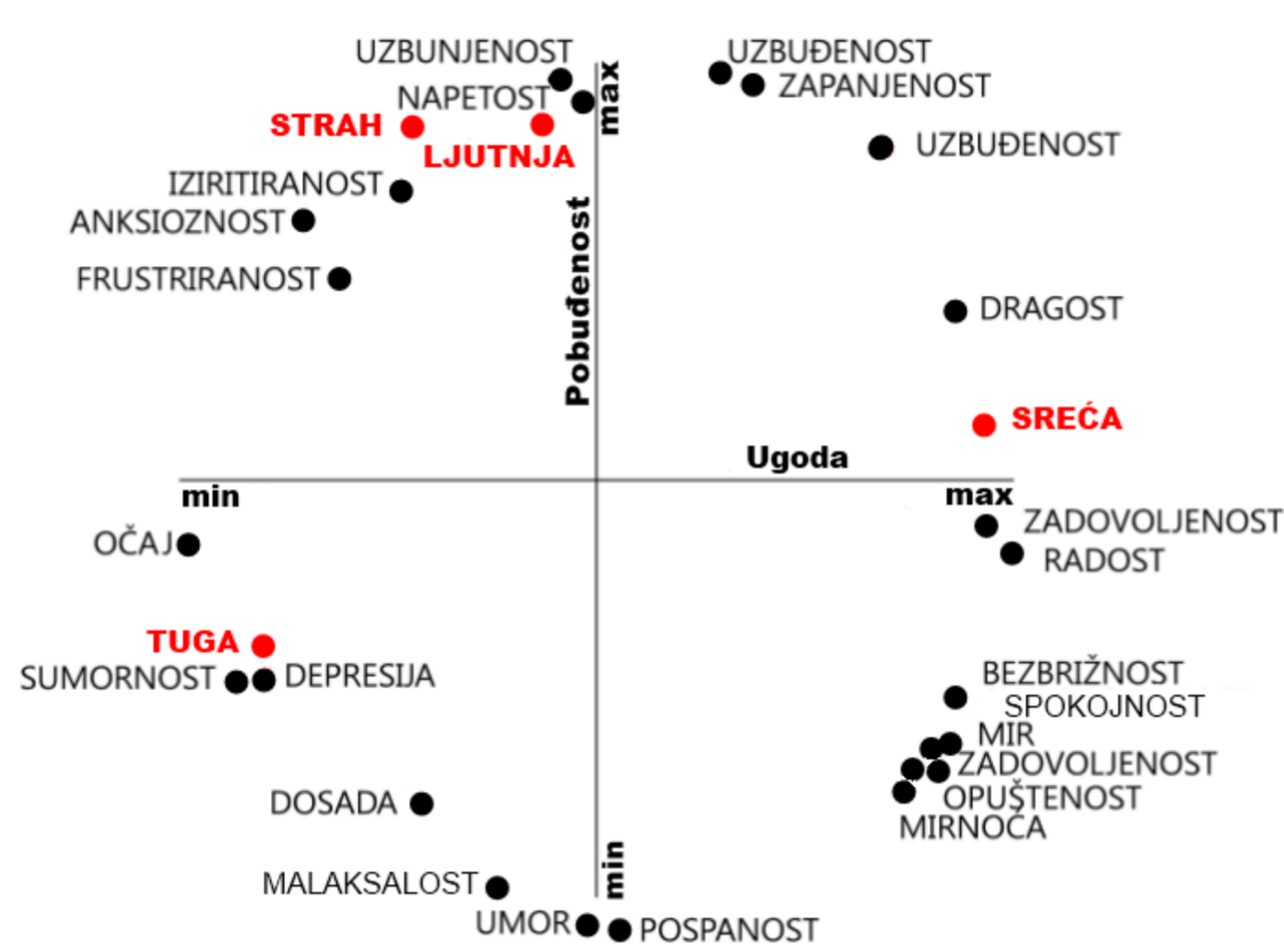


1 Uvod

Emocije su stanja uzbuđenosti organizma koja se izražavaju kroz:

- Specifična ponašanja
- Fiziološke promjene
- Subjektivni doživljaj

Postoji više sustava klasifikacije emocionalnog stanja, a najpoznatiji su klasifikacija na osnovu diskretnih emocionalnih stanja i klasifikacija na osnovu razine ugone i pobuđenosti.



Pored klasifikacija na osnovu lingvističkih značajki paralelno se razvijao i sustav na osnovu akustičkih značajki. (Chmura, M.)

