživača je da s obzirom na cilj rada i snimke kojima raspolaže, pronađe najbolju metodu koja će biti prikladna i čija će se primjena moći opravdati.

Osim spomenutih metoda, vrlo se učestalo koriste statične metode mjerenja formantata te dugotrajne ili dinamičke. Za razliku od statičnih metoda, koje mjere formantske frekvencije na kraćem govornom materijalu (najčešće izoliranim riječima) i u jednoj do tri točke stabilnoga dijela vokala, dugotrajna formantska distribucija (LTFD za eng. *long-term formant distribution*) daje nam jasne informacije o prosječnome ponašanju, tj. kretanju svakoga formanta kod pojedinoga govornika na duljim snimkama. LTFD metodom nastoje se zabilježiti opće (iz)govorne navike pojedinca koje se zadržavaju tijekom cjelokupnoga trajanja govornika, primjerice labijaliziranje ili palataliziranje (Bašić 2018).

3. Istraživačka pitanja i hipoteze

U ovome je radu postavljeno nekoliko istraživačkih pitanja: hoće li različit a) termin snimanja, b) uvjeti snimanja i c) udaljenost te položaj glave u odnosu na mikrofona, utjecati na vrijednosti formantskih frekvencija te fundamentalne frekvencije? Osim

7 Možda upravo to pridonosi činjenici da je prosječna fundamentalna frekvencija jedna od rijetkih mjera za koju postoji statistika za populaciju (Jessen 2008).

utjecaja spomenutih uvjeta na fundamentalnu frekvenciju i formantske frekvencije, u radu su ispitane i razlike u dugotrajnome prosječnome spektru (LTASS). S obzirom na dosadašnje radove u području forenzične i akustičke fonetike te neistraženosti pojedinih okogovornih uvjeta snimanja, postavljeno je pet hipoteza:

H1: neće biti značajne razlike u vrijednostima formanata između snimanja u jutarnjim i popodnevnim satima

H2: značajno će se razlikovati vrijednosti fundamentalne frekvencije između snimanja u jutarnjim i popodnevnim satima

H3: neće biti značajne razlike u vrijednostima formanata i fundamentalne frekvencije između snimanja u studijskim uvjetima i u tihoj sobi sa sniženom razinom buke

H4: različita udaljenost i položaj glave u odnosu na mikروفon značajno će utjecati na vrijednosti formanata i fundamentalne frekvencije te

H5: svi će neoptimalni uvjeti snimanja u LTASS prikazima biti akustički osiromašeni u određenim dijelovima spektra u odnosu na optimalne uvjete.

4. Metodologija

4.1. Govornici

Za potrebe rada snimljena su dva izvorna govornika hrvatskoga jezika, jedan muški govornik i jedna ženska govornica, urednoga govorno-jezičnog statusa. S obzirom na prethodna istraživanja koja pokazuju varijabilnost akustičkih mjera među spolovima, u analizi su zastupljena oba spola. Izabrani govornici u dobi su od 30 (Ž) i 40 (M) godina. Izgovor obaju govornika procijenjen je kao neregionalan.

4.2. Govorni materijal i uvjeti snimanja

Prije početka snimanja govornicima su dane jasne i jednostavne upute o tijeku snimanja te o njihovim zadacima. Sav govorni materijal dobili su na uvid i na jedno probno čitanje prije početka snimanja. Govorni materijal sastojao se od triju dijelova: u prvome dijelu govornici su čitali dvosložne riječi (ukupno 15 riječi za vokale [i], [a] i [u]), nakon čega su čitali nefrikativni tekst za procjenu timbralnih karakteristika glasa. U posljednjem dijelu snimanja prikupljen je spontani govor u trajanju od najmanje četiri minute. Sva tri govorna materijala snimana su u svim uvjetima: tihoj sobi, studiju, u jutarnjem (jedan sat nakon buđenja) i popodnevnom terminu (oko 14 h, odnosno otprilike šest sati nakon buđenja), preblizu mikروفonu (do pet cm udaljenost od mikrofona), predaleko od mikrofona (udaljenost od 30 cm od mikrofona) te s glavom u položaju od 90° u odnosu na mikروفon.

Optimalni uvjeti snimanja, koji su bili referentni u analizi podataka i usporedbi različitih uvjeta snimanja, postignuti su snimanjem u studijskim uvjetima u popodnevnim satima, na optimalnoj udaljenosti od 10-ak cm od mikrofona te optimalnom položaju glave u odnosu na mikروفon (0°, odnosno direktna usmjerenost usana govornika na mikروفon). Obama je govornicima rečeno da materijal čitaju neutralno te da ako pogriješe, ponove dio materijala u kojem su pogriješili. Također, u slučaju zamora čitanjem, imaju mogućnost stanke u snimanju kada im bude

potrebna. Za potrebe prikupljanja spontanoga govora, govornicima je dana uputa da govore na bilo koju temu što duže, a u slučaju da nije prikupljeno dovoljno materijala, predložene su im teme poput opisa njihova uobičajenoga radnog dana, hobija, omiljene knjige ili filma, neke igre iz djetinjstva, grada u kojem žive i sl.

4.3. Akustičke metode

U radu su korištene četiri akustičke metode: formantska analiza, tradicionalne i suvremene metode analiziranja fundamentalne frekvencije te dugotrajni prosječni spektar. Formantska analiza provedena je pomoću programa za akustičku obradu zvuka Praat (Boersma i Weenink 2015), u kojem su mjerene vrijednosti prvih triju formanata (F1, F2 i F3) u trima točkama središnjega stabilnog dijela vokala, gdje su formantske konture gotovo paralelno položene. Pri procjeni formantskih frekvencijskih vrijednosti posebna se pozornost usmjerila na formantsku integraciju (spajanje (najčešće) dvaju formanata u jedan) kod ženske govornice te kod obaju govornika pri izgovoru vokala [u] zbog frekvencijske blizine F1 i F2.

