
UDK 808.62-41

808.63-41

Izvorni znanstveni rad

Prihvaćeno 15.6.1994.

Juraj Bakran,
Filozofski fakultet, Zagreb
Aleš Dobnikar,
Institut Jožef Stefan, Ljubljana

**KONTEKSTUALNO NEOVISAN AKUSTIČKI OPIS MJESTA
ARTIKULACIJE OKLUZIVA: USPOREDNA VIŠEJEZIČNA ANALIZA**

SAŽETAK

Na temelju akustičke analize prijelaza okluziv-vokal u slovenskom i u hrvatskom standardnom govoru, izračunata je takozvana jednadžba lokusa (loкус equation). Ponovo je ustanovljena visoka linearna povezanost frekvencije početka vokalskog tranzijenta drugog formanta nakon okluziva i frekvencije drugog formanta vokala. Jednadžba lokusa time se potvrđuje kao način invarijantnog opisa mjesta artikulacije okluziva.

Ključne riječi: akustička analiza, okluzivi, hrvatski jezik, slovenski jezik

UVOD

Jedan od temeljnih problema koji je usmjeravao istraživanja govora posljednjih pedesetak godina jest odnos kontekstualno uvjetovanog varijabiliteta i potrebe pronalazaženja akustički nepromjenjivih oblika govornog signala. Fonetski segment i njegova apstrakcija višeg reda, fonem, dugo odolijevaju eksplicitnom određivanju. Uzrok je tome netransparenatnost govornog signala uvjetovana koartikulacijskom prirodom govora. Planirani segment za izgovor ima svojstvo da je diskretan i kontekstualno neovisan, a izgovoreni segment, tome nasuprot, karakterizira preklapanje i kontekstualna ovisnost. Odatle teoretski problem: kako je moguća percepcijska stabilnost na temelju neistovjetnih signala? Problem se vrlo lako ilustrira već samo jednim malim segmentom – izgovorom i percepcijom sloga okluziv + vokal.

Jedan od najranijih radova koji su navijestili problem invarijantnosti jest Potter et al.(1947). Tada je ideja hub-a definirana kao mjesto na frekvencijskoj skali na kojemu počinje, u izoliranom izgovoru, drugi formant CV sloga. Zaključeno je da labijalno mjesto artikulacije /p,b/ uzrokuje hub u niskom dijelu spektra, alveolarno-dentalno mjesto artikulacije /t,d/ u srednjem dijelu, ali da velarni glasovi /k,g/ imaju hub vrlo ovisan o vokalu koji slijedi okluziv – uz prednje vokale je u visokom dijelu, uz srednje je u srednjem dijelu spektra, a uz stražnje u niskom dijelu spektra.

Delattre et al. (1955) definirali su *lokus* kao mjesto na frekvencijskoj skali odakle počinje F2 vokala ili na to mjesto smjera. Odnosno, pokusi sa sintezom pokazali su da se bolji rezultati postižu ako se "izbriše" prvih 50 ms tranzijenta, koji počinje za jedan okluziv s istog mjesta.

Stevens i House (1956) pobliže određuju ideju lokusa kao varijabilnu točku na frekvencijskoj skali koja za bilabijale varira od 700 do 1500, za velare 600 do 2500, a relativno je stabilna za alveolare na 2000 Hz.

Ni Lehiste i Peterson (1961) također nisu našli invarijantan lokus za okluzive. Također najmanje varira početna frekvencija F2 za alveolare.

Analizom s pomoću sinteze Stevens et al. (1966) objavili su da je lokus za /d/ oko 1700 Hz i da je vrlo stabilan. Velike varijacije F2 ustanovljene su s labijalima, jer tu jezik ima gotovo potpunu anticipacijsku slobodu.

Fant (1973) smatra da je ideja lokusa suviše pojednostavljena, da se ograničava na sintezu s dva formanta i nema veze s prirodnim govorom.

Neuspjeh pronalaženja invarijantnih svojstava tranzijenata potaknuo je neke istraživače da takva svojstva potraže u drugom dijelu akustičkog signala: Stevens i Blumstein (1978) te Blumstein i Stevens (1979) pokušavaju tražiti invarijantna svojstva okluziva u spektru šuma eksplozije. Njihovi spektralni uzorci prilično su dobro razdvajali okluzive, s 85% točnosti. No kad se suprotstave informacije sadržane u relativno stabilnom šumu eksplozije kontekstualno osjetljivim informacijama sadržanim u tranzijentima, ispitanici sigurnije razdvajaju kategorije prema ovim drugima (Bakran et al. 1991).

