

## PROZODIJSKE KARAKTERISTIKE EMOTIVNOG GOVORA: ANALIZA DINAMIČKOG PONAŠANJA GRUPE AKUSTIČKIH OBELEŽJA

Mirjana Rajković, *Institut bezbednosti, Beograd,*  
Slobodan Jovičić, *Elektrotehnički fakultet u Beogradu,*  
Miodrag Đorđević, *mijadjo@drenik.net*  
Zorka Kašić, *Defektološki fakultet u Beogradu*

**Sadržaj** – Ovaj rad prikazuje rezultate istraživanja vezanih za dinamičko ponašanje izabranog skupa akustičkih obeležja u emotivnom govoru. Merenja su vršena na delu korpusa govorne ekspresije emocija i stavova u srpskom jeziku (GEES), sa neutralnim govorom i četiri emotivna stanja: ljutnja, radost, strah i tuga. Prethodni rezultati bazirani na statističkim merenjima na nivou rečenice, na tzv. statičkim obeležjima, pokazuju dobru separaciju između emocija, osim kod radosti i ljutnje. Ovo istraživanje, odnosi se pre svega na analizu varijacija kontura osnovne učestanosti i intenziteta govora, kao i varijacije spektra u određenim podopsezima, na nivou sloga. Naglašeni slogovi su izdvojeni za potrebe analize i unutar njih su vršena merenja skupa dinamičkih obeležja. Različite kombinacije ovih obeležja načinjene su kako bi se analizirale mogućnosti multidimenzionalne separacije emocija. Rezultati su pokazali superiornost dinamičkih obeležja u poređenju sa statičkim u diskriminaciji veoma ekspresivnih emocija, što se može videti kod separacije ljutnje i radosti.

### 1. UVOD

Govorna komunikacija je izuzetno kompleksan oblik komuniciranja među ljudima. U svakoj govornoj poruci, pored lingvističkih obeležja koja direktno prenose informacioni deo poruke, utisnuta su i brojna nelingvistička obeležja.

Emocije su najsloženiji element govora i svakako najznačajnije nelingvističko obeležje. Na osnovu emocija mogu se identifikovati razni aspekti čovekovog fizičkog i psihičkog stanja uključujući godine starosti, pol, inteligenciju, raspoloženje, itd. [3]. Zbog toga analiza ekspresije emocija postaje sve značajnija u sferi moderne govorne tehnologije, kao što je komunikacija čovek-računar, prepoznavanje govora i govornika, kao i sinteza govora.

Osnovni nosioci ekspresije emocija u govoru su varijacije govornih obeležja na tri nivoa: *suprasegmentnom*, odnosno prozodijskom (specifične varijacije frekvencije, intenziteta i trajanja), *segmentnom* (promene kvaliteta artikulacije) i *intrasegmentnom* (opšti kvalitet glasa, odnosno govora) [4]. Jasno je da postoji veliki skup akustičkih obeležja u govoru koji doprinosi ekspresiji emocija, ali sa različitim težinskim koeficijentima značajnosti.

U dosadašnjim istraživanjima, analizirana su sledeća akustička obeležja sa ciljem da izvrše diskriminaciju pet primarnih emocija: osnovna učestanost [5], energija signala [6], trajanje govora/pauza [6], dugo-vremneni spektri [7], formantna struktura vokala /a/ [8] i ostalih vokala [9], intonaciona kriva [11] i glotalna pobuda glasa /a/ [10]. Sve ove analize dale su samo parcijalni rezultat. Naime, nije postignuta potpuna separacija pet osnovnih emocija, već samo njihovo grupisanje. Pokazano je da su ljutnja i radost emocije sa veoma sličnim akustičkim obeležjima, dok s druge strane, strah i tuga predstavljaju drugu grupu sličnih

emocija. Ovakvi rezultati su dobijeni analizom akustičkih obeležja na dužem govornom segmentu - na celoj rečenici, odnosno statističkim usrednjavanjem njihovih vrednosti na nivou rečenica – *statička obeležja*. Na taj način se zapravo gube značajne informacije o pojedinim emocijama na kraćim govornim segmentima.