Za razliku od formanata, fundamentalna frekvencija analizirana je automatski, uz pomoć skripte iz forenzične fonetike (Harrison 2013) na montiranim uzorcima spontanoga govora u trajanju od najmanje dvije minute. U prvome mjerenju određene su granične frekvencije od 65 Hz (donja granična frekvencija) i 350 Hz (gornja granična frekvencija) za ženski glas, prema preporukama Lindha (2006), Varošanec-Škarić i sur. (2016) te Biočina i sur. (2016). Međutim, zbog laringalizacije i oktavnih skakanja, donja je granična frekvencija povišena za 25 Hz (na 90 Hz), čime su dobiveni precizniji rezultati.

Od mjera fundamentalne frekvencije, korištene su sljedeće: prosječna vrijednost, minimalna i maksimalna vrijednost, medijan, bazična i alternativna bazična vrijednost. Mogli bismo reći da su prosječna, minimalna i maksimalna vrijednost tradicionalne mjere fundamentalne frekvencije koje se najčešće koriste, dok su preostale mjere suvremenije i koriste se u posljednjih 15-ak godina, ponajviše u području forenzične fonetike.

U radu su također urađeni LTASS prikazi za svakoga govornika u svim uvjetima snimanja. S LTASS prikazom optimalnoga uvjeta snimanja (u studijskim uvjetima u popodnevnim satima, na optimalnoj udaljenosti od mikrofona te optimalnom položaju glave u odnosu na mikrofona) uspoređen je svaki LTASS prikaz iz neoptimalnoga uvjeta snimanja, na temelju čega su utvrđene timbralne razlike kod svakoga govornika zasebno.

5. Rezultati

5.1. Rezultati formantske analize

Analizom formantskih frekvencija u različitim uvjetima snimanja nastojala se ispitati robusnost formanata, tj. odrediti stupanj varijabilnosti pojedinoga formanta (F1-F3). U tablici 1. prikazani su uvjeti snimanja u paru (primjerice jutro/popodne, optimalna/premala udaljenost od mikrofona itd.) između kojih je potvrđena značajnost razlike u frekvencijskim vrijednostima pojedinoga formanta.

Tablica 1. Prikaz statistički značajnih razlika u formantima između različitih uvjeta snimanja

UVJETI	Ž	M
jutro / popodne	F2	F1
studio / tiha soba	F2 i F3	F1 i F2
optimalna / premala udaljenost	F2 i F3	F3
optimalna / prevelika udaljenost	F1	-
optimalan / neoptimalan kut (90°)	F2	F1

Rezultati analize pokazali su statistički značajne razlike na razini $p < 0,05$, u najmanje jednom formantu u ispitivanim parovima uvjeta snimanja (osim kod muškoga govornika između optimalne i prevelike udaljenosti od mikrofona). Sagledamo li rezultate na razini spolova, možemo reći da je kod muškoga govornika najvarijabilniji (najmanje robustan) prvi formant te da su drugi i treći formant podjednako stabilniji od F1 (tri statistički značajne razlike za F1 i po jedna razlika za F2 i F3). S druge pak strane, za žensku je govornicu najrobustniji prvi formant (samo jedna statistički značajna razlika), nešto manje robustan F3 (dvije razlike), dok je drugi formant najvarijabilniji (čak četiri statistički značajne razlike – u gotovo svim ispitanim parovima uvjeta). Stupanj varijabilnosti prvih triju formanta ispitali su Jovičić i sur. (2015) analizirajući utjecaj različitih položaja u kojima se koristi mobilni uređaj na spektralne karakteristike govornoga signala. Rezultati provedenoga istraživanja pokazali su najveće distorzije kod F1 (30 %), nešto manje kod F2 (15 %) i najmanje kod F3 (5 %). Drugim riječima, kao najmanje stabilan i robustan pokazao se F1, a kao najrobustniji F3, isto kao i kod ženske govornice u ovome radu.

O nestabilnosti prvog formanta pisali su i Harrison (2013), Zhang i sur. (2013), Byrne i Foulkes (2004) te Künzel (2001). Potonji je autor usporedbom frekvencijskih vrijednosti formanta iz snimke snimljene u studijskim uvjetima i snimke telekomunikacijskoga prijenosa utvrdio da se vrijednosti prvoga formanta statistički značajno povisuju za čak 29 %. U forenzičnoj fonetici ističe se da se vrijednosti formanta prikupljene iz različitih vrsta snimki ne bi smjele uspoređivati, posebice ako je riječ o snimkama visoke kvalitete zvuka i snimkama prikupljenima u telefonskoj i mobilnoj transmisiji (Nolan 2005). Problematika takvih analiza ogleda se u akustičkim artefaktima, tj. frekvencijskim izobličenjima propuštenoga zvuka (posebice kod formanta čije se frekvencijske vrijednosti nalaze blizu graničnih frekvencija propuštanja: oko 400 Hz i oko 3 400 Hz). Upravo iz tog razloga drugi formant pokazao se kao najrobustniji na uvjet prijenosa zvuka tzv. *telex* mrežom.

U ovome se radu, na razini obaju analiziranih govornika, treći formant istaknuo kao najrobustniji s najmanjim brojem statistički značajnih razlika u ispitivanim parovima uvjeta snimanja, kao i kod Jovičića i sur. (2015). U forenzičnoj fonetici navodi se da F3 ima najveću diskriminatornu ulogu među formantima, u usporedbi više govornika. Razlog tomu vjerojatno je u činjenici da nam F3 istodobno daje lingvističke informacije (o vokalu koji je izgovoren) i ekstralingvističke informacije (o timbralnim karakteristikama govornika). O pripadnosti trećega formanta nižim formantima (uz F1 i F2) koji nam daju informacije o vokalskoj kvaliteti i o pripadnosti F3 višim formantima (F4 i F5) koji nam

daju informacije o vokalnoj kvaliteti (timbru) postoje brojna neslaganja među autorima.

5.2. Rezultati mjerenja fundamentalne frekvencije

Osim rezultata formantske analize za potrebe rada analizirana je i fundamentalna frekvencija. Izračunane su prosječna vrijednost, minimalna i maksimalna vrijednost, medijan, bazična te alternativna bazična vrijednost (vidi tablicu 2.). Podebljanim slovima otisnute su najviše i najniže vrijednosti, a podvučene su referentne vrijednosti F_0 dobivene u optimalnim uvjetima snimanja.

