



JEZIK I NJEGOVI UČINCI

Zbornik radova s međunarodnoga znanstvenog skupa
Hrvatskoga društva za primijenjenu lingvistiku
održanoga od 4. do 6. svibnja 2017. godine u Rijeci

Uredile

Diana Stolac

Anastazija Vlastelić

Za izdavača

Damir Agičić

Knjigu uredile

Diana Stolac

Anastazija Vlastelić

Recenzenti članaka

Mislava Bertoša, Goranka Blagus Bartolec, Kristina Cergol Kovačević, Vesna Deželjin, Renata Geld, Maja Glušac, Dejana Golenko, Tanja Gradečak-Erdeljić, Jim Hlavač, Aleksandra Horvat, Lana Hudeček, Snježana Husinec, Nada Ivanetić, Damir Kalogjera, Arijana Krišković, Ivana Kurtović Budja, Jelena Kuvač Kraljević, Radovan Lučić, Marija Lütze Miculinić, Ivana Matas Ivanković, Mihaela Matešić, Daniela Matić, Ana Meštrović, Milica Mihaljević, Vesna Mildner, Borana Morić-Mohorovičić, Jadranka Nemeth Jajić, Kristian Novak, Jasna Novak Milić, Bernardina Petrović, Neda Pintarić, Elenmari Pletikos Olof, Boris Pritchard, Anita Runjić-Stoilova, Renata Šamo, Kristina Štrkalj Despot, Branka Tafra, Sanda Lucija Udier, Lovorka Zergollern Miletić

Jezična je redakcija autorska.

Recenzentice zbornika

prof. dr. sc. Vlasta Rišner

doc. dr. sc. Sanda Lucia Udier

Grafički urednik

Krešimir Krnic

Grafička priprema

Tvrtko Molnar

Banian ITC

Tisak

Tiskara Zelina

ISBN: 978-953-7963-86-6

CIP zapis je dostupan u računalnome katalogu

Nacionalne i sveučilišne knjižnice u Zagrebu

pod brojem 001005206.

Objavljivanje knjige financijski je pomoglo

Ministarstvo znanosti i obrazovanja RH.

Knjiga je objavljena u rujnu 2018.

BENEDIKT PERAK

Sveučilište u Rijeci, Filozofski fakultet

bperak@ffri.hr

KATARINA DAMČEVIĆ

Institut za filozofiju i semiotiku u Tartu

katarinadamcevic@gmail.com

JANA MILOŠEVIĆ

Sveučilište u Rijeci, Filozofski fakultet

janamilosevic0@gmail.com

O sranju i drugim (ne)primjerenim stvarima: leksičko-korpusna analiza i graf konstrukcija semantičkih domena tabuiziranih, vulgarnih i derogatornih riječi u hrvatskome

Leksikografskom i korpusnom metodologijom pristupa se identifikaciji semantičkih domena tabuiziranih, vulgarnih i derogatornih riječi, te ih se klasificira temeljem sintaktičko-semantičke umreženosti koordiniranom konstrukcijom [*x* i/ili *y*]. Neprimjerenost se ontološki tumači kao specifični dio kulturalnog sustava komunikacijskih strategija društvene interakcije između govornika i sugovornika s ciljem spoznajne aktivacije kognitivno-afektivnih procesa. Unutar dinamičnog psihološko-sociološko-kulturalnog sustava određeni iskazi opojmljivanja socijalne interakcije sagledavaju se kao neprimjereni, te postaju tabuizirani. Svojestvo neprimjerenosti ovisi o kulturalnom kontekstu i komunikacijskoj situaciji. Ono se ogleda upravo u sustavu tabua određene kulture, odnosno onoga što se smatra neprimjerenim u aktualnim uspostavljenim normama i vrijednostima koje sačinjavaju svojevrsni sustav morala određene kulture. Kognitivno-lingvistička korpusna analiza psovki u hrvatskome jeziku provodi se korištenjem hrWac korpusa. Na temelju graf analize pojmova umreženih koordiniranom konstrukcijom istaknuta su semantička i ontološka obilježja neprimjerenih imenskih leksema kao i povezanost s drugim leksemima tvoreći tako značenjski sustav od 1065 umreženih pojmova. Rad pridonosi razvoju postavki kvalitativne analize psovki temeljem selektiranih lema u koordiniranoj konstrukciji, te ukazuje na potrebu razvoja sveobuhvatnije metodologije inter-/kroz-kulturnog istraživanja neprimjerenog jezika.

Ključne riječi: psovke; neprimjereni jezik; tabu; koordinirana konstrukcija; graf teorija

1. Uvod

Zašto su određene riječi neprimjerene da bi se izgovarale u određenim prilikama ili određenoj kulturi? Kako su postale neprimjerene ili čak zabranjene? Što je to neprimjereno u njima? Koje su to riječi ili fraze u hrvatskome? Ovo su neka od pitanja koja se vežu uz proučavanje psovki i neprimjerenog jezika. Cilj je ovoga rada postaviti teorijsku osnovu za sustavno razmatranje navedenih pitanja i ponuditi empirijski pristup identificiranju i klasificiranju neprimjerenih riječi i fraza u hrvatskom jeziku. Počevši od kognitivno lingvističke pretpostavke da se jezičnim izrazima uspostavljaju obrasci interakcije kojima se ostvaruje intersubjektivna komunikacija značenja, riječi se smatraju konvencionaliziranim simboličkim kodom za aktiviranje mentalne simulacije zamjedbe stvarnosti, afektivnih stanja i kognitivnih procjena. Riječi koje su označene kao vulgarizmi, psovke, pejorativi ili derogativi, odnosno uvredljive riječi ili sintagme, karakterizira određena razina iskazivanja neprimjerenog sadržaja sukladno normama i konvencijama određene kulture slijedom čega njihova uporaba postaje ograničena odnosno tabuizirana¹. U rječnicima se takve riječi označavaju posebnim oznakama poput tab. za tabuizirane, derog. za derogatorne riječi, pejor. za pejorative odnosno pogrdnice, vulg. za vulgarizme (vidi: Bujas 2001, Jojić 2015). Koristeći se leksičkim natuknicama kao početnim unosom za ekstrakciju podataka o uporabi takvih imenica u korpusu te semantičkim značajkama srodnosti koju obilježava koordinirana konstrukcija [leksem *x* i leksem *y*] u ovome su radu prikazane semantičke domene neprimjerenih riječi u hrvatskome jeziku.