S druge strane, značajno drugačije diskriminacione karakteristike emocija dobijene su na osnovu doživljaja slušalaca [12]. Posebno je značajno što se na ovaj način mogu potpuno razdvojiti ljutnja i radost, što je bilo nemoguće postići *statičkim* merama.

U ovom radu, polazi se od pretpostavke da su nosioci informacija o emotivnom stanju govornika naglašeni slogovi u rečenici [11]. Izabrane rečenice su segmentirane na slogove i na njima je izvršeno merenje određenog skupa obeležja. Različitim kombinacijama izmerenih obeležja ispitana je mogućnost bolje separacije osnovnih emocija i neutralnog govora. Poseban naglasak je stavljen na mogućnost razdvajanja ljutnje od radosti.

### 2. ORGANIZACIJA EKSPERIMENTA

Govorni materijal, koji je u radu korišćen, je deo korpusa govorne ekspresije emocija i stavova u srpskom jeziku, GEES [1].

#### 2.1 Govorna baza

Govorna baza GEES [1] formirana je sa ciljem određivanja minimalnog skupa akustičkih parametara pomoću kojih se može izvršiti karakterizacija emocija i stavova u ekspresivnom govoru. Korpus sadrži 3510 ekspresivnih iskaza. Pripremljeni govorni materijal izgovaralo je šest glumaca (tri žene i tri muškarca), koji su studenti završnih godina Faulteta dramskih umetnosti (FDU). Snimanje govorne baze je izvršeno profesionalnom digitalnom audio opremom u antisonornoj sobi studija FDU. Prilikom izbora jezičkih elemenata za govorni korpus, vodilo se računa o fonetskim i jezičkim proporcijama.

Za potrebe analize korišćen je jedan manji deo korpusa GEES i to: 3 kratke, semantički neutralne rečenice, u interpretaciji šest govornika. Za svaku emociju smo imali 18 rečenica, odnosno ukupno 90 ekspresivnih govornih iskaza.

#### 2.2 Segmentacija na akcentovane slogove

Prvi korak u obradi izabranih rečenica predstavlja segmentacija na slogove. Segmentacija je izvršena korišćenjem programa *Praat* ručno, uz korišćenje spektrograma, vizuelnog prikaza signala govora, kao i intenzitetske konture, uz neizostavno preslušavanje govornog materijala. Sam posao je obavio ekspert, fonetičar, koji je

osim same segmentacije izvršio i označavanje akcentovanih slogova u rečenicama.

Tabela 1: Izabrani skup obeležja.

Akustička obeležja			
Br	Oznaka	Opis	Dim
1	F0-max	Maksimalna vrednost F0 u akcentovanom slogu.	Hz
2	F0-min	Minimalna vrednost F0 u akcentovanom slogu.	Hz
3	F0-max-poz	Pozicija F0 max u akcentovanom slogu.	s
4	F0-min-poz	Pozicija F0 min u akcentovanom slogu.	s
5	F0-opseg	Opseg F0 u akcentovanom slogu.	Hz
6	F0-nagib	Srednja vrednost izvoda F0 između sukcesivnog F0 max i min u akcentovanom slogu.	Hz/s
7	Tr-sloga	Trajanje akcentovanog sloga	s
8	Tr-F0	Trajanje F0-konture u akcentovanom slogu.	s
9	E1/E	Relativna energija u podopsegu 0-1kHz u odnosu na energiju u celom opsegu.	
10	E2/E	Relativna energija u podopsegu 2.5-3.5 kHz u odnosu na energiju u celom opsegu	
11	E3/E	Relativna energija u podopsegu 4-5 kHz u odnosu na energiju u celom opsegu.	
12	F0-var	Broj promene znaka u varijacijama sukcesivnih odmeraka F0 konture.	
13	E-var	Broj promene znaka u varijacijama sukcesivnih odmeraka E konture.	

### 2.3 Izbor akustičkih parametara

Tabela 1 prikazuje skup od trinaest izabranih obeležja koja su merena na naglašenim slogovima u rečenicama (obeležja 1-11) i na nivou cele rečenice (obeležja 12 i 13). Obeležja 1-8 su merena pomoću programa Praat, dok su obeležja 9-13 računata nakon odgovarajućih merenja.

