
UDK 801.42:575

801.42:316

Izvorni znanstveni rad

Prihvaćeno 14.7.1994.

Nina Bertapelle,
Veleposlanstvo Španjolske, Zagreb

NASLJEDNI I DRUŠTVENI ČIMBENICI GLASA

SAŽETAK

Članak se bavi statičkom i nekim elementima dinamičke boje glasa. Namjera je potvrditi pretpostavku da potomci glasom sliče roditeljima i otkriti je li glas zadan anatomijom ili se uči djelovanjem slušne sprege i imitacijom. Analizirane su snimke govora po 15 majki, kćeri, očeva i sinova. Provedena su akustička mjerenja dugotrajnog usrednjenog spektra, osnovnog tona i intenziteta, a potom su se izračunale standardne devijacije i koeficijenti varijacije osnovnog tona i intenziteta. U kratkovremenskom spektru vokala locirana su i očitana prva tri formanta. Iz njih se izračunala udaljenost od larinksa do prednjih zubi, što predstavlja jedini anatomski parametar. Skalom procjene procjenjivala se sličnost glasova srodničkih i nesrodničkih parova u varijanti normalnih i obrnutih snimaka. Rezultati pokazuju da se muški glasovi spektrom međusobno više razlikuju, kao i da postoji veći stupanj sličnosti među srodničkim glasovima nego među nesrodničkim. Srodnici, međutim, nisu sličniji po prosječnom osnovnom tonu, prosječnom intenzitetu, njihovim promjenama u vremenu, a niti po vokalnoj duljini. Rezultati korelacija akustičkih, anatomskog i psihoakustičkog parametra upućuju na mogućnost postojanja različitih načina slušanja muških i ženskih glasova.

Ključne riječi: boja glasa, nasljedni čimbenici, društveni čimbenici

I. UVOD

Ljudski zvuk koji nastaje titranjem glasnica zove se glas i pojavljuje se u dva oblika: kao pjevani i kao govorni. Ovaj će se rad baviti govornim glasom. Kada fonetičari žele definirati glas, ubrzo postaju svjesni mnoštva čimbenika koji sudjeluju u njegovoj kreaciji, variranju i percepciji. U ovom će se radu zanemariti segmentalne karakteristike ljudskoga govornog zvuka, a uzimat će se u obzir suprasegmentalne karakteristike. Dvije su suprasegmentalne komponente glasa. Jedna je ona koju čine opće kvalitete glasa – boja u širem smislu¹ ili glasovna dinamika (način na koji se glasom "rukuje"). One su vezane za slog ili neke glazbene kvalitete kao što su ritam, tremolo, vibrato, legato, staccato (Skarić 1982: 185), a pokazuju se kao promjene zvuka u vremenu. Druga je suprasegmentalna komponenta boja u užem smislu. Ona je statična, odnosno ima kvazi-trajni karakter. U radnji će se promatrati upravo ta komponenta. Jedan dio boje nepromjenjiv je i u njemu su sadržane i naslijeđene govornikove anatomske osobine i one stečene, vidljive u izgovornim navikama. Drugi dio boje promjenjiv je i u njemu se pokazuju govornikove izražajne mogućnosti. Osim što je kao vokalska boja materijal za prenošenje lingvističkog sadržaja (boja glasa mijenja se kod svakog vokala), ona ima i paralingvističku funkciju kada signalizira afektivni dio poruke, a također i ekstralingvističku funkciju jer je boja nositelj obavijesti o fizičkim, psihološkim i socijalnim karakteristikama govornika. Svi su ti znakovi sjedinjeni u istom fenomenu, a teškoća je fonetičara što njegove elemente mora izlučiti iz cjeline vokalne izvedbe.

Kada želimo raščlaniti ovu prvu, fizičku razinu, susrećemo se s velikim problemom – kako dobiti pouzdane podatke o anatomskoj građi govornih organa. Zbog teškoća pri dobivanju izravnih podataka o geometriji vokalnog trakta, prošlih su se godina istraživanja na tom području provodila tako da se obično ta metoda zamjenjivala neizravnim pristupom, a to znači dobivanjem podataka o artikulatorima iz akustičkih podataka. Različitim napetostima muskulature uobličuje se i preobličuje zadana anatomija u različite rezonantne prostore u kojima se proizvodi boja. Taj je fenomen muskularnog rada vezan za naviku, pa je jasno da on nije jednom zauvijek određen. Naprotiv, promjenjiv je, a na njega djeluju razni psihički i socijalni čimbenici, u prvom redu imitacija drugih osoba. Kada se govori o tomu, kao prvo treba spomenuti edinburškog fonetičara Johna Lavera, koji je iznio svoju koncepciju proizvodnje i opisa (artikulacijski i akustički) različitih boja glasa i na čijim se postavkama temelje sve suvremene studije o boji glasa. Autor je definirao setting (Laver 1980:2), u ovom radu artikulacijski postav, kao trajnu tendenciju muskulature da održava određeni oblik govornih organa, muskularnu naviku koja čini osnovu pokreta što ga proizvode sekvence segmenata.

(1) U angloameričkoj se literaturi za ovaj pojam koriste nazivom *voice quality*, dok literarura romanskih jezičnih područja rabi nazive fr. *timbre*, tal. i šp. *timbro*. Kod nas se udomaćio termin boja glasa.