Tablica 2. Prikaz vrijednosti fundamentalne frekvencije u različitim uvjetima snimanja kod ženske govornice

SPONTANI GOVOR	PROSJEČNA F_0	MEDIJAN F_0	MIN F_0	MAKS F_0	ALT F_B	F_B
<i>tiha soba</i>	187,47	184,49	89,93	329,42	161,24	141,19
<i>ujutro</i>	184,32	179,95	89,95	337,75	156,33	141,46
<i>popodne</i>	<u>186,38</u>	<u>181,68</u>	<u>89,92</u>	<u>334,83</u>	<u>158,75</u>	<u>140,75</u>
<i>preblizu</i>	181,79	178	89,91	349,59	156,36	137,09
<i>predaleko</i>	196,62	193,25	89,94	349,6	162,55	144,08
<i>položaj glave (90°)</i>	186,25	184,27	89,92	339,25	159,15	139,27

Na temelju akustičkih podataka prikazanih u tablici 2. možemo izvući nekoliko zanimljivih rezultata, koji će se u nastavku teksta protumačiti:

- najniže vrijednosti F_0 u svim parametrima (osim u maksimalnoj i alternativnoj bazičnoj) utvrđene su prilikom snimanja na premaloj udaljenosti od mikrofona
- najviše vrijednosti F_0 u svim parametrima (osim u minimalnoj) utvrđene su prilikom snimanja na prevelikoj udaljenosti od mikrofona
- utvrđene minimalne i maksimalne vrijednosti F_0 vrlo su blizu graničnih frekvencija primijenjene skripte (90 i 350 Hz)
- najveća varijabilnost F_0 između analiziranih uvjeta snimanja utvrđena je kod mjere maksimalne vrijednosti F_0
- najmanja varijabilnost F_0 između analiziranih uvjeta snimanja utvrđena je kod mjere minimalne vrijednosti F_0 .

S obzirom na to da su u uvjetu snimanja na premaloj udaljenosti od mikrofona utvrđene najniže vrijednosti fundamentalne frekvencije, možemo reći da snimanje u tom uvjetu dovodi do snižavanja vrijednosti F_0 u gotovo svim parametrima (osim u mjeri maksimalne F_0 gdje je utvrđena vrijednost od 349,59 Hz i u alternativnoj bazičnoj mjeri gdje je utvrđena vrijednost od 156,36 Hz). Ako usporedimo razliku u vrijednostima F_0 u svim analiziranim parametrima između optimalne i prevelike udaljenosti te između optimalne i premale udaljenosti, možemo zaključiti da su veće

razlike kod prevelike udaljenosti od mikrofona. Rezultati su neočekivano pokazali da neoptimalan kut od 90° u odnosu na mikrofona ne donosi značajne promjene u većini mjera fundamentalne frekvencije (gotovo podjednake vrijednosti za parametre prosječne i minimalne F_0 , bazične i alternativne bazične vrijednosti te nešto više vrijednosti za medijan (2,5 Hz) i maksimalnu vrijednost (4,4 Hz)).

O nepouzdanosti minimalnih i maksimalnih vrijednosti F_0 pisali su brojni autori, posebice u području akustičke i forenzične fonetike (Lindh 2006; Arantes i Eriksson 2014; Arantes, Eriksson i Gutzeit 2017; Skarnitzl i Vaňková 2017; Biočina 2019). Problematika njihova korištenja ogleda se u tome da program koji se koristi, odnosno skripta koja se primjenjuje, najčešće u spektru pronalaze (ne)opravdane tragove fundamentalne frekvencije u blizini graničnih frekvencija (u ovome radu za ženski glas primijenjen je raspon od 90 Hz do 350 Hz). Tako će se primjerice registrirati glotalna prženja i laringalizacije na kraju rečenice kao minimalne vrijednosti F_0 , kao što možemo vidjeti iz tablice 2. (minimalne vrijednosti fundamentalne frekvencije iznose 89 Hz u svim analiziranim uvjetima). Izračunamo li standardnu devijaciju za minimalnu F_0 u različitim uvjetima snimanja (0,04), mogli bismo pogrešno zaključiti da je taj parametar fundamentalne frekvencije vrlo robusan i otporan na uvjete snimanja, u odnosu na druge parametre. Međutim, ne-varijabilnost ovoga parametra uzrokovana je očitavanjem neopravdanih tragova fundamentalne frekvencije u nižim dijelovima spektra, blizu granične frekvencije očitavanja F_0 određene skriptom.

Također, maksimalne vrijednosti fundamentalne frekvencije zabilježiti će snažnija oktavna skakanja kao rezultat F_0 , što je posebno istaknuto kod premale i prevelike udaljenosti od mikrofona. Nepouzdanost maksimalne vrijednosti F_0 odražava se i u stupnju varijabilnosti između analiziranih uvjeta snimanja, koja je najviša kod mjere maksimalne F_0 (S. D. = 8,1).

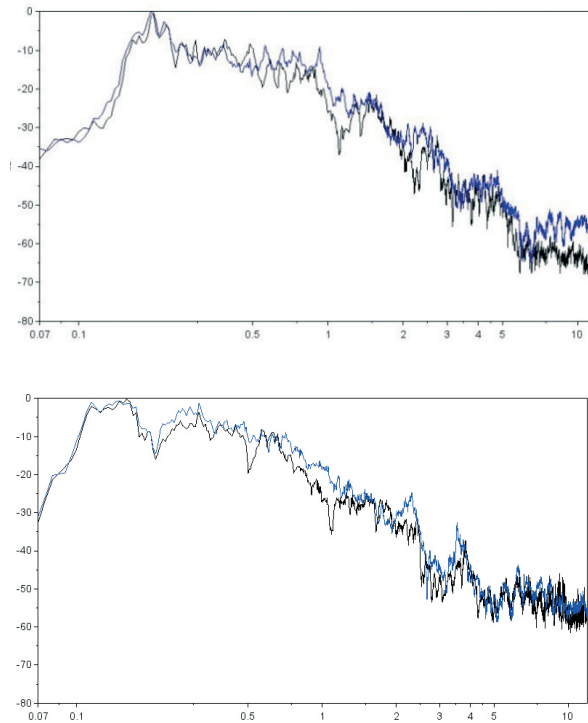
Naposljetku, valjalo bi istaknuti da su se parametri bazične i alternativne bazične F_0 pokazali kao najrobusniji i najotporniji na analizirane uvjete snimanja. Stabilnost tih parametara ogleda se u vrlo maloj varijabilnosti (S. D. = 2,52 za Alt F_b i S.D. = 2,34 za F_b) te u frekvencijskoj razlici između različitih uvjeta snimanja (6,22 Hz za Alt F_b i 6,99 za F_b). O pouzdanosti i robusnosti ovih parametara na različite uvjete snimanja, vokalni napor i vrstu govornih stilova pisali su brojni autori.⁸ Osim toga, rezultati su pokazali da su mjere prosječne F_0 i medijana F_0 dvostruko nestabilnije (S. D. = 5,06 za prosječnu vrijednost i S. D. = 5,34 za medijan; frekvencijska udaljenost od 14,83 Hz za prosječnu vrijednost i 15,25 za medijan u usporedbi različitih uvjeta snimanja).