Temeljem klasifikacije umreženih pojmova predstaviti će se semantičke domene neprimjerenih riječi što pridonosi razvoju postavki za kvantitativnu i kvalitativnu analizu psovki, kao i daljnjem razvitku kros-kulturalne i interkulturalne metodologije istraživanja neprimjerenog jezika.

Struktura rada je sljedeća. Drugo poglavlje daje pregled relevantnih istraživanja neprimjerenog jezika, kako hrvatskih tako i stranih autora. Također se ukazuje na položaj ovog rada u odnosu na navedeni doprinos temi. Treće poglavlje pruža uvid u važnost leksikografskog pristupa za pronalaženje i prikupljanje neprimjerenih riječi, ali objašnjava i nužnost uvođenja drugih izvora informacija koji omogućavaju detaljniji uvid u dinamičan fenomen neprimjerenog jezika. U četvrtome

1 Prema Allan i Burridge, „svaka vrsta ponašanja može biti tabuizirana. Da bi ponašanje bilo propisano mora biti percipirano kao štetno za individualca ili za njegovu/njezinu zajednicu; razina štete može biti kršenje bontona, kao i neka kobnija posljedica“ (Allan, Burridge 2006: 11). Valja istaknuti da spomenuto ovisi o sociokulturalnom kontekstu kulture, stoga ne možemo tvrditi da postoji univerzalan tabu: ono što je tabuizirano u jednom periodu ne mora tako ostati i u drugom. Vrijedan doprinos problematizaciji tabu praksi i njihovih društvenih funkcija moguće je pronaći u djelu Mary Douglas naslova *Purity and Danger: An Analysis of Concepts of Pollution and Taboo* (2002 [1966]). U navedenom ukazuje na rituale čistoće i prljavštine te ističe njihovu isprepletenost, ali i funkciju analogije izražavanja postojećeg društvenog uređenja (Douglas, 2002: 3).

poglavlju objašnjava se korpusni pristup te ističe njegova primjenjivost za aktualan rad. Svojstvo čestotnosti koje se prikuplja korpusnim pristupom ukazuje na razinu konvencionaliziranosti određene jezične jedinice u komunikacijskom sustavu, što pruža mnogo širi uvid u dinamiku neprimjerenog jezika. Peto poglavlje objašnjava konstrukcijski pristup i usredotočuje se na analizu neprimjerenih imenskih leksema u koordiniranoj jezičnoj konstrukciji te raspravlja o rezultatima klasifikacije pomoću graf algoritma za pronalaženje zajednica.

2. Teorijska pozadina i pregled literature

Razmatranje neprimjerenog jezika u hrvatskoj jezikoslovnoj tradiciji razmjerno je nova pojava. Jedno od prvih djela koja problematiziraju psovke na hrvatskome jeziku objavljeno je 1962. godine pod naslovom *Bludna psovka, povijesno-psihološka studija*. Autor studije, fra Ignacije Gavran, promatra psovke kao izraze koji direktno demoraliziraju recipijenta. Nadalje, Mikić i suradnici 1999. godine objavljuju kontrastivno istraživanje *Psovka u hrvatskome i njemačkome jeziku* određujući psovanje kao „kapitalna ubojita uvreda kojom se čovjeku namjerno oduzima njegovo dostojanstvo“ (Mikić et al. 1999: 88). Kao potvrdu svog stava navode da „etimološki gledano riječ *psovati* znači govoriti pogrdne riječi, nekoga grditi, ružiti, kleti, psovkom (ili psovka) obasipati, huliti [...] s prvotnim značenjem činiti, postupati kao pas, dakle čovjeka nedostojno, sramotno“ (Mikić et al. 1999: 89). U skladu sa širim shvaćanjem pragmatičnih obilježja psovki i novijim teorijskim postavkama epistemologije komunikacije, Badurina i Pranjković ističu komunikacijsku funkciju psovke kao „specifična oblika govornoga čina ekspresiva kojim se iskazuju određena govornikova stanja i stavovi te postižu određene komunikacijske namjere“ (Badurina, Pranjković 2016: 233). Takav se teorijski pristup u kojemu se uporaba neprimjerenih riječi smatra sastavnim dijelom jezične prakse s naglašenim funkcijama izražavanja govornikovih emocija, stavova i statusa može naći u nizu recentnijih radova koji proučavaju uporabu neprimjerenog jezika u hrvatskome (Užarević 2012; Memišević i Matešić 2014; Pilch 2011; Lučić 2015; Badurina i Pranjković 2016). Psovanje, kao forma lingvističke aktivnosti, čin je koji upotrebljava tabuizirane riječi s ciljem izražavanja snažnih emocija (Vingerhouts et al. 2013: 287), i to ne nužno samo hedonički negativnih (Jay 2000).

Ljung (2011) u svome istraživanju *Swearing. A Cross-Cultural Linguistic Study* navodi da se u psovanje uključuje: 1) uporaba tabuiziranih riječi, 2) opojmljivanje značenja putem metonimijskih i metaforičkih procesa, 3) formulaična leksička i sintaktička ograničenja, 4) izražavanje osjećaja i stavova govornika (Ljung, 2011: 4). Ljung razlučuje funkcije i teme psovanja, pri čemu funkcije opisuju razloge uporabe psovke, a teme se odnose na opis tabuiziranih semantičko-ontoloških područja. Najzastupljenije su tabu teme, prema Ljungu (2011: 35): religija/nadnaravno, tjelesne izlučevine/skatologija, seksualni organi, seksualne aktivnosti i

majka/obitelj. Manje istaknute teme odnose se na pretke, životinje, smrt, bolest i prostituciju (Ljung 2011: 42-43). Istaknutost tema varira ovisno o društvenom sustavu vrijednosti i prihvaćenom kulturnom modelu. Premda se pretpostavlja da postoje određene kulturološke razlike, teme određene Ljungovom kategorizacijom mogu se zamijetiti i u hrvatskome. Primjerice, Mikić i suradnici (1999) nabrajaju teme životinja, biljaka i drveća, nacije, podrijetla i rase, područje nežive materije, uporabnih predmeta, religije i svjetonazora, ljudskih osobina, odjevnih predmeta i materijala, dijelova tijela, hrane i mirisa, područje bolesti, područje mitskih, religioznih i povijesnih bića, zanimanja/zvanja te napokon, imena i nadimaka.