Osim akustičkih analiza, rade se i psihoakustički pokusi iz kojih se doznaje kako čovjek sluša glas. To je psihoakustičko pitanje postalo posebno zanimljivo kada je čovjek zaželio naučiti stroj prepoznavati nečiji glas, pa su objavljeni mnogi radovi koji boju glasa kao perceptivni fenomen ne shvaćaju kao jednodimenzionalnu, nego kao višedimenzionalnu kvalitetu (Singh, Murry 1977. i 1978; Colton, Estil 1981; Pascal 1984; Pittam 1987). U takvoj koncepciji glas označuje točku prostora, a problem je odrediti broj značajnih dimenzija kojim ga se opisuje. U multidimenzionalnom pristupu boji glasa svrha je otkriti najmanji broj dimenzija koji će adekvatno reprezentirati podatke, pri čemu se iskorištava multidimenzionalno skaliranje i indscal analiza.

II. TEMA ISTRAŽIVANJA

Slijede osnovna pitanja na koja će se pokušati odgovoriti opisanim pokusom.

1. Jesu li glasovi srodnika (majki i kćeri, očeva i sinova) sličniji nego glasovi nesrodnika,² dakle bilo kojih ljudi?
2. Je li glas naslijeđen putem naslijeđene zadane anatomije, ili stečen imitacijom onih s kojima se želimo identificirati, u ovom slučaju roditelja?

Da bi se na to odgovorilo, bilo bi idealno pokus provoditi na uzorku odraslih potomaka koji su rasli uz svoje roditelje, kao i onih koji su živjeli s "usvojenim" roditeljima (uz mogućnost da se eksperimentalnom postupku podvrgnu i njihovi biološki roditelji), no takav je pokus organizacijski teško organizirati, pa se promatrala samo prva skupina, dakle roditelj i odrasli potomak koji su zajedno i živjeli.

A kakva je svrha istraživanja? Je li uopće važno kako je netko "došao" do glasa kakav ima. Prihvatimo li tezu da se glas uči imitacijom, onda je moguće zamisliti da čovjek nauči "krivi" glas, onaj koji ne odgovara njegovim anatomskim karakteristikama i za koji mora uložiti previše napora ionako osjetljivih struktura, što često vodi u vokalnu patologiju. Ovaj se aktualni problem u znanosti nastoji riješiti objektivnom evaluacijom glasa, akustičkom analizom vokalnih parametara te primjenom elektromiografije i elektoglotografije. Na pitanje da li je neki glas zaštitni, najbolje bi odgovorila paralelna interpretacija akustičkih i miografskih rezultata, ali je problem u nedostupnosti laringalne muskulature za aplikaciju elektroda. Jer, ako glas nije "genom zauvijek određen" (Skarić 1982: 190), onda ga se može mijenjati i to na bolje i kultivirati.

(2) Riječju nesrodnici koristit ćemo se da imenujemo parove onih muških i ženskih govornika koji nisu majka i kćer, odnosno otac i sin.

III. METODE

A. *Govornici i govorna građa*

Snimljeno je 60 govornika: 15 očeva i 15 sinova, te 15 majki i 15 kćeri. Govornici su imali između 18 i 70 godina. Svi su bez vokalne patologije i nikada nisu školovali glas. Zadani tekst za čitanje bio je sastavljen od vokala, sonanata i okluziva, a nije sadržao frikative i afrikate. Prije snimanja govornici su se upoznali s tekstom. Također je s njima proveden kratki razgovor radi opuštanja. Snimalo se u gluhoj komori na magnetofon Revox A-77 i mikrofonski Sehnheiser MD 421. Udaljenost od mikrofona do usta bila je 25 cm. Govornici su čitali njima ugodnom jačinom glasa.

B. *Akustička analiza*

1. Uzorak od 128 sek neprekidnoga govora za svakoga govornika analiziran je terčno/oktavnim analizatorom Brüel-Kjaer 2131. Signal digitaliziran 8-bitnim A/D konvertorom analizirao se u realnom vremenu. Usrednjavanje je bilo linearno. Spektri dobiveni ovom terčnom analizom potom su iscrtavani na pisaču Brüel-Kjaer 2307. Relativne vrijednosti amplituda očitane su na 20 terci: za ženske signale između 100 Hz i 8 kHz, a za muške signale između 80 Hz i 6.3 kHz. Donja frekvencija očitavanja i uzimanja u obzir za usporedbu određena je područjem pojavljivanja osnovnog tona, a gornja frekvencija zahtjevom da šum ne prekriva signal na tom području, što je empirijski utvrđeno spektralnom analizom "tišine" na vrpci.

2. Analogne magnetofonske snimke (20 sek svakoga govornika) filtrirale su se niskopropusnim filtrom 4.5 kHz i digitalizirale 12-bitnim A/D konvertorom pri frekvenciji uzorkovanja $F_u=10$ kHz.

2.1. Za detekciju po vrijednosti upotrijebljen je SIFT algoritam (Markel, Gray 1976). Tako je analizirano 14.5 sek govora (5×2.9 sek nasumce izabranih) svakoga govornika. Podaci o F_0 vrijednostima pri vremenu analize $t=10$ msec uneseni su u računalo sa statističkim programom za izračunavanje srednje vrijednosti i standardne devijacije.