5.3. Rezultati dugotrajnih prosječnih spektara

U posljednjem dijelu rezultata ovoga istraživanja obraditi će se dugotrajni prosječni spektri (LTASS-i) kod muškoga i ženskoga govornika te će se utvrditi razlike u timbralnim karakteristikama među ispitivanim uvjetima snimanja.

Na slici 1. prikazana je razlika između optimalnih uvjeta snimanja (plava linija) i neoptimalnoga položaja glave, pod kutom od 90° u odnosu na mikrofona (crna linija).

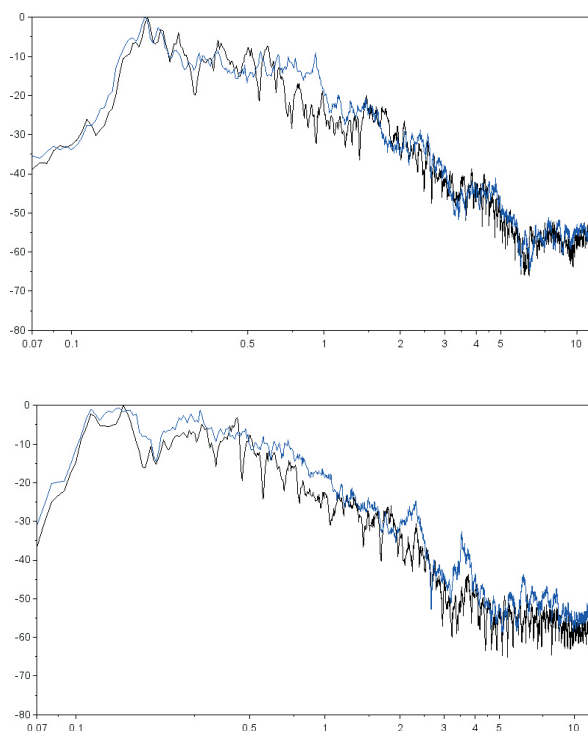
⁸ Traummüller i Eriksson 1995; Lindh 2006; Arantes i Eriksson 2014 itd.



Slika 1. LTASS prikazi ženskoga (lijevo) i muškoga glasa (desno) u optimalnim uvjetima snimanja i s neoptimalnim položajem glave

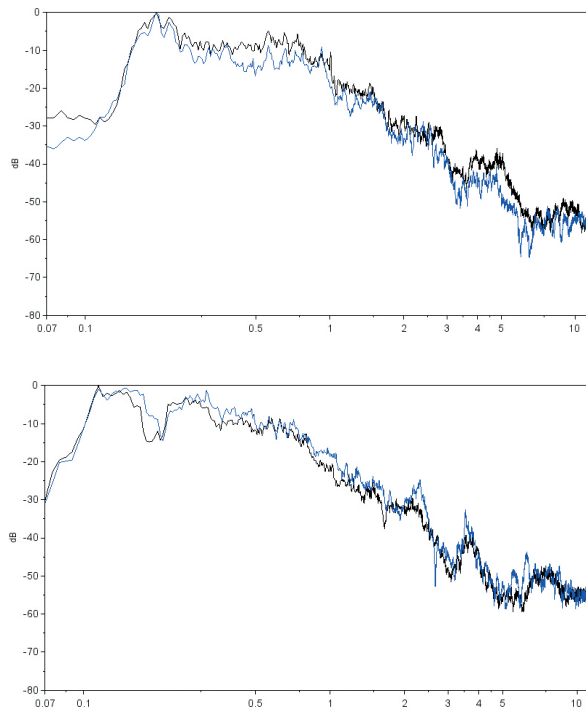
S obzirom na to da je na LTASS prikazima glas iste ženske osobe lijevo i iste muške osobe desno, u različitim uvjetima snimanja, može se očekivati da prikazane razlike nisu rezultat različitoga timbra, već neoptimalnih uvjeta u kojima su govornici snimljeni. Na LTASS prikazu lijevo možemo zamijetiti da je neoptimalan položaj glave doveo do manjega gubitka zvučne energije u području punoće (400 – 800 Hz) i srednje velikoga gubitka u području zvonkosti (800 – 2 000 Hz), okruglosti (2 000 – 2 500 Hz) te pucketavosti (5 000 – 10 000 Hz). Kod muškoga glasa akustički je blago *osiromašeno* područje voluminoznosti (do 300 Hz), punoće (400 – 800 Hz) i okruglosti (2 000 – 2 500 Hz) te u nešto većoj mjeri u području zvonkosti (800 – 2 000 Hz).

Na slici 2. prikazana je razlika između optimalnih uvjeta snimanja (plava linija) i neoptimalne udaljenosti od mikrofona (predaleko) (crna linija).