U hrvatskoj leksikografskoj tradiciji objavljeno je niz rječnika koji navode neprimjerene riječi i klasificiraju ih kao vulgarizme (vulg), tabu (tab), pogrdnice (pogr) ili derogative (derog) (Anić i Pranjković 1998, Bujas 2001, Jojić 2015, Hrvatski jezični portal). Ovaj se rad ograničava na identifikaciju neprimjerenih leksema i semantičkih domena koje se mogu uspostaviti analizom imenica, ne analizirajući glagolske i pridjevske lekseme, niti sintaktičke formulaičnosti koje su svojstvene neprimjerenim izrazima. Osim identifikacije leksema korištenjem leksičkih resursa, određuje se konvencionalizirani ontološki logični asocijativni odnos između leksičkih pojmova koji tvore semantičke domene, odnosno teme neprimjerenog jezika pomoću korpusne metodologije i konstrukcijske analize koordinirane konstrukcije (Perak 2017).

3. Leksikografski pristup identifikaciji neprimjerenih riječi

Leksičke jedinice koje se smatraju neprimjerenim inicijalno su prikupljene korištenjem hrvatsko-engleskog rječnika (Bujas 2001) i Velikog rječnika hrvatskoga standardnog jezika (VRH) (Jonjić 2015). U klasu neprimjerenih riječi Bujas (2001) i Jonjić (2015) bilježe imenice koje su izdvojene u tablici 1.

Tablica 1. Tabuizirane, vulgarne i derogativne prema rječničkoj klasifikaciji Bujas (2001).

Tabu (30)	brzojebac, buzić, buzurant, drkađija, drkanje, fafanje, fukanje, govнар, govнарija, govno, jebač, jebačica, jebačina, jebanje, jebarnik, jebežljivac, jebežljivost, jebica, jebivjetar, kurac, liz, minda, pica, pička, pizdarija, redaljka, ševa, ševac, ševica, šupak
Vulgarne (48)	bljuvotina, buhoserina, drek, dupe, dupence, dupetara, fanfulja, fr-njoklica, guz, guzara, guzatost, guzica, guzonja, hračak, hrakotina, isprdak, kenjaža, kurba, kurbar, kurva, kurvanje, kurvar, kurviš, luk, muhoserina, perioda, pičkaranje, pimpek, pizda, pizdek, pizdun, podjebavanje, popišanac, prdac, prdalo, prdež, prdizvek, prdonja, prkno, radodajka, seronja, sračka, sranje, sroljo, zajebancija, zajebant, zajebavanje

Derogativi (179)	aparatčik, Azijat, babetina, balavac, balavander, balkanac, balkanka, balkanština, balonja, bezmogaš, bijednik, blondać, bukvan, buncalo, buržujština, čelonja, cendralo, cendravic, cerekalo, četnikuša, Čifut, ciganin, ciganjenje, ciganka, ciganluk, cincar, cincarenje, cincarka, cincarluk, cmakanje, cmomantijaš, črčalo, crnja, cvrkutalo, debil, digić, dođoš, došaptavalo, dotepenac, drljotina, drndalo, droljetina, fačkanje, fakultetlija, fifica, fiškal, fiškalstvo, frazer, frazerstvo, gegalo, glavešina, glavonja, glondža, glumatalo, glumatanje, goljo, govorancija, grbonja, grmalj, groktalo, ipara, izležavalo, izmišljalo, izmotavalo, jugoslavenčina, Jugovina, kalkulantstvo, kaputaš, kenjkalo, kenjkanje, kenjkavac, kenjkavost, kimavac, kiseliš, klempavost, klempo, klipan, kljakavac, kljuse, kljusina, kmezavac, kmezavica, koferaš, kompomiser, komunjara, konferencijaš, konferencijaštvo, kožurina, kravetina, krčmetina, kreatura, krele, krepavalo, kretenčina, krezo, krtinja, krpetina, kuferaš, laloka, manipulator, mantijaš, mitingaš, mizerija, mrcina, mućkaroš, muktaš, naptalinac, nakot, narfrakanost, nikogović, njuškalo, nosonja, novinarčić, novotar, novotarstvo, oficirčić, oficirčina, pandur, pandurstvo, parolaš, parolaštvo, partizančina, partizanstina, pašće, peder, pederstvo, pijančina, pijandura, pipalo, pisarija, piskaralo, podrepaš, poluinteligent, popišulja, posranac, potrkalo, prčvarnica, prđavac, prikripa, pripuz, prištavac, prizemnost, prnjavor, profuknjača, purgerstvo, raga, rakijetina, reformaš, režimlija, rječetina, selendra, seljačina, serator, skot, sladunjavost, šminkeraj, soj, srbenda, starkelja, strančar, strančarenje, strvinarenje, titulaš, trbonja, trčkaralo, ućo, uličarka, vodorina, vrdalo, žaca, žbir, ženska, ženturača, zrikavac, zriko, zuba, zubo, zviždalo, zvrndalo
---------------------	--

Tabuizirani leksemi većinom se odnose na seksualne organe i aktivnosti, a vulgarni leksemi na izlučevine, skatološke i seksualne teme. Derogativi, odnosno pogrdnice nešto su tematski širi i posjeduju neke morfosintaktičke obrasce tvorbe poput uvećanica s dočetakom na *-erda*, *-ander* (*balavander*), intenzifikatora *-tina*, *-čina* (*pijančina*, *seljačina*), i čestog metonimijskog profiliranja osobnosti izvedene dočetakom *-aš* (*kuferaš*), *-vac* (*prištavac*), *-alo* (*pipalo*, *vrdalo*), *-lija* (*režimlija*), *-čik* (*aparatčik*).