3.1. Srednji intenzitet tijekom govorenja procijenjen je sa SPL metra 01, čiji je mikrofonski bio udaljen od govornikovih usta 25 cm.

3.2. Iz uzorka (pod točkom 2.1) 14.5 sek govora očitane su vrijednosti kratkovremenske energije E , uz period analize $t=10$ msec. Srednja energija signala po uzorku iskazana je kao digitalizirana naponska razina na A/D konvertoru. Izračunala se srednja vrijednost i standardna devijacija, pa se standardna devijacija preračunala u decibele.

4. U PCM prikazu riječi "vrlo", uzastopnim skeniranjem i zumiranjem lociran je vokal \downarrow , kao dio vokalnog r , što je provjereno vraćanjem tog dijela signala u analognu formu i slušnom verifikacijom. Spektralna analiza obavila se brzom Fourierovom transformacijom (FFT) $N=256$ točaka uz

primjenu Hammingovog prozora. Analizirano je 25 msek kvazistacionarnog dijela toga vokala. Primijenjeno je linearno centralno glaćenje (Stamenković) $K=5$ za ženske glasove i $K=3$ za muške glasove. Iz takvog prikaza očitani su prvi, drugi i treći formant dovođenjem kursora na prva tri intenzitetska vrha u spektru.

C. Psihoakustička procjena

1. Priprema građe za slušanje

Prva rečenica zadanog teksta : "Marko je bio moj dragi kolega." svakog od 60 govornika presnimljena je na radnu vrpca. Nakon toga pripremljena je vrpca za ocjenjivanje sličnosti parova, a parovi su se slagali prema sljedećem kriteriju. Naime, da bi se izbjeglo dugotrajno slušanje $15 \times 15 \times 2 = 450$ parova, prema načelu "svaki sa svakim", napravljen je izbor. Prvi član para je majka, a drugi član čine četiri mogućnosti. U prvoj varijanti to je biološka kćer. Ostale tri varijante izabrane su prema koeficijentu korelacije, stupnju slaganja točaka histograma dugotrajnoga usrednjenoga spektra i to: (1) tuđa kćer koja je s danom majkom, prema tom kriteriju, najviše povezana, (2) ona s manjim koeficijentom korelacije i (3) ona koja je na sedmom, "srednjem" mjestu u stupnjevanom redoslijedu. Tako je učinjeno sa svih 15 majki i svih 15 očeva. Na taj je način pripremljeno 120 parova: 60 ženskih i 60 muških parova nasnimljenih slučajnim redoslijedom. Među članovima para jest stanka od 1 sek, a među parovima 3 sek.

2. Slušanje i procjena

Zadanu građu slušalo je 15 slušača-procjenitelja, diplomiranih fonetičara. Zadatak je bio procijeniti skalom procjene od 7 točaka stupanj sličnosti glasova u paru. Slušanje prvih 20 parova služilo je kao slušna vježba, a nakon toga je počelo ocjenjivanje. Ista građa slušala se na dva načina; kao normalna snimka i kao snimka reproducirana unatrag (backwards). Da bi se izbjegao utjecaj zamora, te su se verzije slušale uz stanku od jednog dana. I prije slušanja druge varijante provedena je slušna vježba na prvih 20 parova. Snimke su se reproducirale na magnetofonu Uher 4000 s KOSS-slušalicama, uz konstantnu ugodnu glasnoću.

IV. REZULTATI

A. AKUSTIČKI PARAMETRI

1. Dugotrajni usrednjeni spektar LTAS

Rezultat analize jest krivulja spektra što daje obavijest o energetskej distribuciji u govornom signalu. Takav dugotrajni spektar zanemaruje različitost vokalnog trakta u različitim dijelovima govora, jer se te dugotrajne promjene u procesu uprosječivanja izglađuju.

Stupanj međusobnog slaganja krivulja spektra svake majke sa svakom kćerkom i svakog oca sa svakim sinom izračunat je iz podataka o relativnim vrijednostima amplituda statističkom metodom Pearsonove korelacije. Rezultati su sljedeći:

1.1. Općenito se opaža da su svi koeficijenti vrlo visoki, što govori o visokom stupnju slaganja spektra glasova govornika istog spola.

1.2. Na razini značajnosti $p=0.01$ kod ženskih govornika samo 12 ili 5.3% koeficijenata korelacije ne ulaze u taj interval, a kod muških 44 ili 19.5% parova.

1.3. Da bi se odgovorilo na pitanje jesu li srodnici značajno sličniji od nesrodnika, provjeravala se t-testom značajnost razlika aritmetičkih sredina uzoraka i srodnika i nesrodnika. Rezultat t-testa je za ženske govornike $t=2.94$, a za muške govornike je $t=4.1$, što treba interpretirati kao značajnu razliku. Stupanj povezanosti krivulja srodnika jest, dakle, značajno veći od stupnja povezanosti krivulja spektra nesrodnika, što znači da je spektar glasova srodnika značajno sličniji od spektra glasova nesrodnika istog spola.

1.4. Iz podataka o varijabilitetu rezultata (koeficijenti korelacije kreću se kod muških glasova između 0.40 i 0.92, a kod ženskih između 0.80 i 0.98) zaključuje se da se muški glasovi međusobno više razlikuju od ženskih.