2 Opis problema

Cilj diplomskog rada je bio istražiti postupke klasifikacije diskretnih emocionalnih stanja. Kao osnovu za klasifikaciju trebalo je koristiti lingvističke značajke govora. Sustav se razvijao za hrvatski jezik, a klasifikacija se provodila na razini rečenica. Za lingvističku analizu su se istražili modeli jezika poput n-gram modela i drugih statističkih modela za modeliranje jezika. U svrhu ekstrakcije lingvističkih parametara iz govora bilo je potrebno upotrijebiti već razvijene sustave za automatsko prepoznavanje govora temeljene na skrivenim Markovljevim modelima (HTK).

Primarni problemi

- Emocionalna stanja koja je potrebno klasificirati:
 - Sreća
 - Tuga
 - Strah
 - Ljutnja
 - Neutralno
- Ne postoji korpus emocionalnog govora za hrvatski jezik
- Sakupljanje snimki dugotrajno i teško
- Potrebno izvršiti anotaciju korpusa
- Pronaći odgovarajući lingvistički model
- Primijeniti sustav prepoznavanja govora

3 Konstrukcija KEG-a

KEG (Korpus Emocionalnog Govora)

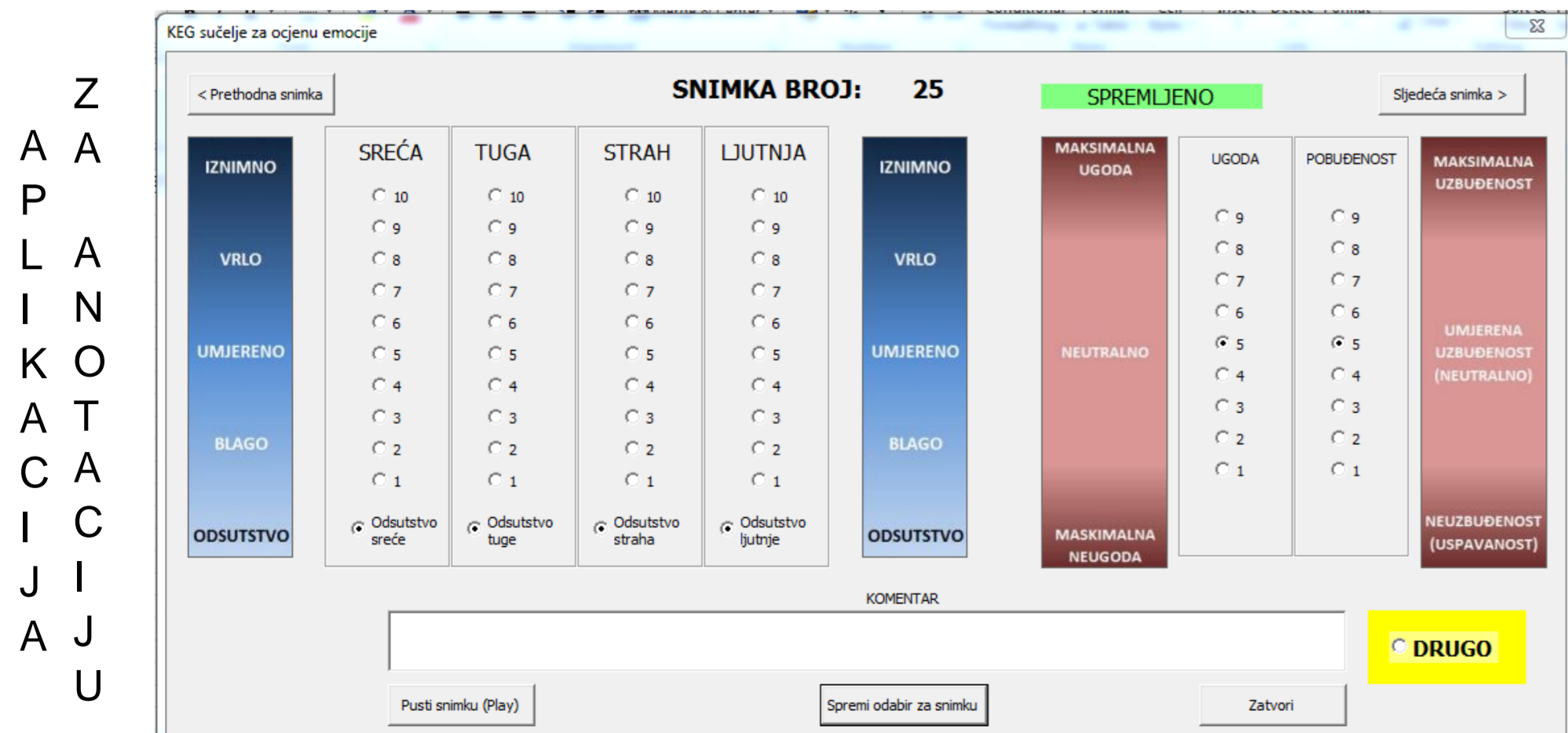
- Faza prikupljanja
- Faza anotacije

SNIMKE

- Stvarne emocije
- Glumljene emocije

KEG statistika nakon anotacije

Emocija	Snimaka (broj.)	Spol po snimci (M/Ž)	Dobna skupina (D/Ado./O dra.)	Vrsta emocija (S/G)	Govornici (broj.)	Spol po govorniku (M/Ž)	Dobna skupina (C/Ado./O dra.)	Ukupno trajanje (mm:ss)	Prosječno trajanje (s)
Sreća	77	25/52	10/1/66	34/43	34	15/19	3/1/30	5:27	4.25
Tuga	83	16/67	1/0/82	44/39	30	7/23	1/0/29	8:57	6.47
Ljutnja	231	133/98	2/2/227	105/126	60	31/29	1/1/58	13:27	3.49
Strah	33	7/26	14/0/19	3/30	16	6/10	3/0/13	2:26	4.42
Neutralno	72	43/29	1/0/71	61/11	50	30/20	1/0/49	10:59	9.15
Ukupno	496	224/272	28/3/465	247/249	146	69/77	7/2/137	41:16	4.99



Za razvoj sustava bilo je potrebno prikupiti materijale za trening, a shodno tome izvršiti i anotaciju istih.

4 Lingvistički modeli jezika i HTK

Pri gradnji lingvističkog modela provedeni su sljedeći koraci:

- Razdvajanje transkripta u klase
- Tokenizacija
- Izračun frekvencije pojavljivanja tokena
- Izračun relativne frekvencije pojavljivanja
- Type-Token ratio
- Konstrukcija skupa jedinstvenih tokena po klasi
- Pregled načina redukcije dimenzionalnosti