Slika 2. LTASS prikazi ženskoga (lijevo) i muškoga glasa (desno) u optimalnim uvjetima snimanja i na neoptimalnoj udaljenosti od mikrofona (predaleko)

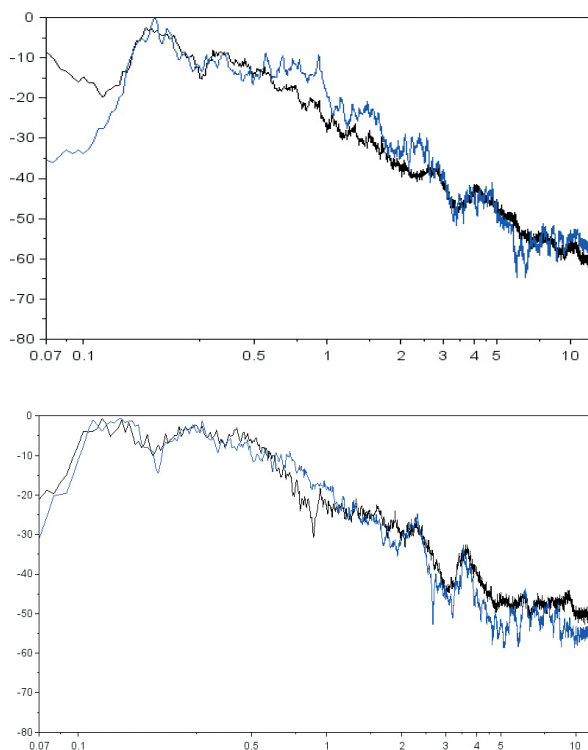
Usporedimo li LTASS prikaze za žensku osobu snimljenu u optimalnim uvjetima i predaleko od mikrofona, možemo zamijetiti manje razlike u raspodjeli zvučne energije. Naime, neoptimalna, prevelika udaljenost od mikrofona dovela je do blagogoga pada u području punoće te srednjega pada zvučne energije u području zvonkosti. Kod muškoga govornika vidljivo je više padova u zvučnome spektru negoli kod ženskoga glasa: u području voluminoznosti i punoće blagogoga intenziteta, u području zvonkosti, okruglosti te više blještavosti srednjega intenziteta. Kao što je već prethodno u radu napomenuto, ispitan je i utjecaj premale udaljenosti govornika od mikrofona (do otprilike pet cm) na različite akustičke parametre. Na slici 3. prikazani su LTASS-i muškoga i ženskoga glasa u opisanome neoptimalnom uvjetu snimanja.



Slika 3. LTASS prikazi ženskoga (lijevo) i muškoga glasa (desno) u optimalnim uvjetima snimanja i na neoptimalnoj udaljenosti od mikrofona (preblizu)

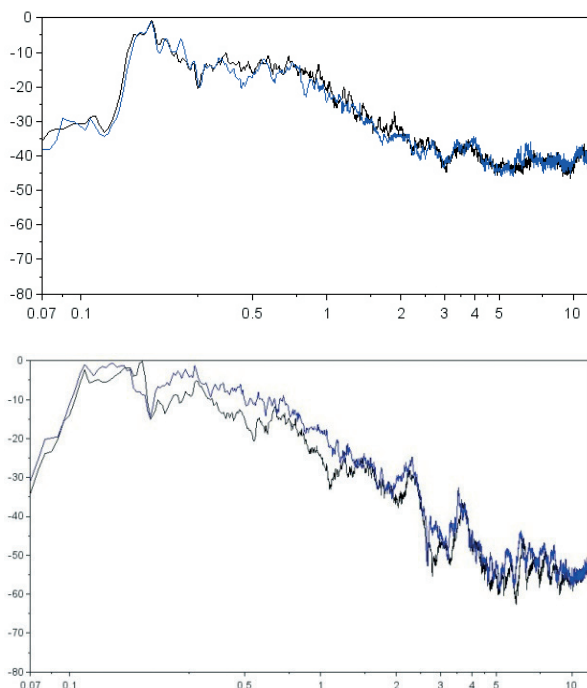
Na LTASS prikazima (slika 3. lijevo) na optimalnoj i premaloj udaljenosti od mikrofona vidljiv je rast zvučne energije u nižem području spektra, do 150 Hz (područje voluminoznosti i punoće) te nešto veći rast u višim dijelovima spektra u užem području pucketavosti (od 5 do 8 kHz). S druge pak strane, kod muškoga govornika zamijećena su manja odstupanja, i to u smjeru pada zvučne energije. Blaži su padovi zvučne energije zabilježeni u područjima voluminoznosti, zvonkosti te okruglosti.

Na slici 4. prikazana je razlika u dvama uvjetima snimanja, u akustičkome studiju te u tihoj sobi sa sniženom razinom buke, kod ženskoga i muškoga govornika. Usporedbom linija na prikazima vidljive su veće timbralne razlike kod ženske govornice. Veća odstupanja primjetna su u području voluminoznosti (rast u neoptimalnim uvjetima, u tihoj sobi) te značajni padovi zvučne energije u područjima punoće, zvonkosti te okruglosti. Kod muškoga glasa primijećen je također porast energije u nižem dijelu spektra (voluminoznost) u uvjetima snimanja u tihoj sobi, ali u blažoj mjeri negoli kod ženske govornice. Veća je razlika prisutna u području zvonkosti (pad zvučne energije u neoptimalnome uvjetu snimanja), dok su u višim dijelovima spektra (blještavost i pucketavost) zamijećeni blaži padovi energije.



Slika 4. LTASS prikazi ženskoga (lijevo) i muškoga glasa (desno) u optimalnim uvjetima snimanja u akustičkome studiju te u tihoj sobi sa smanjenom razinom buke

Posljednji ispitivani uvjet odnosi se na jutarnji i popodnevni termin snimanja. S obzirom na to da vokalna neaktivnost tijekom spavanja dovodi do jutarnje promuklosti i nepripremljenosti glasa na vokalnu aktivnost, pretpostavljalo se da će dugotrajni prosječni spektri glasa iskazati pad zvučne energije, posebice u nižim i višim dijelovima spektra, u usporedbi s prikazima glasa u popodnevnim satima (kada je govorni aparat već dovoljno ugrijan i uvježban). Na slici 5. prikazane su razlike u optimalnim popodnevnim i neoptimalnim jutarnjim uvjetima snimanja ženskoga i muškoga govornika.