2. Visina osnovnog tona

2.1. Srednji osnovni ton za ženske govornike jest 174.25 Hz, a za muške 124.28 Hz, majki 166.54, kćeri 181.84, očeva 125.1, a sinova 123.46 Hz. Iz tih podataka može se uočiti tendencija da se snizi prosječan osnovni ton s porastom kronološke dobi, ali također da je razlika među prosječnim osnovnim tonovima pripadnika dvaju različitih muških naraštaja mnogo manja od one kod žena. Moguće je također zaključiti da muškarci već u ranijoj dobi forsiraju niski ton, pri čemu dominiraju psihološki i društveni razlozi nad anatomsko-fiziološkima.

2.2. Da bi se ustvrdila statistička razlika prosječnog osnovnog tona uzorka parova srodnika i nesrodnika, testirana je značajnost razlika aritmetičkih sredina. Rezultat t-testa jest $t=0.46$ za ženske govornike, a za muške $t=2.61$. Srodnici, dakle, nemaju sličniji prosječni ton od nesrodnika.

2.3. Osim podataka o srednjem osnovnom tonu, zanimljivo je znati i kako se on mijenja u vremenu. Zato su se računale standardna devijacija osnovnog tona preračunata u tonove i koeficijent varijacije izražen u postotku. Statistički se ne razlikuju značajno srodnici od nesrodnika. Također se statistički ne razlikuju muški od ženskih govornika.

3. Intenzitet I

3.1. Srednja vrijednost intenziteta kod ženskih govornika je 71.06 dB, a kod muških je 73.43 dB.

3.2. Upotrebene su dvije mjere promjene intenziteta u vremenu: standardna devijacija preračunata u decibele i koeficijent varijacije izražen u postocima. Srodnici prema tim kriterijima nisu sličniji nego nesrodnici. Muški i ženski govornici ne razlikuju se značajno po ovim kriterijima.

3.3. Usput se ispitala međuzavisnost srednjeg osnovnog tona i srednjeg intenziteta, te koeficijenta varijacije osnovnog tona i koeficijenta varijacije intenziteta. Rezultati pokazuju da kod oba spola veće intonacijske promjene prate i veće intenzitetske promjene. Samo kod muškaraca opaža se da oni koji govore glasnije, govore i višim prosječnim tonom, dok žene koje govore glasnije imaju jači tremolo u glasu.

4. Formanti vokala \bar{e}

4.1. Očitani su formanti, pa su izračunate njihove prosječne vrijednosti.

Tabela 1. Aritmetičke sredine i standardne devijacije formanata vokala \bar{e} u Hz

	F ₁	s ₁	F ₂	s ₂	F ₃	s ₃
očevi	461.9	35.9	1319.8	111.0	2287.7	91.1
sinovi	456.6	40.7	1266.2	74.9	2270.2	91.1
muškarci	459.3	38.5	1293.0	98.4	2280.0	91.6
majke	539.8	65.2	1592.2	100.3	2671.6	155.8
kćeri	556.9	73.0	1528.0	106.8	2606.0	252.7
žene	548.3	69.7	1560.2	108.4	2633.8	212.5

4.2. Iz tih je podataka o distribuciji formanata u uzorku izračunato u kolikom je postotku svaki formant viši kod žena nego kod muškaraca.

Tabela 2. Razlika ženskih u odnosu na muške formante u postotku

F1	F2	F3
+19.4	+20.6	+15.7

4.3. Također je izračunato u kakvom su odnosu srednje vrijednosti drugog i trećeg formanta prema prvom formantu.

Ženski glasovi F1 : F2 : F3 = 1 : 2.84 : 4.83

Muški glasovi F1 : F2 : F3 = 1 : 2.81 : 4.96

Budući da se za taj odnos može reći da malo odstupa od teoretski očekivanog, da su navedene standardne devijacije ponaosob male, a međusobno slične, kao i da su razlike među formantima unutar očekivanih za neutralni vokal, smatra se da su podaci dovoljno pouzdani da bi se mogli dalje obrađivati.

B. Anatomski parametar – Duljina vokalnog trakta l

Ovaj se podatak izračunao iz formanata, prema formuli

$$l = (2k - 1) c / 4F_k \text{ (m)}$$

U gornjem izrazu indeks F_k označava frekvenciju k -og formanta, a $c=353$ (m/s) brzinu zvuka u zraku.

1. Prosječni l muških govornika jest 0.1938 m, a ženskih govornika 0.1655 m, što je razlika od 14.6%.

2. Rezultat t-testa kaže da nema statistički značajne razlike među razlikama duljina srodnika i nesrodnika.

C. Psihoakustički parametri

1. Iz rezultata ocjena sudaca vidjelo se da se varijabilitet ocjena nije bitno razlikovao kod ženskih ili muških glasova, niti kod snimaka normalno ili unatrag reproduciranih.

2. Stupanj slaganja srednjih ocjena 15 sudaca za varijantu normalnog govora sa srednjim ocjenama za varijantu obrnutaoga govora izračunat je metodom Pearsonove korelacije, a rezultati pokazuju da su suci i u varijanti slušanja govora unatrag dosljedno obavljali slušnu zadaću procjene sličnosti.

