
UDK 801.41:534.63-7
Originalni znanstveni rad

Primljeno: 13. 05. 1990.

Juraj Bakran
Filozofski fakultet, Zagreb

DJELOVANJE NAGLASKA I DUŽINE NA FREKVENCIJE FORMANATA VOKALA

SAŽETAK

Klasična predodžba akustičkog vokalskog prostora prikazanog u ravnini $F_1 \times F_2$ opterećena je s više izvora varijabiliteta: razlike između govornika, stil govorenja, neposredan fonetski kontekst itd. Manje istraživani izvori varijabiliteta - tempo artikulacije i naglasak - podijelili su istraživače do kontradiktornih pretpostavki. U ovom se istraživanju na specifičnom korpusu hrvatskoga standardnoga govora pokušava razriješiti da li trajanje ili naglašenost vokala "po sebi" bitno određuju stupanj organiziranosti, odnosno redukcije, vokalskog timbra. Uz neka ograničenja rezultati afirmiraju u posljednje vrijeme često osporavanu tezu Lindbloma da skraćeno trajanje rezultira većim stupnjem redukcije (undershoot) vokala.

UVOD

Bitno u akustičkom opisu vokala jest definirati frekvencije formanta. Prve analize iznosile su frekvencije prvih dvaju ili prvih triju formanta (F1, F2, F3) izmjerene spektralnom analizom kvazi stacionarnog dijela izoliranih ili u naročitoj "neutralnoj" kontekstu izgovoreni vokala (Peterson i Barney, 1952). Iz analize frekvencija formanta proizašla je teorija prema kojoj prva dva formanta sadrže dovoljno informacija za identifikaciju vokala (Delattre et al. 1952). Na planu artikulacije formanti vokala jesu rezonantne frekvencije artikulacijskog trakta. Na taj način formanti objedinjuju artikulacijski, akustički i perceptivni aspekt vokala. Frekvencije formanta nisu nepromjenljive veličine. U klasičnom prikazu u ravni F1xF2 vokali fonološki jednoznačno kategorizirani zauzimaju preklapajuća polja. Uzroci njihovu varijabilitetu su mnogostruki. Najizraženiji je varijabilitet izgovora različitih govornika (u istom govornom standardu), a pogotovo ako su razne dobi i spola. Prihvatimo li izražavanje varijabiliteta u postocima kao:

$$\% \text{ promijene} = 100 \left(\frac{x}{V_{\text{ref}}} - 1 \right) \quad (1)$$

x = izmjerena frekvencija, V_{ref} = referentna frekvencija) frekvencije formanta u izgovoru djece otprilike su 30% više nego u izgovoru odraslih muških govornika (prema Peterson i Barney, 1952) do 100% više (Lieberman i Blumstein, 1988).

Odatle proizlazi kontinuirano nastojanje da se istraži normalizacija vokalskog prostora, što bi onemogućilo da se definiraju invarijantne vokalske karakteristike (za pregled vidi Miller, 1989).

Na redukciju vokala znatno utječe stil govorenja. Većina istraživača radi s vrlo kontroliranim tekstom u laboratorijskim uvjetima. Formalniji govor (čitanje test materijala) znatno se razlikuje od spontanoga, neformalnoga govora u mnogim aspektima pa i znatno manjom redukcijom vokala (Van Bergem i Beinum, 1989).

Nešto manji, ali još uvijek snažan izvor varijabiliteta frekvencija formanta vokala jest konsonantski kontekst. U načelu formanti vokala koji su okruženi konsonantima ne dostižu frekvencije izolirano izgovoreni (produženi) vokala. Prema Stevens i House (1963) varijacije F1 uzrokovane konsonantskim kontekstom jesu oko 13%, a varijacije F2 oko 20%. Primjerice prednji konsonanti (labijali i alveolari) spuštaju F2 prednjih vokala, a dižu F2 stražnjih vokala. Stražnji konsonanti dižu F2 svim vokalima ali najmanje prednjim. Radovi o tom problemu i u vezi s time radovi o definiranju perceptivno relevantnih elemenata najbrojniji su (Broad i Fertig 1970, Stevens et al. 1966, Broad i Clermont 1987, Nearey 1989, Strange 1989).