Jedno od predloženih rješenja problema jest tzv. jednadžba lokusa (*locus equation*) (Sussman et al. 1991). Pokazalo se da je frekvencija drugog formanta (F2) na početku tranzijenta nakon opuštanja okluzije za /b, d, g/ (onset) linearna funkcija frekvencije drugog formanta u sredini vokala (target). Dakle, ako se podaci akustičke analize unesu u dvodimenzionalni dijagram tako da je na apscisi frekvencija drugog formanta sredine vokala (vremenski), a na ordinati frekvencija početka tranzijenta koja se očitava nakon VOT (*voice onset time*), u trenutku kad počinje dio vokala s pravilnim laringalnim impulsima i razvijenim formantima, dobije se "scatterplot" koji se vrlo uspješno može aproksimirati linearnom funkcijom, pravcem,

$$y = ax + b$$

ili

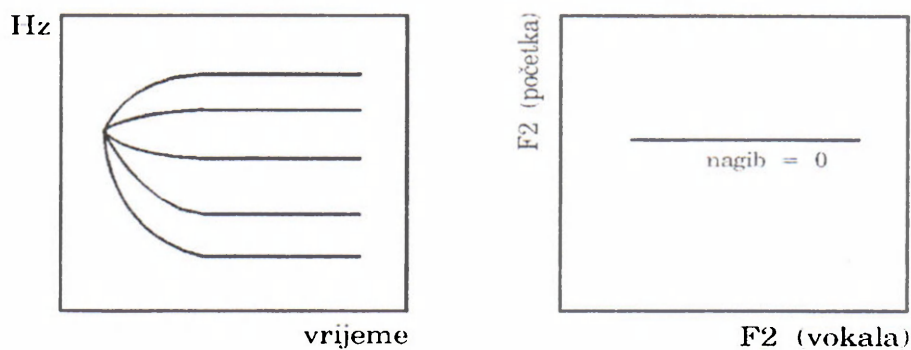
$$F2(\text{početka}) = aF2(\text{vokala}) + b$$

koji je definiran koeficijentom smjera (*slope*) i odsječkom na ordinati. Uspješnost aproksimacije izražava se koeficijentom korelacije. Svaki okluziv /b,d,g/ ima svoju karakterističnu liniju regresije koja se može interpretirati kao način definiranja mjesta artikulacije – drugog reda. Tako je definiran način varijabiliteta i ne mora se uz promijenu vokala navoditi nova definicija mjesta artikulacije pojedinog okluziva. Premda stvarni akustički opis pojedinog okluziva uključuje različite vrijednosti s promjenom vokala, jednadžba lokusa predstavlja invarijantnu definiciju pojedinog mjesta artikulacije.

Jednadžbu lokusa izrazio je prvi Lindblom (1963) na temelju izgovora CVC slogova jednog govornika. Termin lokus mnogo je stariji (Delattre et al. 1955), ali ovdje treba istaknuti da Lindblom nije uzeo Delattrov virtualni lokus, nego stvarno izmjerene vrijednosti frekvencije početka drugog formanta. Klatt (1987) nastojao je iskoristiti jednadžbu lokusa u razvoju algoritma sinteze govora. Načelo očitavanja jednadžbe lokusa možemo ilustrirati primjerom. U slučaju da postoji invarijantan lokus za pojedini okluziv (kao kod Delattre et al. 1955) linija regresije, ako apscisa predstavlja F2 sredine vokala a ordinata tu nepromjenjivu točku lokusa, bit će vodoravna crta na pravcu koji na različitom mjestu siječe ordinatu ovisno o frekvenciji lokusa (sl. 1 a).

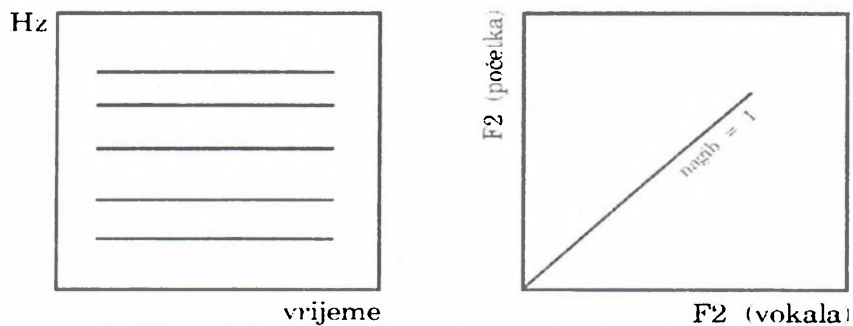
Slika 1. Shematski prikaz raznolikosti jednadžbe lokusa