3. Stupanj međusobnog slaganja sudaca pri ocjenjivanju istih parova normalnih i obrnutih snimaka izrazito je niži kod obrnutih snimaka ženskih glasova (koeficijenti korelacije su od negativnih do 0.2) nego ostalih triju varijanata (koeficijenti korelacije su između 0.2 i 0.5).

4. Stupanj slaganja suca samog sa sobom pri ocjenjivanju normalnih i obrnutih snimaka, izražen koeficijentom korelacije, kreće se između -0.05 i 0.4 što govori da su suci općenito bili nedosljedni u toj slušnoj zadaći, ali i da im je uspješnost bila prilično različita.

5. Testirana je razlika značajnosti prosječnih ocjena sličnosti srodničkih i nesrodničkih parova, i te se ocjene na 1% značajnosti razlikuju i kod muških i kod ženskih snimaka kod normalnih snimaka. Kod obrnutih snimaka ocjene sličnosti glasova srodničkih parova nisu značajno različite od nesrodničkih, iako treba primijetiti da je taj rezultat kod muških obrnutih snimaka puno veći. To upućuje na mogućnost postojanja različitih strategija slušanja ženskih i muških glasova.

D. Korelacije akustičkih, anatomskog i psihoakustičkih parametara

Da bi se znalo u kakvom su međuodnosu spomenuti podaci, korištena je metoda Pearsonove korelacije. Rezultati su prikazani u sljedećoj tabeli:

Tabela 3. Koeficijenti korelacije akustičkih, anatomskog i psihoakustičkih parametara

ŽENSKI GOVORNICI

	r LTAS	Δl	Δf_0	$\Delta s(f_0)$	Δl	ΔSPL s
ocjena normalne snimke	0.3942	-0.0505	-0.1700	-0.0045	-0.0678	-0.0798
ocjena obrnute snimke	0.4624 ^z	-0.1383	-0.4156	-0.2069	-0.0753	0.2726

z = značajnost na 0.02

MUŠKI GOVORNICI

	r LTAS	Δl	Δf_0	$\Delta s(f_0)$	Δl	ΔSPL s
ocjena normalne snimke	0.5701 ^c	0.0077	-0.4585 ^z	-0.0074	-0.0669	-0.2056
ocjena obrnute snimke	0.4777 ^z	-0.0576	-0.2666	0.0944	-0.2667	-0.0466

z = značajnost na 0.02

ŽENSKI I MUŠKI GOVORNICI

	r LTAS	Δl	Δf_0	$\Delta s(f_0)$	Δl	ΔSPL s
ocjena normalne snimke	0.4668 ^z	-0.0182	-0.3091	-0.0088	-0.0698	-0.1483
ocjena obrnute snimke	0.4523 ^z	0.0906	-0.3220	-0.0108	-0.1968	0.0675

z = značajnost na 0.01

Očita je važnost podudarnosti spektra kod procjene sličnosti glasova. Više kod muških nego kod ženskih, više kod ženskih obrnutih nego kod ženskih normalnih snimaka.

Iz navedenih podataka može se zaključiti da se prislićivanje muških i ženskih glasova slušanjem općenito temelji na uspoređivanju spektra i podudarnosti osnovnog tona, ali da ipak postoje različite strategije slušanja i prislićivanja ženskih i muških glasova i to tako da su kod normalnih snimaka muških glasova izrazito značajni spektar i prosječni osnovni ton, dok se kod ženskih glasova čini da to postaje važno tek kod obrnutih snimaka. Razlog je vjerojatno u odsutnosti nekog drugog ili drugih značajnih parametara koji postoje u normalnim snimkama ženskoga govora, ali koji se u rezultatima ovdje korištenih akustičkih analiza nisu odvojili.

V. RASPRAVA

1. Od početka pokusa bilo je jasno da će se vrlo teško odvojiti razine i promatrati samo boja glasa u užem smislu, odvojeno od šire artikulacijske razine, i u akustičkoj analizi i u slušanju. Budući da je boja dugotrajni fenomen, promatran je dugotrajni usrednjeni spektar, pri čemu se postavlja pitanje kako isključiti šumnu komponentu laringealne provenijencije, a naročito kako uspoređivati krivulje spektra. Ovo drugo razni su autori pokušali različito riješiti: Harmegnien "Standardnom devijacijom razlike distribucije" (SDDD), neki drugi kontrastivnim opservacijama (Tarnoszy 1958; Banuls-Terol 1971; Fromby, Monsen 1982), ili su spektru pridruživani indeksi koje se matematički uspoređivalo, kao primjerice odnos energije do i od 1 kHz (Frøpkjaer-Jensen, Prytz 1976), razina intenzitetskih vrhova na 140 Hz, 500 Hz i 2 kHz (Sundberg 1988) ili odnos energije između frekvencijskog opsega 300-800 Hz i 1.5-3 kHz (Kitzing 1986). Neki su se pak autori koristili mjerenjima što su proizvodila brojčane pokazatelje stupnja sličnosti mjerenih spektara: Euklidova distanca (Majewski, Hollien 1974) ili koeficijent korelacije (Zalewski, Majewski, Hollien 1975). Koeficijent korelacije, ovdje korišten, smatra se klasičnim, praktičnim i djelotvornim indeksom za usporedbu spektra. Neosjetljiv je na razliku srednjeg intenziteta, pa ne zahtijeva prethodnu normalizaciju. I u ovom radu taj je postupak opravdao svoje korištenje.