Utjecaj naglaska i tempa govorenja najmanje je istraživani izvor varijacija akustičkog aspekta govora pa i rasporeda formanta. Varijacije tempa artikulacije (TA) unutar određenih granica (4 do 8 slogova u sekundi) čini se da ne djeluju na uspješnost percepcije iako mnoga istraživanja pokazuju da se

akustički oblik govornog signala značajno mijenja ovisno o razlikama u tempu artikulacije (Gay i Hirose, 1973, Gay et al. 1974, Bakran 1984).

Gay (1978) akustičkom analizom nastoji pokazati kako se akustički oblik govora mijenja u funkciji različitog tempa artikulacije. Mjeri trajanje, Fo, intenzitet i formante segmenata uz različiti TA. Neke od njegovih rezultata (u vezi s trajanjem segmenata) komentirali smo prije (Bakran 1984). Analizirajući spektrografske snimke izgovora vrlo ograničenog korpusa, Gay zaključuje da se "target" frekvencije prvih dvaju formanta ne mijenjaju bitno uz promijenjen tempo artikulacije ali mijenja se početna frekvencija tranzijenta, što znači da se brži TA kompenzira ranijim pokretom prema vokalu.

Promatrajući istodobno utjecaj različitog tempa artikulacije i naglašenosti vokala Gay (1978) zaključuje da i TA i naglašenost djeluju na trajanje segmenata, ali da razlike u trajanju ne uzrokuju promjene vokalske boje. Lindblom (1963) zaključuje da tzv. "undershoot" uzrokuje upravo nedovoljno trajanje vokala, a ne naglasak. Van Son i Pols (1989) eksperimentiraju s povezanim govorom (normalnog i brzog tempa artikulacije) profesionalnog spikera i zaključuju, suprotno Lindblomu, da trajanje vokala (regulirano različitim tempom artikulacije) minimalno korelira s frekvencijama formantata i to za čitavog trajanja vokala, pa prema tome redukcije (i "undershoot") nisu rezultat fizioloških ograničenja govornika nametnutih povećanim TA, nego su jezično upravljana svojstva govora.

Namjera je ovog rada da, koristeći se nekim svojstvima hrvatskog standardnog govora pridonese da se riješi dilema: što je presudno za redukciju vokalske boje - trajanje ili naglašenost. U HSG svih se pet vokala ravnopravno mogu naći u naglašenom položaju i to pod dugim i kratkim naglaskom. Za razliku od metodologije spomenutih istraživača trajanje vokala neće biti regulirano namjernim djelovanjem govornika, nego upotrebom fonološki dugih i kratkih vokala i efekta dužine izgovorne cjeline na tempo artikulacije (duža izgovorna cjelina spontano se izgovara bržim tempom artikulacije - Bakran, 1984).

METODA

Govorni materijal i govornici

Analizirane su tri vrste govornog materijala: povezan govor (tekst radijskih vijesti) i liste "hrvatskolikih" logatoma u izgovoru dvojice profesionalnih spikera RTZ (VM i FR) te liste riječi govorne audiometrije koje je čitao zagrebački akademski glumac (IG). Tekst za analizu snimljen je analognom studijskom tehnikom (NAGRA III, Sennheiser MKH 404).