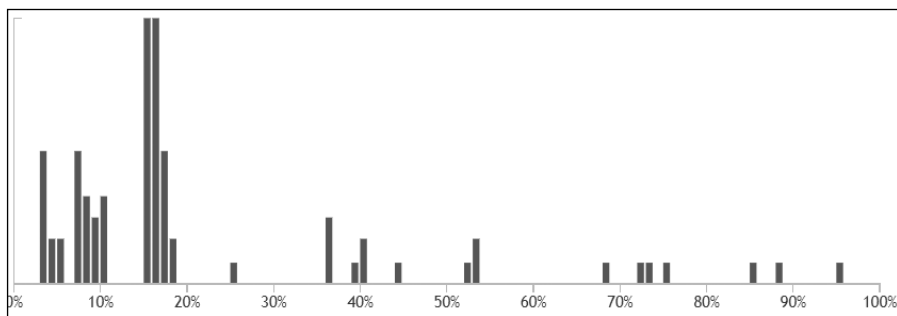
Razlike između rječnika postoje na razini klasifikacije riječi i na razini rječničkih natuknica. Od tabuizama iz Bujasova rječnika nedostaju u Velikom rječniku hrvatskoga (Jonjić 2015) leksemi: *brzojebac*, *buzić*, *buzurant*, *fasanje*, *jebežljivac*, *mindā*, *redaljka*. Nisu zabilježeni vulgarizmi: *fanfulja*, *frnjoklica*, *kenjaža* i *muho-serina*. Nadalje, riječi poput *glupan*, *tupan*, *bedak* nisu klasificirane niti pod jednu vrstu neprimjerenih riječi u oba rječnika. S druge strane moguće je zapitati zašto bi se *liz* klasificirao kao tabu, osim ako nije korišten u tabuiziranom seksualnom kontekstu. Za potrebe strukturiranog popisivanja neprimjerenih riječi u hrvatskome, navedeni rječnici koristan su, ali nepotpun izvor informacija. Štoviše, znajući da neprimjereni jezik predstavlja vrlo dinamičan i kreativan fenomen koji se javlja

iz razgovornog jezika, popis je potrebno proširiti izvorima koji predstavljaju ono što se u suvremenoj terminologiji naziva ‘građanska znanost’ (engl. *citizen science*)². Primjer takvih strukturiranih i polustrukturiranih mrežnih portala koji prikupljaju podatke o neprimjerenim izrazima su Žargonaut (vidi: <http://www.zargonaut.com/tag/vulgarno>, <http://www.zargonaut.com/tag/pogrdno>), Forum.hr tema: Svi nazivi za vaginu na hrvatskom jeziku³, Crol forum tema: Nazivi za njezinu i njegovu stvar...⁴ i brojni drugi koji zaslužuju da se uvrste u popis za ekstrakciju podataka o uporabi neprimjerenog jezika ne samo zbog preskriptivnih razloga, već i kao svjedočanstvo o stvaralaštvu govornika i raznovrsnosti pojmovnog mapiranja.

4. Korpusni pristup

Prednost korpusnog pristupa očituje se u podacima o uporabi jezičnih struktura u komunikaciji. Podaci mogu biti relevantniji ili manje relevantni s obzirom na veličinu i sastav korpusa. Očito je da se u CHILDES korpusu⁵ dječjeg govora s 126 sudionika i 380 tisuća pojava neće pronaći niti jedna od najčešće tabuiziranih riječi. Također, korpusi koji sadrže visoki udio književnih djela izdanih prije 1990. godine i (samo)cenzuriranih novinskih članaka do 2000-ih, poput korpusa Riznice⁶ neće odraziti stvarne vrijednosti čestotnosti neprimjerenih leksema u različitim suvremenim komunikacijskim stilovima kao što to mogu veći korpusi mrežnih domena poput hrvac. Oni obuhvaćaju i slobodnije forme poput mrežnih stranica, blogova, foruma, komentara portala i slično. Za primjer navodimo distribuciju leksema *sranje* u oba korpusa (prikaz 1. i 2.)

Prikaz 1. Distribucija čestotnosti leksema *sranje* u korpusu Riznica (77 pojava, 0.76 milijun)⁷



2 https://en.wikipedia.org/wiki/Citizen_science

3 <http://www.forum.hr/showthread.php?t=631942>

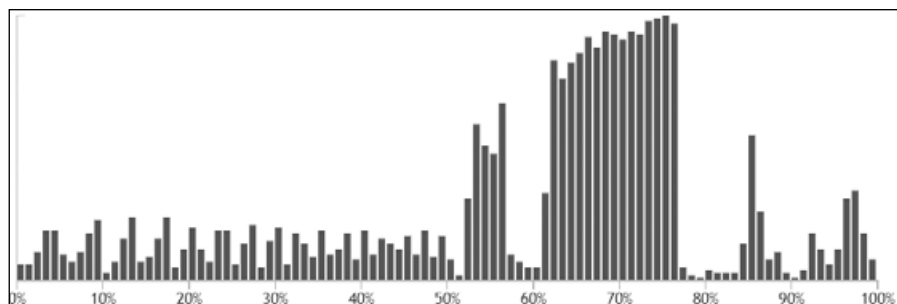
4 <http://www.crol.hr/forum/viewtopic.php?t=5766>

5 https://the.sketchengine.co.uk/corpus/corp_info?corpname=preloaded/childes_hrv&struct_attr_stats=1&subcorpora=1

6 <http://riznica.ihj.hr/philologic/Cijeli.whizbang.form.hr.html>

7 https://the.sketchengine.co.uk/corpus/freq_distrib?q=q%5B%3D%22sranje%22+%7C+lemma_lc%3D%22sranje%22%5D;corpname=preloaded%2Friznica_reldi&refs=%3Ddoc

Prikaz 2. Distribucija čestotnosti leksema *sranje* u hrvac korpusu (47,062 pojavnica, 33.70 na milijun).



Čestotnost uporabe leksičkih jedinica iz rječnika identificirana je u korpusu hrvac22 (Ljubešić i Klubička, 2014) korištenjem SketchEngine API⁸ i pohranjivanjem podataka u lokalnu bazu podataka (Neo4j). Podaci o čestotnosti, relativnoj čestotnosti u hrvac korpusu za imenice klasificirane kao tabuizirane, vulgarne, i derogatorne riječi prema Bujasu navedene su u tablicama 2-4.

Tablica 2. Čestotnost imenica u korpusu hrvac klasificiranih kao tabuizirane riječi (Bujas 2001).