1.a) nepromjenjivi lokus - nema koartikulacije



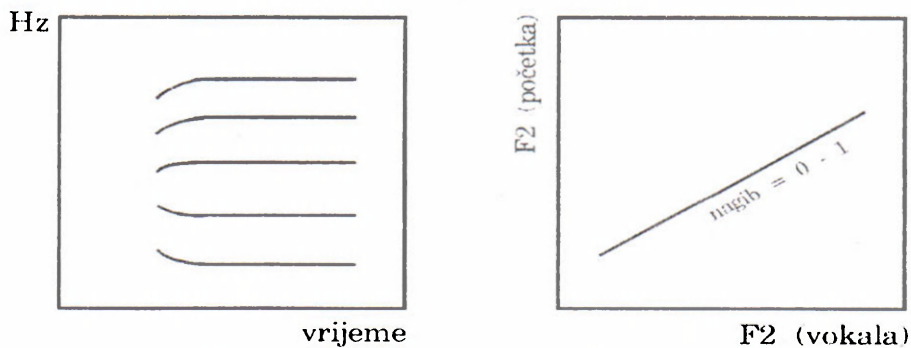
a) Shema jednadžbe lokusa uz stabilni lokus, odnosno bez koartikulacije

1.b) maksimalna koartikulacija



b) Shema jednadžbe lokusa uz maksimalnu koartikulaciju

1.c) stvarna koartikulacija



c) Shema u istom mjerilu jednog konkretnog slučaja

Uz maksimalnu koartikulaciju, kad frekvencija F2 početka vokalskog tranzijenta sasvim slijedi F2 vokala, crta regresije bit će na pravcu koji prolazi kroz ishodište i nagib mu je 1 (sl. 1 b). Stvarne crte regresije definirane na ovakav način bit će između ova dva ekstrema (sl. 1 c).

U jednadžbi lokusa uključena je, osim informacije o mjestu artikulacije, i strategija koartikulacije. U posljednje vrijeme pojavilo se nekoliko radova koji upućuju na specifičnost strategije koartikulacije u različitim jezicima (Smith 1988, 1991; McAlister i Engstrand, 1991; Sussman et al. 1993). Logična je pretpostavka da se u različitim jezicima koji imaju različite repertoare okluziva drukčije ostvaruje CV koartikulacija. S tim u vezi postavlja se pitanje postoje li ipak neka karakteristična mjesta (*hot spots*) za pojedina mjesta artikulacije, ako u novoj ravnini prikazemo na apscisi koeficijent smjera *a* na ordinati odsječak na *y* karakterističnih jednadžbi lokusa.

Sussman et al. (1991) analizirajući jednadžbu lokusa na uzorku od 10 muških i 10 ženskih govornika, uočili su da se jednadžba lokusa za /g/ može razložiti na dvije – jednu uz prednje vokale i jednu za stražnje vokale. Tako izračunate crte regresije pokazuju manje raspršenje. To posebno vrijedi za dio crte uz prednje vokale. Srednje odstupanje od jedinstvene crte regresije iznosi 125 Hz, a izdvojeno za prednje vokale samo 64 Hz. Uz stražnje vokale slično je kao za jedinstvenu crtu regresije – 121 Hz. No jedinstvena se crta regresije bolje (nego posebne za prednje i stražnje vokale) razlikuje od ctra regresije koje predstavljaju druga mjesta artikulacije /b,d/.

Smisao ovog rada jest da na dva još neistražena jezika, slovenskom i hrvatskom, provjeri stupanj linearnosti veze između frekvencija drugog formanta na početku tranzijenta i u sredini vokala. Zatim, da se podaci analize za hrvatski i slovenski prikažu u novom prostoru s pomoću karakterističnih vrijednosti jednadžbi lokusa u usporedbi s jezicima koje opisuju Sussman et al. 1993.