2. Srednja vrijednost F_0 za muške glasove slaže se s ostalim brojnim podacima iz literature (Majewski 1972; Künzel 1989), dok podatak o ženskoj prosječnoj visini osnovnog tona ne odgovara podacima koje su dobili drugi autori. On se obično kreće kod odraslih žena između 200 i 220 Hz. Razlika se pripisuje nepreciznosti korištene metode.

3. Podaci iz literature (Günzburger 1989) govore o značajnu apsolutne visine tona kod žena za psihoakustičke procjene glasa, za razliku od muškaraca. Takve zakonitosti ovdje nisu potvrđene. Dapače, rezultat opisanog pokusa jest da je upravo kod procjene muških glasova statistički značajna visina osnovnog tona za procjenu sličnosti glasa kod normalne snimke, dok je to kod žena tek kod snimaka obrnuto reproduciranih i uz niži koeficijent korelacije.

Standardne devijacije osnovnog tona, preračunate u tonove, nižih su vrijednosti nego one koje se navode u literaturi (Majewski 1972; Saxman 1967) što se može pripisati razlici eksperimentalnog postupka. Ta se građa, naime, čitala. Intonacijske su promjene, kao i u literaturi (Boone 1991) veće kod žena.

Iako se, prema podacima iz literature o značajnosti slušanja intonacijskih promjena kod percepcije ženskog glasa (Graddol, Swan 1983; Ohala 1984) očekivalo da se te intonacijske promjene uvažavaju kod prisličivanja glasa slušanjem, što bi se trebalo pokazati kao visoki koeficijent korelacije ocjena sličnosti normalnoga govora i razlike standardne devijacije osnovnog tona govornika u paru, takav rezultat nismo dobili. Tako treba tek odgovoriti na pitanje radi li se o prirodni pokusa u kojem se nije odvojila prava akustička mjera, što bi predstavljalo

intonacijski dinamizam, iako je u radu korišten isti parametar kao i u navedenim člancima (standardna devijacija osnovnog tona). Prihvati li se ipak teza da je kretanje intonacije kod žena veće, treba se pitati da li je ono fiziološki ili društveno uvjetovano. Tu se treba prisjetiti fiziološki uvjetovanog aperiodiciteta osnovnog tona (*jitter*) koje kod žena ovisi drukčije o osnovnom tonu, nego kod muškaraca, što se tumači različitim korištenjem fonacijskim sustavom (Orlikoff 1990). Također treba napomenuti da je disanje kod žena drukčije, kao i da je tesitura smještena, u odnosu prema fiziološkom rasponu, drukčije nego kod muškaraca. Iz rečenog se može pretpostaviti da veći intonacijski dinamizam jest fiziološki uvjetovan, a ne društvenom prisilom stečen.

4. Ženski govor je prosječno tiši, što su pokazali i pokusi drugih autora (Günzburger 1989). Podaci o SPL sigma vrijednostima odgovaraju nađenima u literaturi (Hanley 1966). Kao druga mjera promjene intenziteta u vremenu promatran je koeficijent varijacije intenziteta, postotak promjene intenziteta u vremenu ili *shimmer* (treperenje). On nije pod govornikovim nadzorom. Vrijednosti dobivene u ovom pokusu odgovaraju onima u literaturi koje su dobivene uz loše reproduksijske uređaje (Doherty, Shipp 1988). Naime, ekstrakcija ovog parametra izrazito je osjetljiva na kvalitetu magnetofona.

5. O međuovisnosti visine i jačine glasa i njihovih promjena u vremenu zaključujemo da kod žena veći prosječni intenzitet prati manji tremolo, a kod muških govornika prosječni je ton u lakoj korelaciji s prosječnim intenzitetom, dok kod žena nema povezanosti tih varijabli. Vjerojatno je riječ o akustičkim podacima o šumnosti i hrapavosti glasa koji se različito vrednuju kod muških i ženskih glasova. Zaključak o različitosti strategija proizvodnje glasa autori (Murry, Singh 1978) izvode iz pokusa o procjeni šumnosti i hrapavosti kod muških i ženskih glasova i korelacije tih psihoakustičkih veličina s akustičkim podacima o promjenama osnovnog tona i trajanja vokala. Razliku između muških i ženskih glasova tumače fiziološkim razlikama i kulturološkim čimbenicima (Mattingly 1866 prema Murry, Singh 1978). Treba se prisjetiti da postoje različiti fonatorni tipovi (Laver 1980), razni modovi fonacije koji dovode glasnice u različite muskularne napetosti i oblike glotisa te posljedično proizvode različite akustičke efekte. U svakom je registru kontrola visine tona i intenziteta različita. Kao primjer može se spomenuti korištenje škripavim registrom na kraju silazne intonacije kod muških govornika (Orlikoff, Baken 1990). Ovi podaci svakako mogu potaknuti na dalje istraživanje na tom području.