	Lema	Pojavnica na milijun	Čestotnost
1	kurac	28.26384	39506
2	govno	16.59587	23197
3	pizdarija	9.684798	13537
4	pička	9.373586	13102
5	šupak	3.816112	5334
6	pica	2.90823	4065
7	ševa	2.056151	2874
8	jebač	1.928804	2696
9	jebanje	1.389368	1942
10	jebivjetar	0.866388	1211
11	jebačina	0.731887	1023
12	govnar	0.688961	963
13	drkanje	0.590947	826
14	ševac	0.21606	302
15	liz	0.207475	290
16	redaljka	0.14094	197
17	jebačica	0.137363	192
18	govnarija	0.06582	92

⁸ [title&iquery=sranje](https://www.sketchengine.co.uk/documentation/api-documentation/)

⁸ <https://www.sketchengine.co.uk/documentation/api-documentation/>

	Lema	Pojavnica na milijun	Čestotnost
19	drkadija	0.060096	84
20	fukanje	0.036487	51
21	mindā	0.025756	36
22	jebica	0.013593	19
23	fafanje	0.012162	17
24	ševica	0.011447	16
25	jebežljivost	0.00787	11
26	jebarnik	0.007154	10
27	jebežljivac	0.003577	5
28	buzurant	0.001431	2

Tablica 3. Čestotnost imenica u korpusu hrwac, klasificiranih kao vulgarne riječi (Bujas 2001).

	Lema	Relativna čestotnost	Čestotnost
0	sranje	33.6675	47059
1	guzica	16.55437	23139
2	zajebancija	8.882799	12416
3	kurva	8.772623	12262
4	dupe	8.128019	11361
5	pizda	5.116052	7151
6	drek	3.711659	5188
7	guz	2.773729	3877
8	seronja	1.774986	2481
9	pimpek	1.521723	2127
10	bljuvotina	1.402246	1960
11	zajebant	0.578784	809
12	isprdak	0.537289	751
13	prdac	0.439275	614
14	podjebavanje	0.429974	601
15	zajebavanje	0.392772	549
16	kurba	0.171704	240
17	radodajka	0.134501	188
18	hračak	0.133786	187
19	kurviš	0.13307	186
20	kurvar	0.088714	124
21	kenjaža	0.065104	91
22	popišanac	0.055804	78
23	kurbar	0.010731	15
24	fanfulja	0.007154	10

	Lema	Relativna čestotnost	Čestotnost
25	guzara	0.002146	3
26	dupetara	0.000715	1
27	frnjoklica	0.000715	1
28	guzatost	0.000715	1

Tablica 4. Čestotnost imenica u korpusu hrvac klasificiranih kao derogatorne riječi (Bujasu 2001).⁹

	Lema	Pojavnica na milijun	Čestotnost
1	peder	13.08238	18286
2	debil	9.671921	13519
3	zubo ⁹	8.140897	11379
4	komunjara	6.179899	8638
5	seljačina	4.45857	6232
6	soj	3.506331	4901
7	manipulator	2.604887	3641
8	balavac	2.53549	3544
9	balkanac	2.48684	3476
10	mrcina	1.862269	2603
11	kreatura	1.302086	1820
12	bijednik	1.282054	1792
13	raga	1.20121	1679
14	selendra	1.184755	1656
15	mizerija	1.1683	1633
16	pandur	1.153991	1613
17	glavešina	1.125374	1573
18	ciganka	0.869249	1215
19	njuškalo	0.867103	1212
20	babetina	0.726163	1015
21	glumatanje	0.71257	996
22	novinarčić	0.710424	993
23	piskaralo	0.706131	987
24	govorancija	0.65462	915
25	aparatčik	0.580215	811
26	klipan	0.560898	784
27	dotepenac	0.553029	773
28	starkelja	0.522265	730
29	krntija	0.520119	727
30	pijandura	0.517973	724

⁹ Leksem *zubo* ima veliku čestotnost jer se pojavljuje kao leksički oblik emotikona *:zubo:* u forumima. Primjerice: „*A nisam ja kriv kaj me majka lepog rodila :zubo:*“.

	Lema	Pojavnica na milijun	Čestotnost
31	kljuse	0.485063	678
32	glavonja	0.444998	622
33	kretenčina	0.394203	551
34	skot	0.364155	509
35	zrikavac	0.340545	476
36	strančarenje	0.302628	423
37	kravetina	0.296189	414
38	krele	0.286173	400
39	serator	0.265425	371
40	nosonja	0.26471	370
41	pripuz	0.232515	325
42	pijančina	0.22393	313
43	fiškal	0.213914	299
44	ženska	0.203898	285
45	žbir	0.197459	276
46	ženturača	0.193167	270
47	prčvarnica	0.189589	265
48	grmalj	0.177427	248
49	nakot	0.173134	242
50	vodurina	0.171704	240
51	šminkeraj	0.139509	195
52	sladunjavost	0.129493	181
53	četnikuša	0.125916	176
54	ciganluk	0.120908	169
55	partizanština	0.118046	165
56	dodoš	0.114469	160
57	digić	0.113038	158
58	partizancina	0.110176	154
59	jugoslavenčina	0.105884	148
60	prizemnost	0.104453	146
61	ciganin	0.103022	144
62	Azijat	0.098014	137
63	krpetina	0.096583	135
64	nikogović	0.092291	129
65	srbenda	0.087283	122
66	droljetina	0.086567	121
67	balonja	0.083706	117
68	balkanština	0.08299	116
69	podrepaš	0.082275	115

	Lema	Pojavnica na milijun	Čestotnost
70	muktaš	0.081559	114
71	kenjkanje	0.077267	108
72	fifca	0.076551	107
73	Jugovina	0.072974	102
74	učo	0.070828	99
75	rakijetina	0.070112	98
76	cmakanje	0.069397	97
77	kiseliš	0.069397	97
78	balavander	0.06582	92
79	mantijaš	0.061527	86
80	posranac	0.061527	86
81	mitingaš	0.056519	79
82	uličarka	0.052942	74
83	krčmetina	0.051511	72
84	profuknjača	0.051511	72
85	cincar	0.048649	68
86	poluinteligent	0.045072	63
87	fakultetlija	0.040064	56
88	trčkaralo	0.040064	56
89	prištavac	0.037202	52
90	prdavac	0.036487	51
91	cendravac	0.035772	50
92	prikrpa	0.035772	50
93	zriko	0.035772	50
94	oficirčina	0.033625	47
95	prnjavor	0.031479	44
96	goljo	0.027902	39
97	novotar	0.025756	36
98	balkanka	0.02504	35
99	režimlija	0.02504	35
100	kuferaš	0.024325	34
101	cerekalo	0.023609	33
102	rječetina	0.022894	32
103	oficirčić	0.022178	31
104	kljusina	0.021463	30
105	pederstvo	0.020748	29
106	frazer	0.01717	24
107	strančar	0.01717	24
108	glondža	0.015739	22