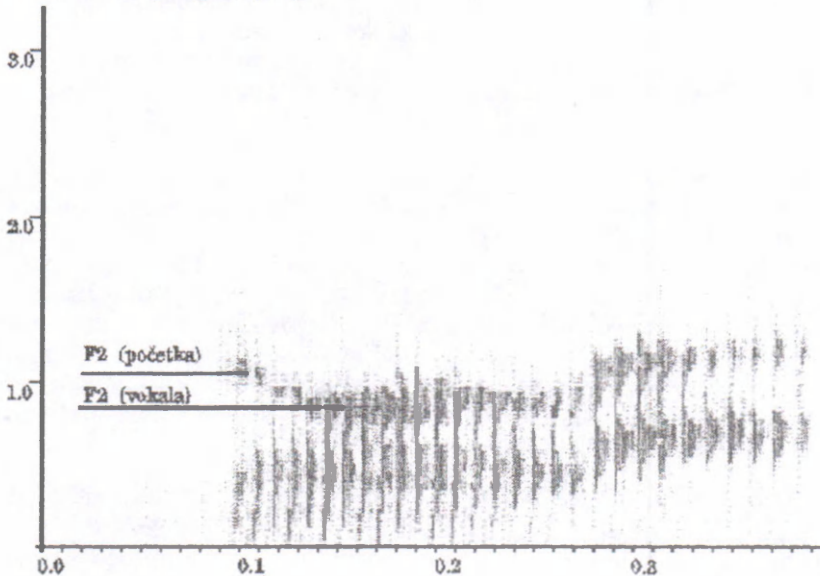
POSTUPAK

A. Ispitanici i stimulusi

Govornici koji su izgovorili riječi koje su se analizirale, neprofesionalni su govornici standardnog hrvatskog i slovenskog govora, studenti Filozofskog fakulteta u Zagrebu i istraživači Instituta "Jožef Stefan" u Ljubljani. Za hrvatski imali smo 5 muških govornika i dva ženska, a za slovenski samo 4 muška govornika. Tekst koji su ispitanici izgovarali sastojao se od popisa dvosložnih riječi koje počinju s /b/, /d/, /g/, a slijede ih svi mogući naglašeni vokali hrvatskog odnosno slovenskog jezika. Hrvatskom je korpusu još pridodan brzi izgovor logoatoma /babababa/, /dadadada/, i /gagagaga/ da se ustanovi utjecaj brzine artikulacije na strategiju koartikulacije opisanu jednadžbom lokusa.

B. Mjerenje

Ispitanici su snimljeni najprije analognom tehnikom a zatim su snimke digitalizirane i analizirane pomoću LSI sustava za analizu govornog signala (*Laughborough Sound Images*, 1991). Frekvencije drugog formanta izmjerene su izravnim očitavanjem širokog spektrograma s pomoću pointera upravljano *mišem*. Mjesto očitavanja frekvencije početka tranzijenta određeno je prvim vidljivim laringalnim impulsom nakon suma eksplzije. Frekvencija F2 sredine vokala očitana je na maksimumu/minimumu konveksnog dijela ili u slučaju uzlaznog/silaznog oblika formanta na sredini ukupnog trajanja.

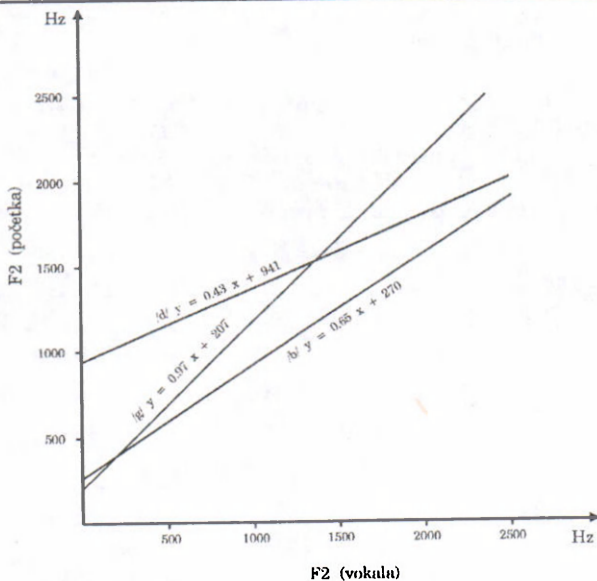


Slika 2. Ilustracija načina mjerenja – (spektrogram)