6. Izračunata duljina vokalnog trakta, što je kod muškaraca za 14,6% veća nego kod žena, odgovara podacima iz literature prema kojima ženski vokalni trakt iznosi 0,87 duljine muškog trakta (Chiba, Kajiyama 1958), gdje se također navodi podatak o 19,3 cm za muške i 16,5 cm za ženske govornike. Nije odgovoreno na pitanje proizlazi li razlika iz nepreciznosti metode ili je skupina ispitanika iz tog istraživanja izrazito visoka. Duljina vokalnog trakta može se samo donekle precizno izračunati. Greška je manja ako se uzmu viši formanti, a postoji podatak u literaturi da pri uzimanju manjeg broja formanta iznosi između 4,4% i 5,6% (Wakita 1977, prema Štamenković 1991). Greška se umanjuje uzimanjem viših formanta ili što većeg broja formanta.

7. Sličnost glasova i u drugim se pokusima procjenjivala iz jedne rečenice (Singh, Murry 1978; Pascal 1984; Pittam 1987) kada je uzorak kontinuirani govor, iako se se različiti perceptivni pokusi procjene boje glasa često rade i na držanom vokalu (Bassich, Ludlow 1986; Horga et al. 1980). Trajanje stimulusa od 3 sek opravdava se podatkom (Bolt, Cooper 1969) da je prepoznavanje glasa slušanjem *steep* funkcija trajanja govornog uzorka do 1,2 sek, a nakon toga je rast prilično spor. Stanka je određena prema uzoru na sličan pokus (Singh, Murry 1978).

Standardne devijacije ocjena sudaca ne razlikuju se bitno kod normalnih i obrnutih snimaka, iako smo pretpostavili da će se suci nenavikli na slušanje govora unatrag manje slagati u ocjenjivanju. Taj podatak pokazuje da sumnjičavost u mogućnosti obavljanja te slušne zadaće nije bila opravdana. Gledajući odvojeno, koeficijent korelacije ocjena za muške glasove veći je od onoga za ženske glasove, što je jednako primjedbama sudaca za vrijeme trajanja pokusa. Ako se koeficijenti usporede s podacima iz literature (Bassich, Ludlow 1986) koji tvrde da se međusudačka povezanost kreće između 0,78 za profesionalce i 0,4 za laike, može se zaključiti da suci-fonetičari nisu bili dovoljno spremni za tu slušnu zadaću.

8. Iz povezivanja akustičkih i psihoakustičkih podataka proizlazi da kod prislićivanja muških glasova slušači doista "slušaju" spektar, što se povezuje s činjenicom eksperimentalno potvrđenom u ovom radu, da među muškim spektrima postoji veći varijabilitet. No kada se snimke slušaju unatrag, onda su za procjenu sličnosti glasa i kod žena i kod muškaraca značajni i spektar i osnovni ton. Iz rečenog bi se moglo pretpostaviti da u odsutnosti nekih elemenata važnih za slušanje i otkrivenih, slušači primjenjuju kod obrnuto reproduciranih ženskih glasova strategiju slušanja kojom su se koristili kod normalne snimke muškog glasa. Čini se, naime, da se kod slušanja boje glasa muških glasova više sluša ona statična, a kod ženskih glasova ona dinamična boja. Tome u prilog govori i podatak o većoj standardnoj devijaciji i koeficijentu varijacije osnovnog tona kod žena, dok je spektar, naprotiv, u ženskom uzorku uniformniji. S tim je u svezi i podatak da su se suci bolje međusobno slagali kod obrnuto reproduciranih snimaka, kod muških glasova. Vjerojatno stoga što su ženski glasovi obrnuto reproducirani izgubili ono što je za strategiju slušanja ženskog glasa važno, pa su slušači bili dezorjentirani. O mogućnosti postojanja različitih strategija slušanja muškog i ženskog glasa pisali su Aronovich (1970) i Singh, Murry (1978). Pokazalo se, također, da ženska statična boja manje varira, kao i da je donekle određena anatomskim čimbenikom, vokalnom duljinom. Iz toga se moglo pretpostaviti da je manje "slobodna" i time manje prikladna da bude presudni element u procjeni sličnosti ili različitosti glasa. Nedostatak prostora za variranje statične boje omogućuje elementima dinamične da budu važni za percepciju ženskog glasa. Kod muških glasova, naprotiv, statična boja više varira i uopće nije povezana s vokalnom duljinom, ali jest značajna za procjenu sličnosti glasa. Možda je s time u vezi i činjenica da muškarci čine većinu pacijenata koji liječe glas zbog njegove loše fokusiranosti (Boone 1991:159). Kada bi se moglo dokazati da se glas "uči" imitacijom, sve navedeno govorilo bi u prilog tome da se kod žena uči ona dinamična, a kod muškaraca ona statična boja.

VI. ZAKLJUČAK

1. Spektar glasova srodnika (majki i kćeri, očeva i sinova) sličniji je nego spektar glasova nesrodnika.
2. Spektar više varira među muškim, nego među ženskim govornicima.
3. Sličniji spektar procjenjuje se slušanjem kao sličniji glas.
4. Glasovi srodnika procjenjuju se slušanjem sličniji nego glasovi nesrodnika samo kod normalnih snimaka.
5. Prosječan osnovni ton, prosječan intenzitet i njihove modulacije nisu sličniji u srodnika, od onih u nesrodnika.
6. Vokalna duljina nije sličnija u srodnika, od onih u nesrodnika.
7. Veća sličnost vokalne duljine ne rezultira većom sličnošću spektra kod muških glasova, a samo donekle kod ženskih glasova.