Digitalna obrada

Analogne magnetofonske snimke filtrirane su niskopropusnim filtrom 4.5 kHz i digitalizirane 12-bitnim A/D konvertorom pri frekvenciji uzorkovanja $F_u=10$ kHz. Spektralna analiza obavljena je FFT transformacijom u $N=128$

točaka uz primjenu Hammingovog prozora. Budući da korak spektralne analize za određene parametre uzrokovanja i broj FFT točaka ($\Delta F=40\text{Hz}$) nisu neposredno primjereni prikazu formanta, primijenjeno je linearno centralno gladenje (Stamenković et al. 1990):

$$F_i = \frac{1}{2k+1} \sum_{j=-k}^k f_{i+j} \quad (2)$$

gdje je f_i vrijednost i -te spektralne komponente (frekvencija $i \neq 5000/128\text{Hz}$). Dobivena je efektivna širina filtera $\Delta F'=3\Delta F=120\text{Hz}$ uz $k=1$. Na taj je način onemogućeno razlučivanje formanta s točnošću $\Delta F=40\text{Hz}$.

Problem definiranja formanta

Frekvencije formanta nije moguće egzaktno odrediti i to ne samo zbog dinamike govora nego i zbog relativno labave definicije. Instrumentarij se ne pojavljuje kao usko grlo analize. Formanti kao rezonantne frekvencije artikulacijskog trakta nisu vidljivi i mjerljivi direktno nego posredno u djelovanju na "izvor" - laringalni (harmoničan) ton ili šum. Akustičkom analizom prirodnoga govora ta dva elementa nije moguće izdvojiti. Zbog toga se nerado rade eksperimenti sa ženskim govornicima. Naime, harmonici ženskih glasova dvostruko su razmaknuti od muških što povećava nesigurnost procjene "istaknutog mjesta" u spektru. Za definiranje frekvencije formanta optimalan bi bio kontinuiran linearni spektar izvora. Posebno je teško identificirati F1 kad se približava F0 i F2 kad se približava F3, jer se stapaju u istu "masu". Nije moguće poštovati formalni kriterij "prvo intenzitetski istaknuto mjesto na spektru", jer često F0 ima intenzitetski višu razinu od F1. Rezultat analize prema tome ne ovisi samo o varijabilitetu fenomena i instrumentariju, nego također o subjektivnoj procjeni, iskustvu autora.

EKSPERIMENT I

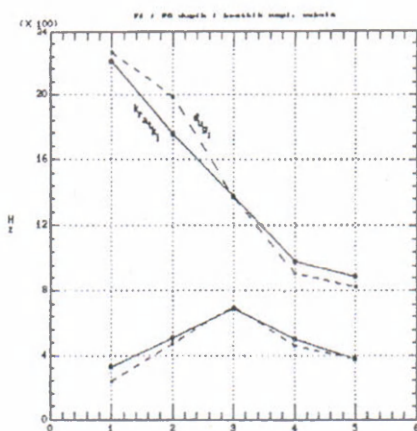
Utjecaj dužine naglašanih vokala na frekvencije formanta

Analizirani materijal bile su liste dvosložnih riječi (originalno snimljene za govornu audiometriju). Govornik je bio zagrebački akademski glumac IG.D/Kontrolirana varijabla samo je dužina naglašenog vokala. Silaznost - uzlaznost akcenta nije posebno promatrana. Zanimaren je i utjecaj konsonantskog konteksta.

U tablici 1. i na dijagramu slika 1. prikazane su srednje vrijednosti F1 i F2 dugih i kratkih naglašanih vokala jednoga govornika, koje su dobivene na temelju 15 do 30 mjerenja za svaki podatak.

Tablica 1.

		i		e		a		o		u	
		F1	F2	F1	F2	F1	F2	F1	F2	F1	F2
dugi	M	238	2264	471	1989	707	1371	461	905	379	822
	S	47	52	43	65	33	102	16	63	52	86
kratki	M	329	2213	511	1754	691	1374	499	983	375	886
	S	101	73	65	102	35	104	49	129	71	98



slika 1

Kao što se iz dijagrama vidi dugi naglašeni vokali artikuliraju se ekstremnije: prednji /i,e/ imaju razmaknutije formante (F1 poprima niže vrijednosti, a F2 više vrijednosti), a stražnji vokali imaju niže frekvencije obaju formantata. Vokal /a/ u izgovoru našega govornika nije pomicao frekvencije formantata. Međutim, statistički se znatno razlikuju samo F1 dugog i kratkog naglašenog /i/ ($t=2.93$) te F2 dugog i kratkog /e/ ($t=3.05$). Dobiveni rezultati slažu se s predviđanjem proizašlim iz teorije podbačaja ("undershoot") te s rezultatima istraživanja prema kojima se uz brži tempo artikulacije frekvencije formantata približavaju neutralnom položaju (Lindblom, 1963).