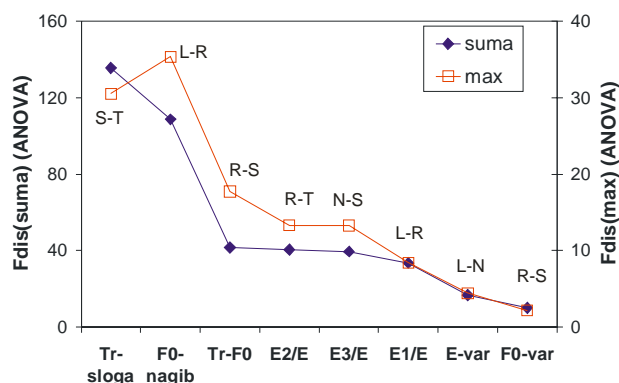
## 3. REZULTATI

### 3.1 ANOVA testovi

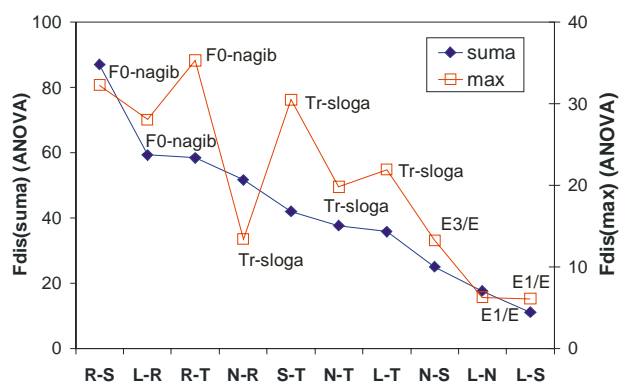
Da bi se odredila značajnost svakog od izabranih obeležja u distinkciji parova emocija, poslužili smo se *One-Way ANOVA* testom. ANOVA test (na bazi faktora F) smo sprovedeli na svim rečenicama i za sve govornike zajedno. Cilj ovog testa je da se izračunaju faktori značajnosti pojedinih obeležja u distinkciji emocija  $F_{dis}$ , pri čemu je nivo značajnosti definisan za verovatnoću od 0,05. Faktori značajnosti su određivani za sve parove emocija (ukupno 10 parova emocija).

Rezultati  $F_{dis}$  su najpre rangirani po značajnosti svakog obeležja u distinkciji pojedinog para emocija, kao i zbirno za sve parove emocija. Na slici 1 je prikazan grafik rangirane ukupne značajnosti  $F_{dis}(suma)$  (ANOVA) obeležja za sve parove emocija. Druga kontura se odnosi na maksimalnu vrednost značajnosti odgovarajućeg obeležja  $F_{dis}(max)$  (ANOVA) (naznačeno je uz tačku na grafiku koji je to par

emocija). Grafik obuhvata 8 obeležja od izmerenih 13 u tabeli 1 (obeležja 6-13), pošto su obeležja koja se odnose na F0



Sl.1. Grafik ukupne značajnosti svakog obeležja (suma), kao i maksimalne pojedinačne značajnosti (max).



Sl.2. Grafik ukupne značajnosti svih 8 obeležja (suma), kao i maksimalne značajnosti (max) odgovarajućeg obeležja po svakom paru emotivnih iskaza.

unutar naglašenog sloga (1-5) sadržana u obeležju F0-nagib (redundantnost informacija).

Na slici 2 je prikazan grafik rangirane ukupne značajnosti  $F_{dis}(suma)$  (ANOVA) svih 8 (prethodno naznačenih) obeležja u distinkciji određenog para emocija. Na istom dijagramu uporedo je prikazan i grafik značajnosti samo jednog obeležja (naveden uz odgovarajuću tačku grafika) čija je vrednost maksimalna u distinkciji odgovarajućeg para iskaza  $F_{dis}(max)$  (ANOVA).

U cilju iznalaženja minimalnog broja obeležja koji obezbeđuju dovoljno dobru separaciju emocija, potrebno je, kao i u slučaju *statičkih mera* [13] izvršiti procenu individualne značajnosti obeležja. U tu svrhu je izvršeno rangiranje svih obeležja za svaki par iskaza, a zatim su izdvojena po tri najbolje rangirana. U tabeli 2 data su ta obeležja i numeričke vrednosti njihovih značajnosti.