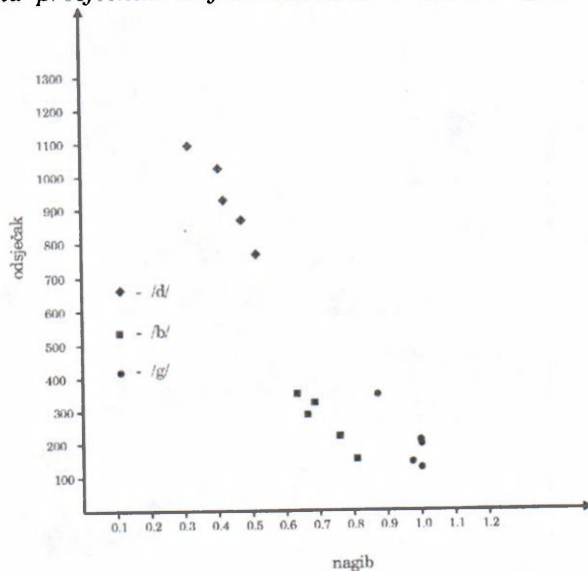
C. Rezultati za hrvatski

Tablica 1. muški govornici

	nagib			odsječak			r ²		
	b	d	g	b	d	g	b	d	g
1	.681	.322	.867	325	1090	352	.977	.858	.947
2	.758	.483	.968	226	877	145	.965	.926	.955
3	.626	.406	1.000	355	1027	128	.909	.879	.947
4	.662	.420	1.024	289	937	198	.959	.886	.890
5	.812	.521	1.004	155	778	214	.980	.953	.912
ar.sr.	.647	.430	.972	270	941	207	.957	.900	.930



Slika 3. Pravci regresije za pojedine okluzive u hrvatskom prema prosječnim vrijednostima za 5 muških govornika



Slika 4. "scatter" dijagram "muških" crta regresije

Analiza varijance za sva mjerenja pokazuje nesumnjivu razdvojenost kategorija po mjestu artkulacije: $F(\text{odsječak})=52.39$ ($p=1.75E-09$), $F(\text{nagib})=74.11$, ($p=5.36E-11$).

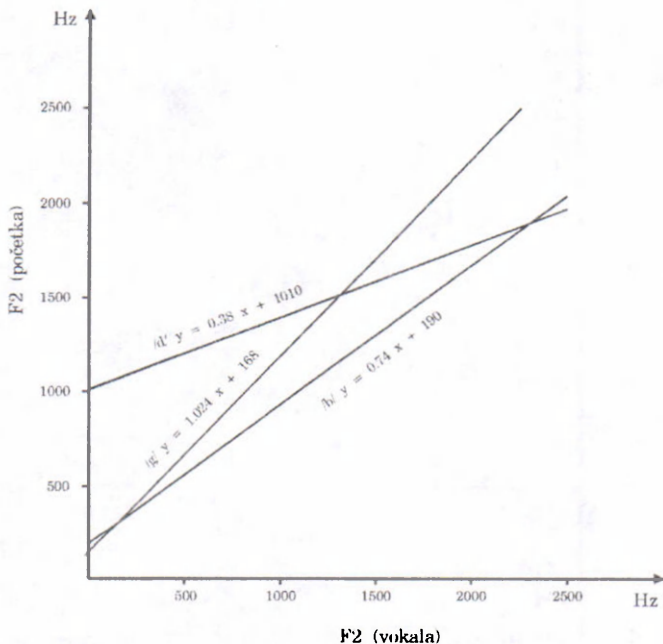
D. Rezultati za slovenski

Korpus za analizu slovenskog govora u ovom istraživanju predstavlja izgovor dvosložnih riječi koje počinju sa /b,d,g/. Četiri izvorna muška govornika (Ljubljana) izgovorilo je po 7 riječi (različiti vokali) za svaki od okluziva, i to u 4 različita redoslijeda. Na temelju 7 riječi i 4 ponavljanja za svakoga govornika posebno se izračunala linija regresije:

Tablica 2.

govornici	nagib			odsječak			r ²		
	b	d	g	b	d	g	b	d	g
1	.755	.303	1.011	227	1186	150	.953	.752	.855
2	.659	.488	1.093	217	780	44	.902	.867	.949
3	.696	.262	.943	236	1225	336	.952	.617	.808
4	.842	.451	1.054	80	849	143	.968	.895	.931
ar.sr.	.738	.376	1.024	190	1010	168	.944	.780	.885

Napomena: Nagib jednadžbe lokusa, kako je u uvodu spomenuto, može biti između 0 i 1. Maksimalna koartikulacija ne može prouzrokovati veću vrijednost od 1 ako ideja lokusa uopće ima smisla. Veće vrijednosti od jedan u opisu rezultata ilustriraju nepreciznost mjerenja.