LITERATURA

- Bassich, C., Ludlow, C.** (1986). *The use of perceptual methods by new clinicians for assessing voice quality.* Journal of Speech Hearing Disorders, 51, 125-133.
- Boone, D.** (1991). *Is your voice telling on you?* San Diego, Singular Publishing Group Inc.
- Colton, R.** (1974). *Factor analysis of voice perception.* Journal of the Acoustical Society of America, 55, 436.
- Colton, R., Estill, J.** (1981). *Elements of voice quality; Perceptual acoustic and physiologic aspects.* Speech and Language – Advances in basic research and practice, 5, 311-403.
- Dalleas, B., Dalleas, A.** (1987). *Qualités acoustiques de la voix.* Revue de Laryngologie, 108, Suppl., 373-377.
- Elbaz, P. et al.** (1987). *Analyse informatique de la voix.* Revue de Laryngologie, 108, Suppl., 383-385.
- Esling, J. et al.** (1991). *Social distribution of long-term average spectral characteristics in Vancouver English.* In: *Proceedings of the XII International Congress of Phonetic Science.* Aix en Provence. s.n., vol. 2, 182-184.
- Frøpkjaer-Jensen, B., Prytz, S.** (1975). *Registration of voice quality.* B. & K. Technical Revue, 3, 3-17.
- Graddol, D., Swann, J.** (1983). *Speaking fundamental frequency; Some physical and social correlates.* Language and Speech, 16, 351-366.
- Günzburger, D.** (1989). *Voice adaptation by transsexuals.* Journal of Clinical Linguistics and Phonetics, 3, 163-172.
- Harmengies, B.** (1988). *SDDD: A new dissimilarity index for the comparison of speech spectra.* Pattern Recognition Letters, 8, 153-158.
- Kitzing, P.** (1986). *LTAS criteria pertinent to the measurement of voice quality.* Journal of Phonetics, 14, 477-482.
- Laver, J.** (1980). *The phonetic description of voice quality.* Cambridge, Cambridge University Press.
- Laver, J.** (1979). *Voice quality; A classified research bibliography.* Amsterdam, John Benjamins.
- Löfquist, A.** (1986). *The long-time average spectrum as a tool in voice research.* Journal of Phonetics, 14, 471-475.

- Löfquist, A., Mandersson, B.** (1987). *Long-time average spectrum of speech and voice analysis*. *Folia Phoniatica*, 39, 221-229.
- Mermelstein, P.** (1966). *Determination of the vocal-tract shape from measured formant frequencies*. *Journal of Acoustical Society of America*, 39, 1283-1294.
- Murry, T., Singh, S.** (1977). *Multidimensional classification of abnormal voice qualities*. *Journal of the Acoustical Society of America*, 61, 1629-1635.
- Pascal, D.** (1984). *Qualite de la voix et caractéristiques individuelles*. *Bulletin de l'Institut Phonétique de Grenoble*, 14, 73-127.
- Pittam, J.** (1987). *Discrimination of the voice qualities and prediction to perceptual ratings*. *Phonetica*, 44, 38-49.
- Singh, S., Murry, T.** (1978). *Multidimensional classification of normal voice qualities*. *Journal of the Acoustical Society of America*, 64, 81-87.
- Scherer, K., Giles, H.** (Ed.). (1979). *Social markers in speech*. Cambridge, Cambridge University Press.
- Wendler, J. et al.** (1986). *Classification of voice qualities*. *Journal of Phonetics*, 14, 483-488.
- Zalewski, J., Majewski, W., Hollien, H.** (1975). *Cross correlation of long-term speech spectra as a speaker identification technique*. *Acustica*, 34, 20-25.

Nina Bertapelle,
Embassy of Spain, Zagreb

HEREDITARY AND SOCIAL FACTORS OF VOICE

SUMMARY

The article deals with static and some elements of dynamic voice quality. Its aim is to verify the common belief that children's voice is very much alike to their parents' one and to find out whether the voice is determined by anatomy or is it learned by acoustic feedback and imitation. Speech recordings of 15 mothers and their 15 daughters as well as 15 fathers and their 15 sons have been analyzed. The direct acoustic measurements of long-term average spectrum, fundamental frequency and intensity have been made, after which the standard deviation and coefficient of variation of both fundamental frequency and intensity have been calculated. Besides, the first three formants from the short-term spectrum of vocal] have been located and taken with considerable accuracy and they were used for the vocal tract length calculation that figures as the only anatomical parameter. By seven-point scale the similarity was rated to the pairs of voices that belong to related/unrelated pairs in two variants: normal and backwards tape playing. The results of spectral slopes correlation show that male voices differ more among themselves than the female voices do. Besides, the greater degree of similarity between the pairs of related than unrelated has been noticed. However, the relatives do not show the greater degree of similarity between themselves concerning the average fundamental frequency, average intensity and their change during period of time. Neither they are alike concerning the anatomic criterion. The psychoacoustic results prove the well done listening task of similarity rating all the voice recordings, except the female backwards recordings. From correlation of acoustic, anatomic and psychoacoustic parameters it has been concluded that listeners, during the similarity ratings of male and female voices use the comparisons of spectrum and fundamental frequency, although when listening female voices, they give less importance to those elements which impose the possibility of existence of some elements that were not isolated in this research, as well as the possibility of existence of different strategies of listening male and female voices.

Key words: voice quality, hereditary factors, social factors