Posmatranjem slika 1 i 2, kao i tabele 2, možemo reći da od dinamičkih obeležja, najveću značajnost imaju obeležja F0-nagib i to za parove R-S, L-R i R-T, dok je drugo po značaju obeležje Tr-sloga i to za parove N-R, S-T, N-T i L-T, što odgovara saznanju da obeležje - trajanje govora/pauza izdvaja tugu od ostalih emocija [6]. Što se tiče ostalih obeležja, F0-var i E-var se mogu odbaciti, obzirom na veoma male vrednosti njihovih značajnosti, dok se obeležja iz domena energije rangiraju negde između. Takođe je činjenica

da su vrednosti faktora značajnost kod dinamičkih obeležja do 20 puta manje u odnosu na odgovarajuće vrednosti kod statičkih obeležja.

Tabela 2: Najznačajnija tri obeležja za svaki par iskaza.

Parovi iskaza	OBELEŽJA / Vrednosti Fdis		
L-N	E1/E 6,26	E-var 4,41	F0-nagib 2,83
L-R	F0-nagib 28,07	Tr-sloga 15,73	E1/E 4,05
L-S	E1/E 6,11	Tr-sloga 1,3	F0-var 1,2
L-T	Tr-sloga 21,95	E1/E 6,14	E-var 3,92
N-R	Tr-sloga 13,41	Tr-F0 11,3	E2/E 10,8
N-S	E3/E 13,26	F0-nagib 2,57	E-var 2,12
N-T	Tr-sloga 19,82	E3/E 12,63	Tr-F0 2,31
R-S	F0-nagib 32,31	Tr-sloga 26,04	Tr-F0 17,7
R-T	F0-nagib 35,33	E2/E 13,3	F0-var 2,15
S-T	Tr-sloga 30,51	Tr-F0 4,39	E2/E 3,47

### 3.2 Multidimenzionalno skaliranje (MDS)

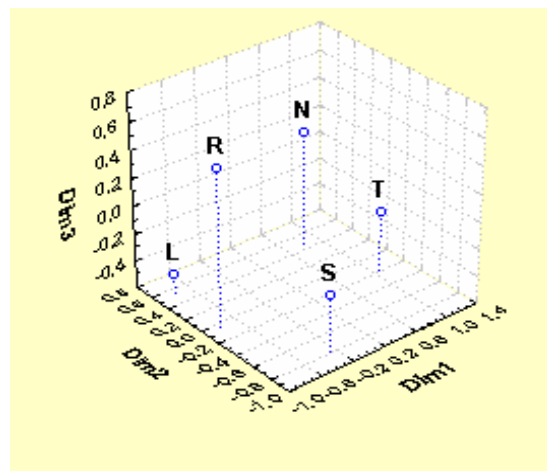
U cilju nalaženja multidimenzionalne raspodele osnovnih emocija na osnovu njihovog perceptivnog doživljaja koristili smo MDS (*multidimensional scaling*) analizu. Ova analiza kreira optimalnu prostornu reprezentaciju emocija na bazi njihove perceptivne sličnosti, a zatim nam omogućava da upoređujemo subjektivne doživljaje sa određenim objektivnim akustičkim obeležjima. Matrica konfuzije (5 stimulusa x 5 odziva) je dobijena u postupku evaluacije govornog korpusa [1] i prikazana je u tabeli 3. Ova matrica se koristila u MDS analizi kao matrica sličnosti za analizirane emocije.

Tabela 3: Zbirna matrica konfuzije u percepciji emocija za sve govornike (u %).

Spikeri	Stimulusi	Odziv- prepoznavanje emocije (%)					Napomena: 100% <=> 4692
		N	L	R	S	T	
Svi	N	<b>93.54</b>	1.56	0.17	0.92	3.82	
	L	0.90	<b>94.52</b>	3.47	1.07	0.04	
	R	1.53	5.07	<b>89.26</b>	4.03	0.11	
	S	1.17	2.28	2.30	<b>93.16</b>	1.09	
	T	6.20	0.04	0.13	0.60	<b>93.03</b>	

Zatim je urađena MDS procedura u programu STATISTICA 5.0. Za startnu konfiguraciju korišćena je Guttman-Lingoes konfiguracija. Posle 26 iteracija i Kruskal stress faktora vrednosti 0,000048, dobijena je finalna konfiguracija u tri dimenzije. Na slici 3 prikazana je

trodimenzionalna raspodela emocija. Emocije su veoma dobro razdvojene u trodimenzionalnom prostoru, što odgovara polaznoj matrici prema kojoj su sve primarne emocije prepoznate sa više od 85%.