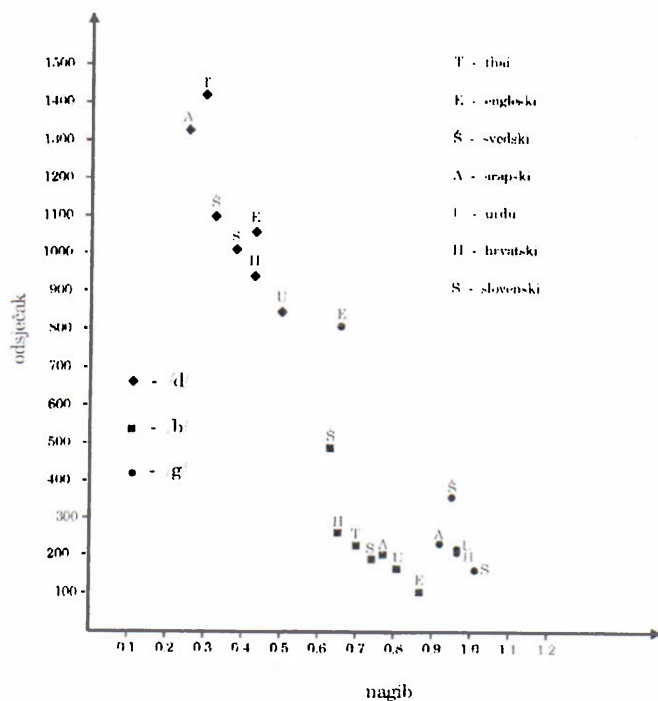


Slika 5. Crte regresije koje predstavljaju srednje vrijednosti za slovenski.

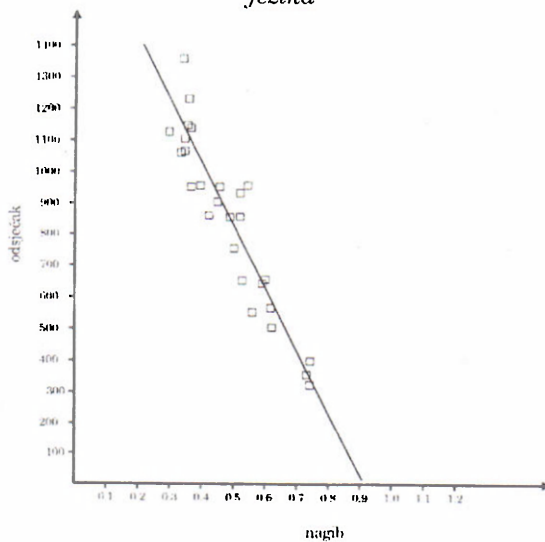
E. Usporedba s drugim jezicima

Tablica 3. Zajednička tablica s rezultatima za druge jezike
(Sussman et al. 1993)

okluziv	jezik govor	nagib	odsječak
b	thai	.70	228
b	engleski	.87	106
b	švedski	.63	487
b	arapski	.77	206
b	urdu	.81	172
b	hrvatski	.65	270
b	slovenski	.74	190
d	thai	.30	1425
d	engleski	.43	1073
d	švedski	.32	1096
d	arapski	.25	1307
d	urdu	.50	857
d	hrvatski	.43	941
d	slovenski	.38	1010
g	thai		
g	engleski	.66	807
g	švedski	.95	360
g	arapski	.92	229
g	urdu	.97	212
g	hrvatski	.97	207
g	slovenski	1.02	168



Slika 6. Zajednički "scatter" dijagram za svih 7 opisanih jezika



Slika 7. "Scatter" dijagram parametara jednadžbi lokusa pojedinih govornika za hrvatski standardni govor, uključujući ženske govornike i brzu artikulaciju.

F. Priroda razlike koartikulacijskih strategija

Pomnijim promatranjem jednadžbi lokusa pojedinih govornika, čija različitost zapravo odražava različitost individualnih strategija koartikulacije, lako se može uočiti da se u ravnini nagib/odsječak pojedini rezultati ne raspršuju slučajno oko neke srednje vrijednosti, nego da tvore jasan izdužen "oblak", koji na ovoj još višoj razini apstrakcije sugerira linearnu povezanost tih dviju dimenzija – parametara jednadžbe lokusa. Na slici 7. prikazani su na istom dijagramu rezultati analize te izračunati parametri crta regresije samo za /d/ pojedinih izgovora muških i ženskih govornika uključujući i različitu brzinu artikulacije (ukupno 27 crta regresije).