Slika 5. LTASS prikazi ženskoga (lijevo) i muškoga glasa (desno) u optimalnim uvjetima snimanja u popodnevnim satima te u jutarnjim satima

Rezultati kod ženske govornice pokazali su da u jutarnjim satima snimanja dolazi do blažega pada zvučne energije u spektralnim područjima punoće, zvonkosti i okruglosti, dok u drugim dijelovima spektra gotovo i nema razlike. Kod muškoga govornika zabilježen je veći pad zvučne energije i u više spektralnih područja: u području voluminoznosti, punoće, zvonkosti (veći pad zvučne energije) te blagi pad u području niže blještavosti.

6. Zaključak

Akustička snimanja metodološki su postupak kojim se prikupljaju materijali u svrhu daljnje govorne i jezične analize u sociofonetici i sociolingvistici. Stalnim razvojem tehnologije izloženi smo brojnim uređajima za snimanje i obradu govornoga zvuka, koji su nam omogućili kvalitetnije snimke, jednostavnije provođenje analiza itd. Međutim, među istraživačima su i dalje ostala neodgovorena neka istraživačka metodološka pitanja, na koja se ovim radom nastojalo odgovoriti. U radu su se opisali uvijek prisutni unutarnji i vanjski čimbenici koji utječu na rezultate akustičkih analiza te su se iznijeli potrebni koraci kojima bi se broj i veličina učinka ometajućih čimbenika umanjili. Osim toga, u radu su ispitani neoptimalni uvjeti snimanja (položaj glave od 90° u odnosu na mikrofona, predaleko/preblizu od mikrofona, u tihoj sobi sa sniženom razinom buke te u ranijim jutarnjim satima) te se utvrdila razlika u vrijednostima akustičkih

parametara (osnovnoga tona i formanata) u odnosu na optimalne uvjete snimanja. Rezultati različitih mjera fundamentalne frekvencije pokazali su da su bazična i alternativna bazična vrijednost fundamentalne frekvencije najrobusnije, odnosno da su najmanje podložne varijabilnosti zbog neoptimalnih uvjeta snimanja. S druge pak strane, minimalne i maksimalne vrijednosti F_0 pokazale su se kao najmanje pouzdane, posebice zbog oktavnih skokova i glotalizacije na krajevima rečenice.

Rezultati formantske analize pokazali su robusnost prvoga formanta kod ženske govornice, nešto manje robusan F_3 , dok je drugi formant najvarijabilniji, što je potvrđeno i kod Jovičića i sur. (2015). S druge pak strane, kod muškoga govornika najvarijabilniji je prvi formant, a drugi i treći formant podjednako su stabilniji od F_1 . Rezultati dugotrajnih prosječnih spektara pokazali su da u neoptimalnim uvjetima snimanja uglavnom dolazi do pada zvučne energije, posebice u nižim i srednjim dijelovima spektra. Do timbralnih promjena u višim dijelovima spektra došlo je kod premale udaljenosti od mikrofona, u snimkama iz tihe sobe te s neoptimalnim položajem glave.

Rezultati ovoga rada pokazali su da uvjeti snimanja mogu značajno utjecati na rezultate akustičke analize pa tako i na interpretaciju rezultata. S obzirom na to, u sociofonetskim i sociolingvističkim istraživanjima valjalo bi u najvećoj mogućoj mjeri umanjiti utjecaje unutarnjih i vanjskih ometajućih čimbenika te kontrolirati snimanje (opremu, vlastito ponašanje i ponašanje govornika) za vrijeme cijeloga trajanja. Ovim je radom potvrđeno da idealni tehnički uvjeti za osiguravanje pouzdanoga akustičkog materijala uključuju snimanje govornika u studijskim uvjetima na udaljenosti od 10 do 15 cm od mikrofona i najranije nekoliko sati nakon buđenja govornika. Također, idealna pozicija za snimanje bilo bi sjedenje nasuprot govorniku (položaj lice prema licu), kako bi se u što većoj mjeri ograničilo i ublažilo okretanje glave govornika u odnosu na mikrofona. S obzirom na to da istraživački uvjeti snimanja katkad ne mogu zadovoljiti kriterije idealnoga, svakako bi bilo preporučljivo sve govornike snimati u identičnim uvjetima i tako bi svi nepovoljni uvjeti snimanja u jednakoj mjeri utjecali na kvalitetu govornih zapisa.

Budući da su rezultati ovoga rada dobiveni na temelju govora dvaju govornika (jednog muškog i jednog ženskog), u budućnosti bi bilo korisno ispitati postavljene istraživačke hipoteze na većem broju govornika.

Literatura

- Arantes, P.; Eriksson, A. 2014. "Temporal stability of long-term measures of fundamental frequency." U Nick Campbell, Dafydd Gibbon i Daniel Hirst (ur.), *Proceedings of the 7th International Conference on Speech Prosody* (str. 1149–1152). Dublin: International Conference on Speech Prosody.
- Arantes, P.; Eriksson, A.; Gutzeit, S. 2017. "Effect of Language, Speaking Style and Speaker on Long-term F_0 Estimation." U *Proceedings of INTERSPEECH* (3897–3901). Stockholm: Interspeech.
- Baldwin, J.; French, P. 1990. *Forensic Phonetics*. London: London Pinter Publishers Limited.
- Bašić, M. 2012. *Kvantitativna sociolingvistička analiza crikveničkoga govora*. Doktorski rad. Zagreb: Filozofski fakultet.