EKSPERIMENT II

Razlika u frekvencijama formantata naglašenih i nenaglašenih vokala u povezanom govoru

Materijal za analizu jest tekst radio vijesti u izgovoru profesionalnog spikera RTZ (FR). Tempo artikulacije analiziranog teksta bio je relativno brz, 7.5

slogova u sekundi.

Kontrolirana varijabla samo je naglašenost - nenaglašenost vokala. Ni dužina ni silaznost - uzlaznost nisu kontrolirani.

Spektar je analiziran samo u središnjem dijelu vokala, što je bilo moguće jasno izdvojiti.

Srednje vrijednosti su određene na temelju 10 (kratko naglašeni /c/) do 20 (nenaglašeni /a/) mjerenja.

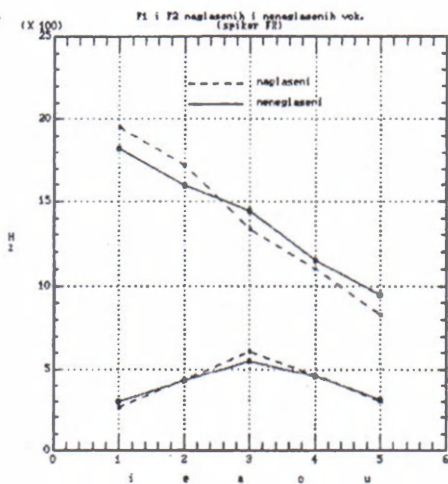
Rezultati su iskazani u tablici 2. i na slici 2.

Tablica 2.

		i		e		a		o		u	
		F1	F2	F1	F2	F1	F2	F1	F2	F1	F2
naglašeni	M	263	1956	439	1720	707	1333	462	1102	296	828
	S	50	89	104	130	51	120	79	158	28	195
nenaglašeni	M	299	1823	428	1594	553	1445	456	1149	310	943
	S	43	116	89	78	51	169	85	104	44	208

t=3.01

t=2.78 t=3.01 t=2.12



slika 2.

EKSPERIMENT III

Utjecaj tempa artikulacije na frekvencije formanta

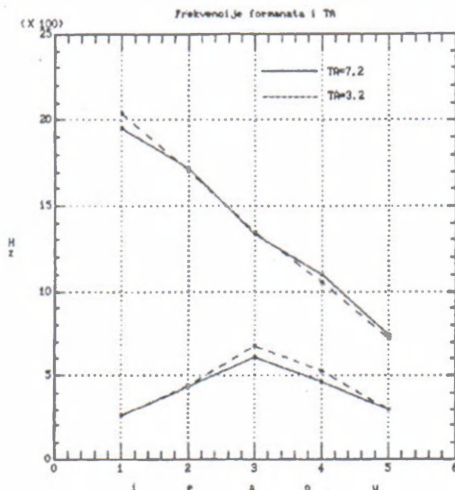
Izmjerene su F2(i, e) u dvosložnim logatomima s kratkim naglaskom (10 mjerenja za svaki vokal). TA povezanog govora jest 7.2 sloga u sekundi, a TA izoliranih logatoma 3.2 sloga u sekundi. Izolirani izgovor upotrijebljen je ovdje namjerno da se postigne spor izgovor (Bakran, 1984).

Usporedba s frekvencijama naglašanih slogova iz povezanog teksta iznesena je u tablici 3. i prikazana je na dijagramu slika 3.