Takav oblik povezanosti parametara jednadžbi lokusa navodi na zaključak da jednostavna aritmetička sredina, odnosno prosječan nagib i prosječan odsječak na ordinati, nisu dovoljno dobar način prikaza. Ovako raspršeni podaci bolje će se predstaviti novom crtom regresije, pravcem koji najbolje pristaje svim pojedinim vrijednostima. Za podatke na slici 7. to je pravac $y = -1918x + 1782$. U ovom radu nećemo potanje analizirati uzroke takvom raspršenju. Svakako postoji povezanost brzine artikulacije s načinom koartikulacije: uz brži tempo artikulacije nagib crta regresije smanjuje se, a odsječak na osi y povećava se. To je otprilike poznati fenomen "podbačaja" (*undershoot*).

Tablica 4. Alternativni način izražavanja srednjih vrijednosti za hrvatski i slovenski

hrvatski			
	nagib	odsječak	r ²
b	-1.23	1139	.87
d	-1.95	1794	.92
g	-1.15	1294	.73
slovenski			
	nagib	odsječak	r ²
b	-.78	764	.70
d	-2.06	1784	.99
g	-1.85	2069	.92

Na ovaj se način može opisati rezultate analize za engleske okluzive prema Sussman et al. (1991).

Pojedine crte regresije izračunate su na temelju analize izgovora 10 muških i 10 ženskih govornika. Opravdanost takvog prikazivanja može se procijeniti prema koeficijentu korelacije (r²).

Tablica 5. Rezultati Sussman et al. (1991) izraženi s pomoću novih crta regresije umjesto aritmetičkim sredinama.

	b			d			g		
	odsječak	nagib	r ²	odsječak	nagib	r ²	odsječak	nagib	r ²
muški	1479	-1.60	.74	2157	-2.53	.90	2117	-1.98	.90
ženski	1593	-1.66	.63	2273	-2.27	.85	2345	-2.08	.96
svi	1366	-1.44	.62	2309	-2.63	.63	1955	-1.64	.77

RASPRAVA I ZAKLJUČAK

Osnovna je namjera ovog rada bila da se provjeri primjerenost jednadžbe lokusa za definiciju invarijantnih svojstava tranzijenata spoja okluziva i vokala te da se rezultati analize hrvatskog i slovenskoga standardnog govora usporede s drugim jezicima. Nedvojbeno možemo zaključiti da su rezultati analize slovenskog i hrvatskog standardnoga govora potvrdili visoku linearnu povezanost F2 početka tranzijenta i F2 vokala, koja je izražena visokim koeficijentom korelacije. Prema našim rezultatima niti za /g/ ne može se opravdati neka drukčija vrsta povezanosti. Analiza varijance koja se izračunala na temelju odsječka na ordinati ili koeficijenta smjera jednadžbe lokusa nedvojbeno razdvaja kategorije različitog mjesta artikulacije. Usprkos specifičnim strategijama koartikulacije različitih jezika, smatramo da je jednadžba lokusa dovoljno otporna da može poslužiti za fonetski opis kategorije okluziva. Prednost opisa jednadžbom lokusa sastoji se u tome da se ne opisuje pojedinačna realizacija CV sloga, nego pojedino mjesto artikulacije kroz cijeli vokalski prostor odabranog jezika (govora). Ako se koeficijentom smjera jednadžbe lokusa i odsječkom na y osi definira novi prostor, u tom se prostoru na različitim mjestima smještaju pojedine alofonske varijante, individualne koartikulacijske karakteristike i razlike među jezicima. No pažljivijim promatranjem dijagrama na slici 7. s rezultatima pojedinih govornika, može se uočiti da se oni ne raspršuju simetrično u ravnini nagib odsječak, nego da tvore izdužen "oblak" kojeg će bolje predstavljati pravac nego jednostavna srednja vrijednost¹.