- Biočina, Z. 2019. *Prepoznavanje i akustička analiza govora otoka Brača*. Doktorski rad. Zagreb: Filozofski fakultet.
- Biočina, Z.; Varošaneć-Škarić, G.; Bičanić, J. 2016. "Comparison of F_0 measures for male speakers of Croatian and Serbian." U *25th Annual Conference of the International Association for Forensic Phonetics and Acoustics* (str. 79–80). York: IAFPA.
- Boyd, Z.; Elliott, Z.; Fruehwald, J.; Hall-Lew, L.; Lawrence, D. 2015. "An Evaluation of Sociolinguistic Elicitation Methods." U *The Scottish Consortium for ICPhS 2015* (ur.), *The 18th International Conference of the Phonetic Sciences* (str. 1–5). Glasgow: ICPhS.
- Braun, A. 1995. "Fundamental frequency - How speaker-specific is it?." U Angelica Braun, Jens P Köster (ur.), *Studies in forensic phonetics* (str. 9–23). Trier: Wissenschaftlicher Verlag Trier.
- Byrne, D.; Dillon, H.; Tran, K.; Arlinger, S.; Wilbraham, K.; Cox, R. i sur. 1994. "An International comparison of long-term average speech spectra." *The Journal of Acoustical Society of America*, 94(4), 2108–2120.
- Byrne, C.; Foulkes, P. 2004. "The 'Mobile Phone Effect' on vowel formants." *International Journal of Speech, Language and the Law*, 11(1), 83–102.
- Fraser, H. 2014. "Transcription of indistinct forensic recordings: Problems and solutions from the perspective of phonetic science." *Language and Law*, 1(2), 5–21.
- Fraser, H. 2017. "'Assisting' listeners to hear words that aren't there: dangers in using police transcripts of indistinct covert recordings." *Australian Journal of Forensic Sciences*, 1–11.
- Galović, F. 2012. "Fonološke značajke govora Milne na otoku Braču." *Čakavska rič*, 40 (1-2), 87–100.
- Gold, E.; French, Peter J. 2011. "International practices in forensic speaker comparison." *International Journal of Speech, Language, and the Law*, 18(2), 293–307.
- Harrison, P. T. 2013. *Making Accurate Formant Measurements: An Empirical Investigation of the Influence of the Measurement Tool, Analysis Settings and Speaker on Formant Measurements*. Doktorski rad. York: Sveučilište u Yorku.
- Hayward, K. 2000. *Experimental Phonetics*. Harlow: Pearson Education Limited.
- Hollien, H. 1990. *The Acoustics of Crime: The New Science of Forensic Phonetics*. New York: Plenum Press.
- Hollien, H. 2002. *Forensic Voice Identification*. London: Academic Press.
- Huber, J. E.; Stathopoulos, E. T.; Curione, G. M.; Ash, T. A.; Johnson, K. 1999. "Formants of women, children and men: The effect of vocal intensity variation." *The Journal of Acoustical Society of America*, 106(3), 1532–1542.
- Hudson, T.; de Jong, G.; McDougall, K.; Harrison, P.; Nolan, F. 2007. " F_0 statistics for 100 young male speakers of standard Southern British English." U Jürgen Trouvain, William John Barry (ur.), *16th International Congress of Phonetic Sciences* (str. 1809–1812). Saarbrücken: ICPhS.
- Jessen, M.; Köster, O.; Gfroerer, S. 2003. "Effect of increased vocal effort on average and range of fundamental frequency in a sample of 100 German-speaking

- male subjects." U Maria-Josep Sole, Daniel Recasens, Joaquin Romero (ur.), *5th International Congress of Phonetic Sciences* (str. 1623–1626). Barcelona: Causal Productions.
- Jessen, M. 2008. "Forensic Phonetics." *Language and Linguistics Compass*, 2(4), 671–711.
- Jovanović, N. M.; Jovičić, S. T. 2011. „Uticaj načina upotrebe mobilnoga telefona na formante srpskih vokala." U *19th Telecommunications forum TELFOR 2011* (str. 1083–1086). Beograd: IEEE.
- Jovičić, S. T.; Jovanović, N. M.; Subotić, M.; Grozdić, Đ. 2015. "Impact of mobile phone usage on speech spectral features: some preliminary findings." *The International Journal of Speech, Language and the Law*, 22(1), 111–125.
- Köster, O.; Schiller, N.; Künzel, H. J. 1995. "The influence of native-language background on speaker recognition." U Kjell Elenius i Peter Branderud (ur.), *International Congress of Phonetic Sciences* (str. 306–309). Stockholm: ICPHS.
- Köster, O.; Schiller, N. 1997. "Different influences of the native language of a listener on speaker recognition." *Forensic Linguistics*, 4(1), 18–27.
- Künzel, H. 2001. "Beware of the 'telephone effect': the influence of telephone transmission on the measurement of formant frequencies." *Forensic Linguistics*, 8(1), 80–99.
- Labov, W. 1972. *Sociolinguistic Patterns*. Philadelphia: Philadelphia University of Pennsylvania Press.
- Ladefoged, P. 2003. *Phonetic Data Analysis. An introduction to Fieldwork and Instrumental Techniques*. Oxford: Blackwell Publishing.
- Langston, K. 2006. *Čakavian Prosody. The Accentual Patterns of the Čakavian Dialects of Croatia*. Bloomington, Indiana: Slavica.
- Lindh, J. 2006. "Preliminary Descriptive F0-statistics for Young Male Speakers." *Working Papers in Linguistics*, 52, 89–92.
- Maher, R. C. 2009. "Audio Forensic Examination - Authenticity, enhancement, and interpretation." *IEEE Signal Processing Magazine*, 84, 84–94.
- Nolan, F. 1983. *The Phonetic Bases of Speaker Recognition*. Cambridge: Cambridge University Press.
- Nolan, F. 2005. "Forensic Speaker Identification and the Phonetic Description of Voice Quality." U William J. Hardcastle, Janet Mackenzie Beck (ur.), *A Figure of Speech: a Festschrift for John Laver* (str. 385–411). New York, London: Routledge.
- Nolan, F.; McDougal, K.; De Jong, G.; Hudson, T. 2009. "The DyViS database: style-controlled recordings of 100 homogeneous speakers for forensic phonetic research." *International Journal of Speech, Language and the Law*, 16(1), 31–57.
- Rathcke, T.; Stuart-Smith, J.; Torsney, B.; Harrington, J. 2017. "The beauty in a beast: Minimising the effects of diverse recording quality on vowel formant measurements in sociophonetic real-time studies." *Speech Communication*, 86, 24–41.
- Rose, P. 2002. *Forensic Speaker Identification*. London, New York: Taylor & Francis.
- Schulman, R. 1989. "Articulatory dynamics of loud and normal speech." *Journal of*