Sl.3. Raspodele emocija i neutralnog govora u trodimenzionalnom prostoru dobijena MDS procedurom.

### 3.3 Korelacija MDS i akustičkih karakteristika

Sledeći korak u nalaženju multidimenzionalne raspodele emocija je ispitivanje da li dimenzije perceptivnog prostora mogu biti u relaciji sa 8 akustičkih obeležja analiziranih u tački 3.1.

Tabela 4 predstavlja rezultate Spearman-ovog testa korelacije rangiranih parova koordinata osnovnih emocija u svakoj dimenziji i akustičkih obeležja. "\*" označava da je korelacija značajna sa verovatnoćom  $p < 0,05$ , odnosno "\*\*\*" sa verovatnoćom  $p < 0,01$ .

Tabela 4: Spearman-ov koeficijent korelacije akustičkih obeležja sa dimenzijama perceptivnog prostora.

Dimenzije	Dim1	Dim2	Dim3
<b>Akustičko obeležje</b>			
F0-nagib	-0,1	-0,1	1**
F0-var	0,5	-0,6	-0,7
E-var	0,6	-0,1	0,7
E1/E	0	-0,3	0,9*
E2/E	-0,9*	-0,1	0,2
E3/E	-0,5	-0,6	0,3
Tr-sloga	0,3	0,0	0,6
Tr-F0	-0,3	0,2	0,5

Dimenzija 1 je korelisana sa obeležjem E2/E, što znači da ovu dimenziju predominantno definiše energija signala. Sa slike 3 se jasno vidi potpuna diskriminacija energetski slabih emocija (N i T) i jakih emocija (L i R), dok se emocija strah (S) nalazi između. Dimenzija 2 nije pokazala značajnu korelaciju ni sa jednim od analiziranih obeležja. Ipak, slika 3 pokazuje da dimenzija 2 pravi jasnu diskriminaciju para emocija L-N od para emocija R-T i emocije S. Takođe, tabela 4 pokazuje da je dimenzija 2 u određenoj korelaciji sa

obeležjima F0-var i E3/E (-0,6). Na primeru emocije straha (S) ova indikacija je sasvim korektna jer su kod straha izražene varijacije u F0 konturi kao i uvećanje energije na višim frekvencijama. I konačno, dimenzija 3 pokazuje jaku korelaciju sa dva obeležja F0-nagib i E1/E. Integralno dejstvo ova dva obeležja omogućava bolju distinkcija emocija, što je uočljivo na slici 3. Treba naglasiti da ovu dimenziju u najvećoj meri karakteriše obeležje F0-nagib.

#### 4. ZAKLJUČAK

Obzirom da je ova analiza rađena na malom broju rečenica, zbog obimnog posla na samim slogovima, možemo reći da predstavlja pilot istraživanje po pitanju *dinamičkih obeležja*. ANOVA analiza između parova emocija je pokazala da su za međusobnu diskriminaciju značajna obeležja F0-nagib, Tr-sloga, kao i energetska obeležja. Ono što je posebno bitno je da parametar F0-nagib sadrži onu informaciju zahvaljavujući kojoj se perceptivno razlikuju emocije radost i ljutnja. Prethodno ispitivana – *statička obeležja* nisu omogućavala dobru separaciju upravo ovih emocija.

MDS analiza je pokazala odličnu trodimenzionalnu prostornu raspodelu (separaciju) emocija ali baziranu na njihovoj perceptivnoj sličnosti, odnosno različitosti. Ispitivanje korelacije dimenzija perceptivnog prostora sa izabranim *dinamičkim obeležjima*, potvrdilo je rezultate ANOVA testova za obeležje F0-nagib, i ukazalo da ovo obeležje zaista možemo koristiti u daljoj multi-dimenzionalnoj analizi diskriminacije emocija. Što se tiče obeležja iz domena energije (E2/E i E1/E), i pored velike korelisanosti sa perceptivnim dimenzijama, vrednosti faktora značajnosti ovih obeležja (prema ANOVA testovima) su relativno male za sve parove emocija, što im umanjuje komparativne vrednosti sa drugim obeležjima.