	Lema	Pojavnica na milijun	Čestotnost
109	Čifut	0.015024	21
110	laloka	0.015024	21
111	pisarija	0.015024	21
112	kaputaš	0.014309	20
113	klempto	0.014309	20
114	bukvan	0.012878	18
115	ciganjenje	0.011447	16
116	parolaš	0.011447	16
117	faćkanje	0.010016	14
118	trbonja	0.010016	14
119	kimavac	0.008585	12
120	mućkaroš	0.00787	11
121	kalkulantstvo	0.007154	10
122	crnja	0.006439	9
123	koferaš	0.006439	9
124	kožurina	0.006439	9
125	potrkalo	0.006439	9
126	drljotina	0.005008	7
127	parolaštvo	0.005008	7
128	purgerstvo	0.005008	7
129	reformaš	0.005008	7
130	klempavost	0.004293	6
131	titulaš	0.004293	6
132	kenjkavac	0.003577	5
133	pašće	0.003577	5
134	popišulja	0.003577	5
135	strvinarenje	0.003577	5
136	frazerstvo	0.002862	4
137	kmezavica	0.002862	4
138	krezo	0.002862	4
139	naftalinac	0.002862	4
140	gegalo	0.001431	2
141	kenjkavost	0.001431	2
142	zuba	0.001431	2
143	čelonja	0.000715	1
144	črčkalo	0.000715	1
145	grbonja	0.000715	1
146	kmezavac	0.000715	1
147	novotarstvo	0.000715	1

Podaci o čestotnosti govore o konvencionaliziranosti određene jezične jedinice u komunikacijskom sustavu distribuirane kulturne kognicije. Konvencionalnost se stoga odnosi na mjeru usvojenosti određenog uzorka jezičnih iskaza i pokazatelj je kulturnog modela određene govorne zajednice u promatranim komunikacijskim stilovima. Međutim, jednodimenzionalne vrijednosti o pojavnosti leksika u smislu izdvojenih pojmova svakako nisu jedini pokazatelj pojmovne strukture, te je za cjelovitiji uvid o uporabi i kulturnim modelima potrebno svaku od jezičnih jedinica istražiti u odnosu na ostale jezične jedinice s kojima gradi značenje. U sljedećem odjeljku prikazuje se korpusna metoda sintaktičko-semantičke analize kompleksnosti opojmljivanja neprimjerenih pojmova pomoću konstrukcijskog pristupa.

5. Konstrukcijski pristup

Konstrukcijski pristup proistječe iz kognitivno lingvističke teorije o konvencionalizaciji značenja u sintaktičko-semantičkim konstrukcijama. Sintaktičko ustrojstvo nastaje jezičnom uporabom i funkcionalno je jer omogućuje smisleno procesiranje perceptivno-informacijske, afektivne i kognitivne razine, odražavajući ontološki odnos pojmovnih referenata i određujući pojmovni sadržaj određenog iskaza. Jezične strukture na taj način opisuju zamijećena ili zamišljena svojstva entiteta, njihove međusobne interakcije i procese. Ontološki stabilne mreže svojstava, entiteta, odnosa i procesa konvencionaliziraju se na razini jezične uporabe, pa se analizom određenih konstrukcija može steći uvid u konvencionalizirane obrasce opojmljivanja stvarnosti za određenu jezičnu zajednicu i kulturu.

Ovaj se rad usredotočio na analizu neprimjerenih imenskih leksema u koordiniranoj jezičnoj konstrukciji [$x^{\text{tabu, derogativ, pejorativ}}$ *i/ili* y] koja ima logičku funkciju spajanja entiteta, a na korpusnoj razini daje uvid o opojmljivanju ontološki srodnih entiteta. Pretpostavka je da će se uvidom u pojmovne mreže apstrahirane korpusnom ekstrakcijom najčešće povezanih imenskih leksema koordiniranom konstrukcijom prikazati umreženja pojmovno srodnih neprimjerenih pojmova i moći klasificirati domene neprimjerenih riječi.

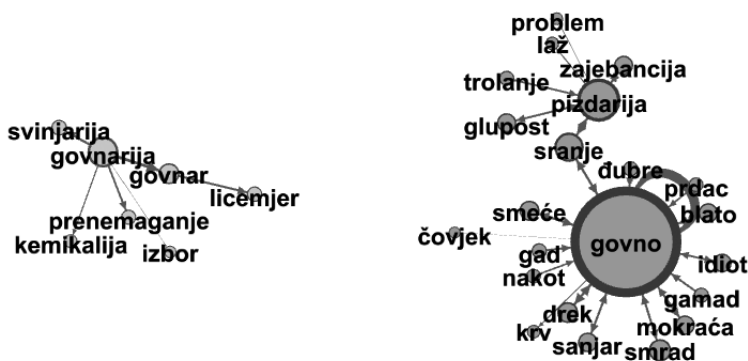
Za potrebe istraživanja korištena je Python skripta namijenjena za ekstrakciju jezičnih podataka u hrwac korpusu na SketchEngine poslužitelju, te pohranu podataka o primjerima, čestotnosti i strukturi gramatičkih konstrukcija. Navedenim se slijedom računalnih postupaka za svaku neprimjerenu imenicu slao upit na SketchEngine API, te identificiralo do 50 ciljnih imenskih supojavnica u koordiniranoj konstrukciji. Podaci o čestotnosti i mjeri povezanosti potom su pohranjeni u jezični model lokalne graf baze Neo4j (<http://neo4j.com>) što je omogućilo stvaranje upita o umrežavanju pojedinačnih (prikaz 3) i višestrukih pojmova (prikaz 4). Struktura umreženosti supojavnica koordinirane konstrukcije prikazala se slanjem podataka o čvorovima, vezama i snazi veza u program Gephi (<https://gephi.org/>).

Iz graf prikaza, stvorenih programom Gephi (<http://gephi.org>), može se zamijetiti da određeni pojmovi tvore snažnija umreženja. U koordiniranom umreženju gore navedenih 152 pojmova identificiralo se 15 snažnije povezanih zajednica (klastera) pomoću Louvain algoritma (Blondel i dr. 2008) (prikaz 5).