- Acoustical Society of America*, 85(1), 295–312.
- Skarnitzl, R.; Vaňková, J. 2017. "Fundamental frequency statistics for male speakers of common Czech." *Acta Universitatis Carolinae Philologica*, 3, 7–17.
- Šimunović, P. 2011. *Čakavska čitanka*. Zagreb: Golden marketing-Tehnička knjiga.
- Šprljan, N. 2015. *Opis govora Selaca na Braču s posebnim naglaskom na akcentuaciju (u usporedbi s okolnim mjestima)*. Doktorski rad. Zagreb: Filozofski fakultet.
- Švec, J. G.; Granqvist, S. 2010. "Guidelines for selecting microphones for human voice production research." *American Journal of Speech-Language Pathology*, 19, 356–368.
- Tillery, J.; Bailey, G. 2003. "Approaches to real time in dialectology and sociolinguistics." *World Englishes*, 22(4), 351–365.
- Traunmüller, H. 1994. "Conventional, biological, and environmental factors in speech communication: A modulation theory." *Phonetica*, 51, 170–181.
- Traunmüller, H.; Eriksson, A. 1995. "The perceptual evaluation of F₀-excursions in speech as evidenced in liveliness estimations." *The Journal of the Acoustical Society of America*, 97(3), 1905–1915.
- Varošanec-Škarić, G.; Biočina, Z.; Bičanić, J. 2016. "Usporedba mjera F₀ muških govornika hrvatskog jezika i srpskog jezika." U Nikolaj Lazić, Elenmari Pletikos Olof (ur.), *Istraživanja govora: Deveti znanstveni skup s međunarodnim sudjelovanjem* (str. 110–112). Zagreb: Hrvatsko filološko društvo.
- Varošanec-Škarić, G.; Biočina, Z.; Kišiček, G. 2017. "Comparison of F₀ measures for male speakers of Croatian, Serbian and Slovenian." U Mária Gósy (ur.), *CAPSS2017 Workshop on Challenges in Analysis and Processing of Spontaneous Speech* (str. 40–41). Budimpešta: CAPSS.
- Vermeulen, J. 2009. *Beware of the Distance: Evaluation of Spectral Measurements of Synthetic Vowels Re-recorded at Different Distances*. Doktorski rad. York: Sveučilište u Yorku.
- Vermeulen, J.; Cambier-Langeveld, T. 2017. "Outstanding cases: about case reports with a 'strong' conclusion." U Gordana Varošanec-Škarić, Anita Runjić Stoilova (ur.), *26th Annual Conference of the International Association for Forensic Phonetics and Acoustics* (str. 31–33). Split: Hrvatsko filološko društvo.
- Zhang, C.; Morrison, G. S.; Enzinger, E.; Ochoa, E. 2013. "Effects of telephone transmission on the performance of formant-trajectory-based forensic voice comparison – Female voices." *Speech Communications*, 55, 796–813.

Internetski izvori

- Boersma, P.; Weenink, D. 2015. *Praat: doing phonetics by computer* (verzija 6.0.31). <http://www.praat.org/>. (13. 12. 2018.).
- Sjölander, K.; Beskow, J. 2017. *Wave Surfer*. <https://sourceforge.net/projects/wavesurfer/>. (6. 11. 2017.).

Prilozi

Popis slika

- Slika 1. LTASS prikazi ženskoga (lijevo) i muškoga glasa (desno) u optimalnim uvjetima snimanja i s neoptimalnim položajem glave
- Slika 2. LTASS prikazi ženskoga (lijevo) i muškoga glasa (desno) u optimalnim uvjetima snimanja i na neoptimalnoj udaljenosti od mikrofona (predaleko)
- Slika 3. LTASS prikazi ženskoga (lijevo) i muškoga glasa (desno) u optimalnim uvjetima snimanja i na neoptimalnoj udaljenosti od mikrofona (preblizu)
- Slika 4. LTASS prikazi ženskoga (lijevo) i muškoga glasa (desno) u optimalnim uvjetima snimanja u akustičkome studiju te u tihoj sobi sa smanjenom razinom buke
- Slika 5. LTASS prikazi ženskoga (lijevo) i muškoga glasa (desno) u optimalnim uvjetima snimanja u popodnevnim satima te u jutarnjim satima

Popis tablica

- Tablica 1. Prikaz statistički značajnih razlika u formantima među različitim uvjetima snimanja
- Tablica 2. Prikaz vrijednosti fundamentalne frekvencije u različitim uvjetima snimanja kod ženske govornice

METHODOLOGY FOR ACOUSTIC RECORDING

Abstract

Sociolinguistics and sociophonetics place the native speaker in the center of their research. Since different research goals demand a different methodological setting it is very important to assess which steps are necessary and sufficient to answer the research questions. Thus the authors of this paper recorded more than 300 speakers, which gave them ample methodological experience but also opened a lot of questions. In many previous studies, the importance of the optimal conditions of recording which contribute to a spontaneous spoken reaction of the speaker is emphasised, especially for the purpose of collecting audio material of dialect speech. The influences of different speaking styles, microphone distance (Vermeulen 2009; Fraser 2017) and head position (Jovičić et al. 2015) on acoustical parameters in English, Dutch, Serbian, etc. were examined in a smaller number of acoustic studies. This paper will look into conditions of acoustic recording and their influence on different acoustic parameters (formant frequencies and fundamental frequency). For the purposes of this research two speakers (one male and one female voice) with a healthy voice and speech and language development were recorded. In the first part of the research, the speakers read a short text, while in the second part spontaneous speech was recorded. Both speakers were recorded in the morning and in the afternoon in two different settings: in a sound booth and in a room with a lower noise level, both with a high-quality recorder. The speakers were also recorded at a different distance from the microphone and with different head positions. Three research questions were formulated: will different a) terms of recording b) conditions of recording and c) distance and head position relative to the microphone influence the values of formant frequencies and the fundamental frequency? Apart from that, this paper will emphasize the methodological details that should be taken into consideration while collecting audio material for acoustic analysis purposes. From the fundamental frequency measurements, the basic value and median have been shown to be the most robust, and from the formant frequencies, the third formant. The most significant differences between values are in the non-optimal distance of the speaker from the microphone. This research aims to contribute to the clarification of specific acoustic issues and to facilitate future researchers to set up the necessary methodological settings in sociolinguistic, sociophonetic and phonetic-forensic studies due to their undoubted influence on data analysis and results.

Key words: sociolinguistics, acoustic recording, methodology, fundamental frequency, formant frequencies