Tablica 3.

		i		e		a		o		u	
		F1	F2	F1	F2	F1	F2	F1	F2	F1	F2
brzo	M	263	1956	439	1720	607	1347	462	1102	296	828
	S	50	89	104	130	51	120	79	158	28	195
sporo	M	300	2038	428	1704	679	1345	531	1054	310	720
	S	43	68	89	74	26	115	85	20	44	79

$t=2.37$ $t=4.38$ $t=2.81$
 $p=0.03$ $p=0.003$ $p=0.01$



slika3

Prema dijagramu vidi se sličan trend kao u eksperimentu Uz sporiji tempo artikulacije frekvencije formanta pomiču se prema ekstremnijim vrijednostima.

Vrijednosti F1(u), F2(u) i F1(i) nisu pouzdane zbog problema mjerenja opisanih u poglavlju o metodi. Ispod tablice označene su statistički značajne razlike u promjeni formantskih frekvencija.

EKSPERIMENT IV

Da bi se provjerilo znače li dobivene različite frekvencije formana degradaciju vokala ili su to perceptivno irelevantne varijante, proveden je eksperiment identifikacije.

Metodološki problem jest u tome što kategoričko opredjeljivanje izvornih govornika u vokalskom sistemu od samo pet vokala funkcionira i uz znatan stupanj dezorganizacije vokala. Zbog toga je trebalo dodatno otežati identifikaciju vokalske boje ali "linearno", tj. svim vokalima podjednako. Iz lista riječi (spiker IG) izdvojena su po četiri vokala s različitim formantima. Prosječna razumljivost tako izdvojenih vokala bila je 93%. Od svakog su vokala zatim izdvojeni dijelovi u trajanju od 5, 3, 2, i 1 perioda (trajanje isječaka uz F0 oko 100 Hz jest 50 do 10 ms). Ispitanici (studenti fonetike) nastojali su identificirati vokalsku boju tih kratkih isječaka. Prethodno im je objašnjeno kako su stimuli kreirani, a oni su trebali zanemariti neprirodan zvuk stimulusa te svakako se opredijeliti za jedan od pet vokala. U tablici 4. navedeni su formanti odabranih vokala i prosječan postotak identifikacije:

Tablica 4.

	F1	F2	% i dent.
i1	236	2283	99
i2	275	2165	74
i3	236	2244	98
i4	275	2244	100
e1	472	1968	95
e2	440	1653	73
e3	433	1968	95
e4	511	1732	95
a1	708	1220	100
a2	748	1456	99
a3	708	1417	99
a4	708	1300	96
o1	472	866	97
o2	472	940	89

o3	472	826	94
o4	472	984	96
u1	433	944	58
u2	433	866	63
u3	433	940	82
u4	315	670	99

Na prvi se pogled vidi da razlike u razumljivosti vokala /a/ unatoč različitim frekvencijama formantata nisu značajne. Ako ostale vokale klasificiramo samo prema frekvenciji drugog formanta i to prema udaljenosti od /a/ vidi se da bolji postotak identifikacije imaju vokali kojima je F2 udaljeniji od /a/. Koeficijent korelacije između frekvencijske udaljenosti F2 od /a/ dobivene prema:

$$dF = F2 (i 1 \dots u4) - 1400 \quad (3)$$

i stupnja identifikacije (postotak iz tablice 4.) iznosi

$$\text{za } i1 \dots i4 \quad r=0.926$$

$$\text{za } e1 \dots e4 \quad r=0.728$$

$$\text{za } o1 \dots o4 \quad r=0.176$$

$$\text{za } u1 \dots u4 \quad r=0.759$$

Prema ovim rezultatima može se zaključiti da su različite realizacije istog vokala ne samo prema akustičkoj analizi nego i perceptivno bitno različite (bez obzira na to što za /a/ i /o/ nismo dobili značajnu razliku) i to je suprotno procjeni Gaya (1978) i rezultatima koje objavljuju Van Son i Pols (1989).