(1) Računarsku opremu osigurao je FELMET iz Zagreba

REFERENCIJE

- Bakran, J., D. Horga, M. Stamenković** (1991). *Percepcija mjesta artikulacije bezvučnih okluziva*. Govor, 8/9, 1-31.
- Blumstein, S.E., K.N. Stevens** (1979). *Acoustic invariance in speech production: Evidence from measurements of the spectral characteristics of stop consonants*. JASA 66, 1001-1017.
- Blumstein, S.E., K.N. Stevens** (1980). *Perceptual invariance and onset spectra for stop consonants in different vowel environments*. J. Acoust. Soc. Am. 67, 648-662.
- Delattre, P.C., A.M. Liberman, F.S. Cooper** (1955). *Acoustic loci and transitional cues for consonants*. J. Acoust. Soc. Am. 27, 769-773.
- Fant, G.** (1973). *Stops in CV-syllables, in Speech Sounds and Features*. Edited by G. Fant, MIT, Cambridge, MA.
- Klatt, D.H.** (1987). *Review of text-to-speech conversion for English*. J. Acoust. So. Am. 82, 737-793.
- Lahiri, A., L. Gewirth, S.E. Blumstein** (1984). *A reconsideration of acoustic invariance for place of articulation in diffuse stop consonants: Evidence of cross-language study*. J. Acoust. Soc. Am. 76, 391-404.
- Lehiste, I., G.E. Peterson** (1961). *Transitions, glides and diphthongs*. J. Acoust. Soc. Am. 33, 268-277.
- Lindblom, B.** (1963). *On vowel reduction*. Rep. No.29, The Royal Institute of Technology, Speech Transmission Laboratory, Stockholm, Sweden.
- Llisterri J., West M.** (1987). *Invariant acoustic correlates for place of articulation in Catalan voiceless stops*. Proceedings of XI ICPhS Tallin vol 4, 44-47.
- McAlister, R., O. Engstrand** (1991). *Some cross language aspects of coarticulation*. Proceedings of the XIIth Congress of Phonetic Sciences, Aix en Province, France, 5, 18-21.
- Poch-Olive D.** (1987). *Acoustic correlates for places of articulation in Spanish stop consonants*. Proceedings of XI ICPhS Tallin vol 4, 44-47.
- Potter, R.K., G.A. Kopp, H. Green** (1947). *Visible speech*. Van Nostrand, Princeton, NJ.
- Smith, C.L.** (1988). *A cross-linguistic contrast in consonant and vowel timing*. J. Acoust. Soc. Am. 86, S 84.

- Smith, C.L.** (1991). *The timing of vowel and consonant gestures in Italian and Japanese*. Proceedings of the XIIth Congress of Phonetic Sciences, Aix en Province, France, 4, 234-237.
- Stevens, K.N., S.E. Blumstein** (1978). *Invariant cues for place of articulation*. J. Acoust. Soc. Am. 64, 1358-1369.
- Stevens, K.N. & A.S. House** (1956). *Studies of formant transitions using a vocal tract analog*. J. Acoust. Soc. Am. 28, 578-585.
- Stevens, K.N., A.S. House and A.P. Paul** (1966). *Acoustical description of syllabic nuclei: An interpretation in terms of a dynamic model of articulation*. J. Acoust. Soc. Am. 40, 123-132.
- Sussman, H.M., H.A. McCaffrey, S.A. Matthews** (1991). *An investigation of locus equations as a source of relational invariance for stop place categorization*. J. Acoust. Soc. Am. 90, 1309 - 1325.
- Sussman, H.M., K.A. Hoemeke, F.S. Ahmed** (1993). *A cross-linguistic investigation of locus equations as a phonetic descriptor for place of articulation*. J. Acoust. Soc. Am. 94, 1256 - 1268.
- Van Son, R.J.J.H., L.C.W. Pols** (1990). *Formant frequencies of Dutch vowels in a text, read at normal and fast rate*. J. Acoust. Soc. Am. 88, 1683-1693.

Juraj Bakran,
Faculty of Philosophy, Zagreb
and
Aleš Dobnikar,
Institute of Jožef Stefan, Ljubljana

CONTEXTUALLY INDEPENDENT ACOUSTIC DESCRIPTION
OF THE OCCLUSIVES POINT OF ARTICULATION:
COMPARATIVE MULTILINGUAL ANALYSIS

SUMMARY

On the basis of acoustic analysis of the transition of occlusive to vowel in Slovenian and Croatian standard speech forms, the locus equation has been calculated. A high linear connection of the frequency of the beginning of vowel transient of the second formant following the occlusive and the frequency of the second formant of the vowel has been confirmed again. The locus equation is therefore confirmed as a way of invariant description of the point of articulation of occlusives.

Key words: *acoustic analysis, occlusives, Croatian language, Slovene language*