- Primitivni unigram sustav
- Bayesov tekstualni klasifikator

$$P(c|d) \propto P(c) \cdot \prod_{1 \leq k \leq n_d} P(t_k|c)$$

$$c_{map} = \arg \max_{c \in C} \left[\log \hat{P}(c) + \sum_{1 \leq k \leq n_d} \log \hat{P}(t_k|c) \right]$$



Prepoznavanje govora se obavilo alatom HTK (engl. *Hidden Markov Model Toolkit*). Svi potrebni podatci su prilagođeni već razvijenom ASR (engl. *Automatic Speech Recognition*) sustavu, te se prošlo kroz tri faze: priprema podataka, trening i test.

5 Rezultati

Matrica konfuzije primitivnog unigram sustava

		ULAZI U SUSTAV					Točno prepoznatih
		SREĆA	TUGA	STRAH	LJUTNJA	NEUTRALNO	
IZLAZ IZ SUSTAVA	SREĆA	35	9	4	9	3	35
	TUGA	5	34	1	18	5	34
	STRAH	3	3	2	5	2	2
	LJUTNJA	9	16	7	125	14	125
	NEUTRALNO	6	7	3	17	36	36
	NIJE PREPOZNATO	18	12	16	55	11	232
\sum ulaznih_emocija		76	81	33	229	71	490
UKUPNO							47.34 %

Matrica konfuzije za Bayesov tekstualni klasifikator

		ULAZI U SUSTAV					Točno prepoznatih
		SREĆA	TUGA	STRAH	LJUTNJA	NEUTRALNO	
IZLAZ IZ SUSTAVA	SREĆA	30	3	0	0	3	30
	TUGA	4	25	0	8	3	25
	STRAH	3	0	2	1	1	2
	LJUTNJA	39	51	30	213	37	213
	NEUTRALNO	2	3	1	6	31	31
	NIJE PREPOZNATO	0	0	0	0	0	301
\sum ulaznih_emocija		75	82	33	228	72	490
UKUPNO							61.42 %

Sustav se trenirao sa 90 % materijala KEG-a, odnosno sa 447 rečenica. Ostatak od 49 rečenica (10 %) upotrijebljen je za testiranje tako utreniranog sustava. Rezultati prepoznavanja alatom HTK na ovom korpusu dosežu svega 13 %, što je izravna posljedica iznimno velikog broja govornika i loše kvalitete snimki u kategoriji stvarnih emocija.

6 Zaključak

Ovim diplomskim radom po prvi puta na Zavodu za elektroničke sustave i obradu informacija se proučila mogućnost automatske klasifikacije diskretnih emocionalnih stanja na osnovu lingvističkih parametara govora. U svrhu tome razvio se i prvi korpus hrvatskog emocionalnog govora, koji može poslužiti kao osnova za daljnji razvoj i istraživanje. Iz rezultata je vidljivo kako je čista lingvistička analiza moguća uz relativno zadovoljavajuću točnost. Isto ne možemo reći i za sustav prepoznavanja govora, što je posljedica samih karakteristika KEG-a i trenutačne nesavršenosti funkcioniranja sustava prepoznavanja govora.