DISKUSIJA

Ako je istina da vokali trebaju određeno vrijeme kako bi se adekvatno realizirali (premda oni ne spadaju u grupu vremenski strukturiranih glasova), nije jasno zašto govornici ne uzmu dovoljno vremena za njihovu realizaciju. Paradoksalno je da je redukcija najizraženija upravo u spontanom govoru kad govornici proizvoljno određuju i mijenjaju tempo artikulacije (Van Bergem i Beinum, 1989).

U analizi veličine i mogućih uzroka redukcije ne treba zaboraviti da je osnovni cilj govora prijenos informacije i da slušači imaju na raspolaganju više simultanih izvora.

Interesantna je hipoteza (Van Son i Pols, 1989) da povećan koartikulacijski efekt može značiti raspoređivanje informacije o jednom segmentu na više njih i time olakšava razumijevanje. Spektralna redukcija također ne mora se nužno shvatiti kao degradacija govora nego kao signal nenaglašenosti, dakle kao prozodijska kategorija.

Treba imati na umu da u eksperimentima s različitim tempom artikulacije, posljedica čega su novi odnosi trajanja segmenata, nemamo prirodnu situaciju. Profesionalni govornici, nakon upute da variraju tempo artikulacije, zadržavaju

približno isti stupanj organizacije govora, to jest sličnu razinu spektralne redukcije vokala (Van Son i Pols, 1989). Prema svemu sudeći trajanje i redukcija vokalske boje posljedica su zajedničkog uzroka. Tako možemo razumijeti povezanost tih dvaju atributa. Pretpostavimo li da je "reguliranje" trajanja primijenjeno u odabiru analiziranog korpusa zaista samo promjena samo jedne dimenzije (to jest da se dugi i kratki naglašeni vokali razlikuju samo po trajanju), rezultati opisanih mjerenja idu u prilog tezi da je trajanje u korelaciji sa stupnjem izgrađenosti vokala.

Premda se tezi da je trajanje presudan uzrok koji regulira redukciju vokalskog spektra (Lindblom 1963) suprotstavlja niz novijih istraživanja (Gay, 1978; Engstrand, 1988; Van Son i Pols, 1989; i djelomice Van Bergem i Beinum, 1989) ovdje izneseni rezultati mjerenja vokalskih formanata u hrvatskom standardnom govoru afirmiraju Lindblomovu tezu. U protivnom moglo bi se primjerice zaključiti da su dugi naglašeni vokali u hrvatskom "naglašeniji" od kratkih, jer se kod njih registrira manja redukcija frekvencija formanata.

ZAKLJUČAK

U opisanim eksperimentima (I-IV) provjeren je utjecaj naglašenosti i dužine vokala na frekvencije prvih dvaju formanata u hrvatskom standardnom govoru, te odraz tog utjecaja na identifikaciju vokala.

Vokali u naglašenom slogu manje su reducirani od vokala u nenaglašenom slogu. (Redukcijom, ovdje u spektralnom smislu, nazivamo pomak frekvencija formanata prema neutralnom shwa). Dugi naglašeni vokali manje su reducirani od kratkih naglašenih. Izgovor u sporijem tempu manje reducira vokalsku boju nego izgovor u brzom tempu (uz rezervu da su uspoređivani izolirano izgovoreni logatomi s čitanjem povezanog teksta). Sve su spomenute razlike u frekvencijama formanata statistički značajne, a manifestiraju se i u različitom stupnju identifikacije vokala.

REFERENCije

Bakran, J. (1984)

Model vremenske organizacije hrvatskoga standardnog govora. Disertacija, Sveučilište u Zagrebu, (neobjavljeno)

Van Bergem, D.R., F.J. Koopmans-van Beinum (1989)

Vowel reduction in natural speech, Proceedings of European Conference on Speech Communication and Technology, Eurospeech '89 editori: JP Tubach i JJ Mariani vol. 2, 285-288.

Broad D.J., Clermont F. (1987)

A methodology for modeling vowel formant contours in CVC context. J. Acoust. Soc. Am. 81, 155-165.