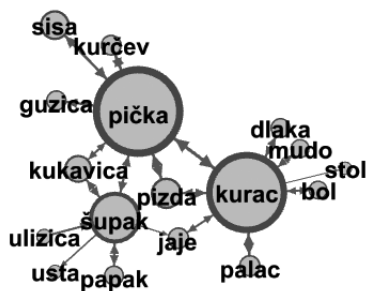
Prikaz 5. 15 klastera tabuiziranih supojavnica koordinirane konstrukcije



1. Osobe vezane uz samozadovoljavanje
2. Procesi vezani uz seks koji se naplaćuje ili pornografiju



3. Nečiste i nekorisne osobe, stvari, procesi i ishodi
4. Nečistoće, izlučevine, pojmovi i komunikacija



5. Seksualni organi



6. Mjesto za seksualne aktivnosti



7. Psihološka i biološka stanja vezana uz seksualne aktivnosti



8. Učesnici seksualnih aktivnosti



9. Nazivi za nekorisne osobe



10. Seksualna predigra

minda

11. Naziv za vulvu



12. Seksualni organi, hrana vezana uz istopisnicu naziva za ženski spolni organ



13. Grupne seksualne aktivnosti



14. Naziv za seksualne aktivnosti i pojmovi ptica, osobnih imena¹⁰ povezanih zbog istopisnice



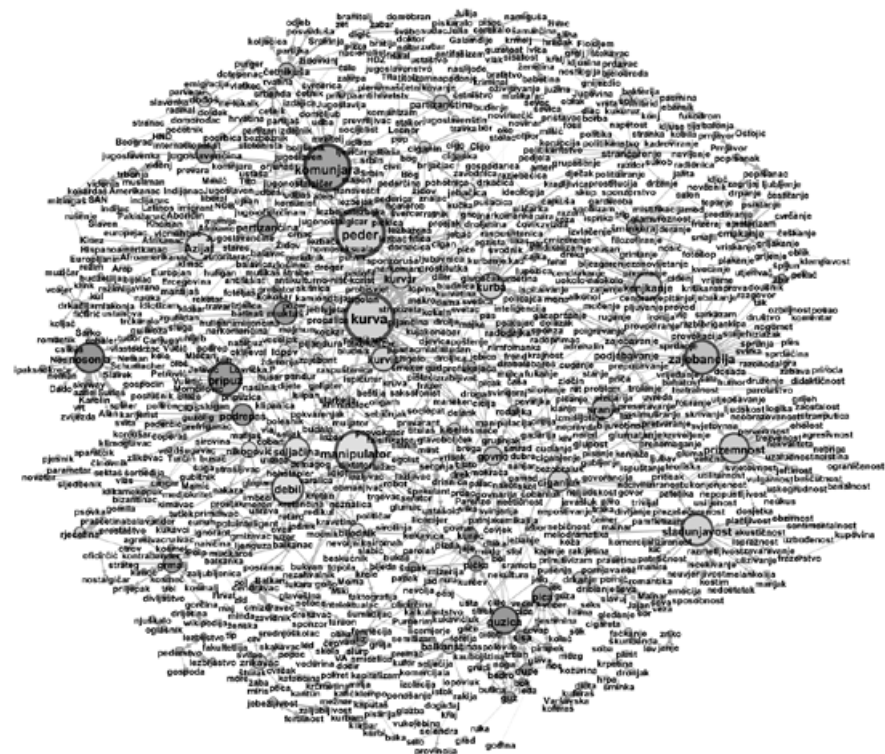
15. Akteri spolnih aktivnosti.

Povezivanje tabuiziranih leksema u klaster koordinirane konstrukcije potvrđuje seksualnu aktivnost kao glavnu tabu temu. Dodatna vrijednost klasteriranja pojmova proizlazi iz pod-kategorizacije teme seksualnosti. Pod-kategorizirane teme poput seksualnih aktera, nominaliziranih aktivnosti i njihovih objekata zabilježene su u opisu svakog klastera povezanih leksema (prikaz 5). Zamjetno je i da se svi pojmovi naizgled ne uklapaju u tematizaciju klastera. Djelomično su tome razlog greške koje nastaju lematizacijom ili pri mašinskom (ne)razlučivanju istopisnica (poput imenica *ševa*, *pica*). Tu su i otvorenosti tumačenja leksičkog značenja, poput *redaljka*, u kojem se tek u kontekstu može zaključiti radi li se o prisilnoj ili svojevoljnoj naravi grupne seksualne aktivnosti. Naposljetku, manje uklopivi leksemi poput *muzičar* u klasteru 1 samozadovoljavača, svjedoče o komunikacijskim metodama prijenosa značenja i metonimijskog proširenja umreženja leksičkog značenja neutralnih ili pozitivnih riječi prema negativno obilježenim tematskim domenama. U smislu identifikacije pojmovno raznorodnih leksema u koordiniranom klasteru se ovaj postupak može koristiti za psiho-lingvističku analizu procesiranja pojmovnih dvoznačnosti ili za određivanje metonimijskoga i metaforičkoga prijenosa značenja.

Ukoliko primijenimo postupak gradbe koordinirane pojmovne mreže za sve identificirane neprimjerene riječi u Bujasovu rječniku, povezujući svaku tabuiziranu, vulgarnu i derogativnu riječ iz tablica 2-4 sa svim imenskim supojavnicama u koordiniranoj konstrukciji zabilježenim u hrwac22 korpusu, dobivamo graf sa 1065 imenica (čvorovi) povezanih sa 1818 veza.

10 U mreži se pojavljuju *Ševa*, *Jajan* i *Malnar*, akteri popularnog showa „Noćna mora“ zbog nerazlučivanja osobnih entiteta od imenskog pojma *ševa*.