Sledeći korak u analizi emocija bi bio da se uradi MDS analiza za sva do sada obrađivana obeležja (*statička obeležja*), gde bi se mogli očekivati bolji rezultati s obzirom na velike vrednosti faktora značajnosti kod pomenutih obeležja.

#### LITERATURA

- [1] S.T. Jovičić, Z. Kašić, M. Đorđević, M. Vojnović, M. Rajković, J. Savković, "Formiranje korpusa govorne ekspresije emocija i stavova u srpskom jeziku- GEES", *TELFOR, Beograd 2003*.
- [2] M. Rajković, M. Đorđević, S. Jovičić, "Karakterizacija emotivnog govora reprezentativnim skupom akustičkih obeležja", *ETRAN-2004, Čačak 2004*.
- [3] Schötz, S., "Linguistic & paralinguistic variation in speaker recognition & text-to-speech synthesis", *GSLT: Speech Technology 1, term paper, 2002*.
- [4] Z. Kašić, "Funkcija suprasegmenta u govornom izrazu", *Beogradska defektološka škola, br. 2-3, str. 113-124, Beograd, 2000*.
- [5] I. Đukić, S. Jovičić, M. Đorđević, "Intonacione karakteristike emotivnih ekspresija u srpskom govornom diskursu", *TELFOR, Beograd 2003*.
- [6] M. Rajković, D. Rakić, S. Jovičić, M. Vojnović, M. Đorđević, "Intenzitetske i vremenske karakteristike emotivnih ekspresija u srpskom govornom diskursu", *TELFOR, Beograd 2003*.

- [7] M. Vojnović, "Spektralne osobine emotivnog govora", *ETRAN-2004, Čačak 2004*.
- [8] M. Vojnović, "Uticaj emotivnog stanja govornika na formantnu strukturu vokala /a/", *DOGS 2004, Sombor 2004*.
- [9] M. Đorđević, S. Jovičić, M. Rajković, Z. Kašić, "Analiza značajnosti akustičkih obeležja u distinkciji primarnih emocija u emotivnom govoru", *DOGS 2004, Sombor 2004*.
- [10] M. Vojnović, "Uticaj emotivnog stanja govornika na glotalnu pobudu kod izgovora vokala /a/", *DOGS 2004, Sombor 2004*.
- [11] M. Vojnović, "Intonation changes provoked by different emotional states of speakers" *Speech and Language 2004, Beograd 2004*.
- [12] M. Rajković, M. Đorđević, S. Jovičić, "Klasifikacija osnovnih emocija pomoću Plutchik-ovog točka", *DOGS 2004, Sombor 2004*.
- [13] M. Đorđević, M. Rajković, "Kvantifikovanje značajnosti akustičkih obeležja u inter i intra spikerskoj distinkciji emotivnog govora", *ETRAN-2004, Čačak 2004*.

**Abstract** - This paper presents the results of investigation of dynamic behaviour of a set of acoustic features in emotional speech. Measurements were performed on a part of Serbian emotional database (GEES), with neutral speech and four emotional states: anger, happiness, fear and sadness. Previous results based on statistical measures at a sentences level, referred as "static measures", showed good separation between emotions, except between anger and happiness. This investigation analyzes the variations of intensity and pitch contour, and subband spectral variations at the syllable level. The stressed syllables were extracted for final analysis and different measures were done inside them. Various combinations of this features were made in order to analyze the possibilities of multidimensional emotion separation. The results show superiority of dynamic features compared to static ones in separation of very expressive emotions, as in the case of anger and happiness.

#### PROSODIC CHARACTERISTICS OF EMOTIONAL SPEECH: MEASUREMENTS OF DYNAMIC BEHAVIOUR OF A SET OF ACOUSTIC FEATURES

M. Rajković, S. Jovičić, M. Đorđević, Z. Kašić