Dellatre, P.C., Liberman, A.M., Cooper, F.S. Gertsman, L.J. (1952)

An experimental study of the acoustic determinants of vowel colour, Word 8,

- 195-210. U: Acoustic phonetics, editor: D.B. Fry, Cambridge UP 1976.
- Engstrand, O. (1988)**
Articulatory correlates of stress and speaking rate in Swedish VCV utterances, *J. Acoust. Soc. Am.* 85, 1863-1875.
- Gay, T., Ushijima, T., Hirose, H., Cooper, F.S. (1974)**
Effect of speaking rate on labial consonant- vowel articulation, *J. of Phonetics* 2, 47-63.
- Gay T. (1978)**
Effect of speaking rate on vowel formant movements. *J. Acoust. Soc. Am.* 63, 223-230.
- Gay, T., H. Hirose (1973)**
Effect of speaking rate on labial consonant production: A combined electromyographic high-speed motion picture study, *Phonetica* 27, 203-213.
- Lieberman, P., Blumstein, S. (1988)**
Speech Physiology, Speech Perception and Acoustic Phonetics, Cambridge U.P. New York (prema Nearey, 1989)
- Lindblom, B. (1963)**
Spectrographic study of vowel reduction, *J. Acoust. Soc. Am.* 35, 1773-1781.
- Miller J.D. (1989)**
Auditory - perceptual interpretation of the vowel. *J. Acoust. Soc. Am.* 85, 2114-2134.
- Nearey T.M. (1989)**
Static, dynamic and relational properties in vowel perception. *J. Acoust. Soc. Am.* 85, 2088-2114.
- Peterson, G.E., H.L. Barney (1952)**
Control methods used in a study of the vowels. *J. Acoust. Soc. Am.* 24, 175-184.
- U: Acoustic phonetics, editor: D.B. Fry, Cambridge UP 1976.
- Van Son, R.J.J.H., L.C.W. Pols (1989)**
Comparing formant movements in fast and normal rate speech, *Proceedings of European Conference on Speech Communication and Technology, Eurospeech '89* editori: J.P. Tubach i J.J. Mariani vol 2, 665-668.
- Stamenković, M., J. Bakran, M. Miletić, P. Tancig (1990).**
AGOS - programski sustav za analizu govora. Informatička tehnologija u primijenjenoj lingvistici, Zagreb, Filozofski fakultet, (u tisku)
- Stevens, K.N., A.S. House (1963)**
Perturbation of vowel articulations by consonantal context: An acoustical study, *J. Speech Hearing Res.* 6, 111-128 (prema Stevens et al. 1966)
- Stevens, K. N., A.S. House, A.P. Paul (1966)**
Acoustical description of syllable nuclei: An interpretation in terms of a dynamic model of articulation. *J. Acoust. Soc. Am.* 40, 123-131.
- Strange W. (1989)**
Evolving theories of vowel perception. *J. Acoust. Soc. Am.* 85, 2081-2088.
- Strange W. (1989a)**
Dynamic specification of coarticulated vowels spoken in sentence context. *J. Acoust. Soc. Am.* 85, 2135-2154.

Juraj Bakran
Faculty of Philosophy, Zagreb

*THE EFFECT OF STRESS AND DURATION ON THE FRE-
QUENCY OF VOWEL FORMANTS*

SUMMARY

The traditional concept of the acoustic vowel space represented in terms of F1xF2 is made intricate and embroiled by the existence of several sources of variability including: differences between speakers, speech style, immediate phonetic context etc. The less investigated sources of variability - articulation tempo and stress have divided the investigators and provoked contradictory assumptions. The present investigation, based on the specific corpus of the Croatian standard speech, is trying to establish if the duration or stress of vowels are „per se” essential in determining the degree of organization i. e. reduction of the vowel timbre. Apart from certain limitations the results confirm Lindblom’s recently so often disputed idea according to which reduced duration results in a higher degree of vowel reduction.