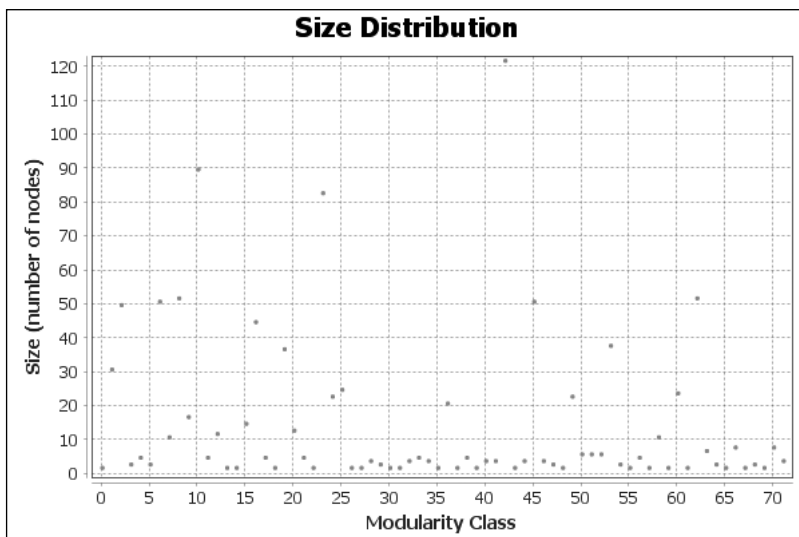
Prikaz 6. Graf prikaz koordinirane konstrukcije [x i/ili y] tabuiziranih, vulgarnih i dero-gatornih lema navedenih u tablicama 2-4. Graf je izrađen pomoću Force Layout prikaza koji približava snažnije povezane čvorove. Veličina čvora označava zbroj mjera snage veza i broj veza s drugim čvorovima.



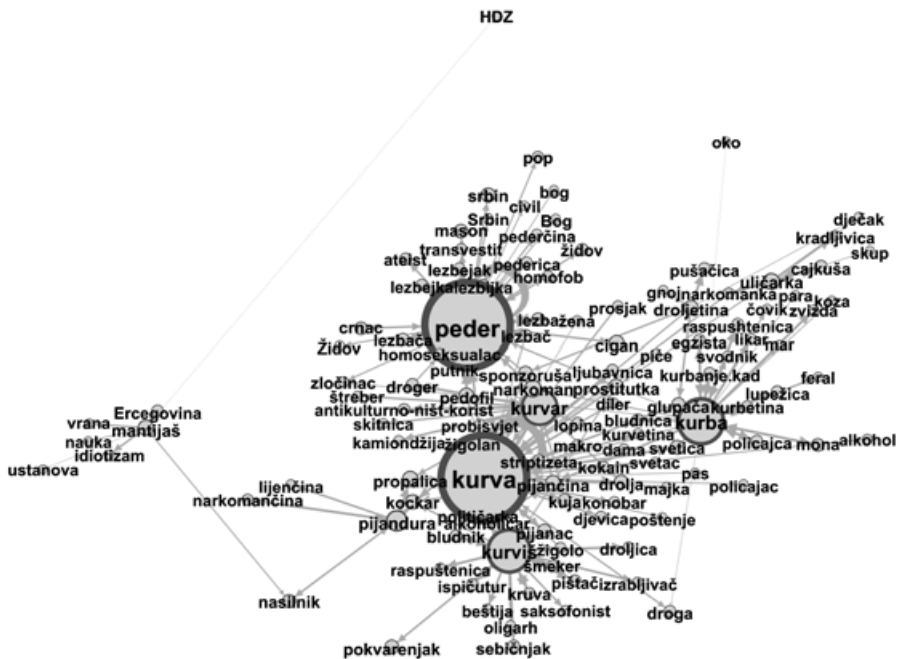
Primjenom Louvain algoritma¹¹ u gornjem grafu identificirane su 72 zajednice s modularnom rezolucijom 0.84, od kojih samo 12 zajednica ima više od 30 čvorova (prikaz 7).

11 Louvain algoritam odnosi se na metodu kojom je moguće izmjeriti gustoću veza unutar zajed-nica u usporedbi s vezama između samih zajednica: <http://research.pomona.edu/complex-networks/2015/10/02/louvain-algorithm/>

Prikaz 7. Analiza modularnosti grafa tabuiziranih, vulgarnih i derogatornih lema. Na x osi nalaze se brojevi modula zajednica, na y osi naveden je broj čvorova u zajednici.



U nastavku se predstavlja osam pojmovnih klastera koji imaju 50 ili više čvorova. Najveća zajednica pojmova povezanih koordiniranom konstrukcijom s centralnim čvorovima *peder*, *kurva*, *kurba*, *kurviš* (prikaz 8) odnosi se na osobe sklone neprimjerenoj seksualnoj djelatnosti prema standardima hrvatske kulture, nasilništvu, krađi i varanju. *Kurva* i *peder* istaknute su leme u ovoj domeni, pri čemu je moguće zapaziti veze s pojmovno oprečnim pojavnicama. Primjerice *kurva* i *drolja* koordiniraju se s pojmovima *dama*, *svetica* i *djevica* koje obilježavaju *poštenje* i nevinost unutar diskursa dominantno vezanog uz kršćanstvo. Nadalje, unutar klastera seksualnosti uočava se povezanost s entitetima vezanim uz devijantno ponašanje poput *pijanac*, *diler* i *kokain*.



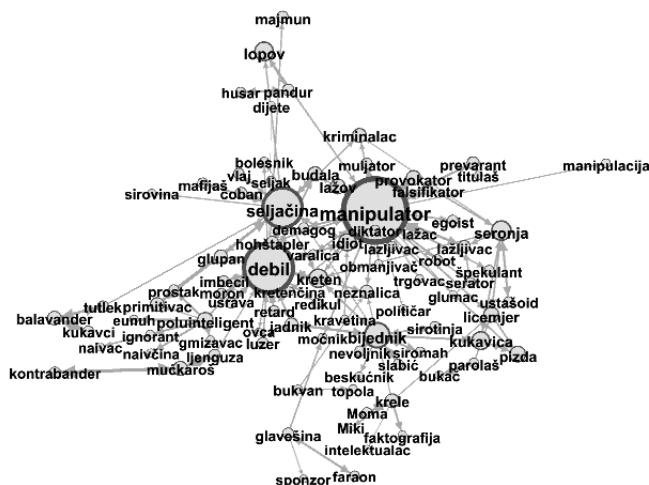
Skupina pojmova povezana leksimima *sladunjavost*, *ciganluk*, *ševa* odražava vrstu interakcija koje se smatraju neciviliziranim, razmetljivim ili ispraznim.

Prikaz 10. Graf prikaz klastera *ciganluk*, *sladunjavost*, *ševa* koordinirane konstrukcije.



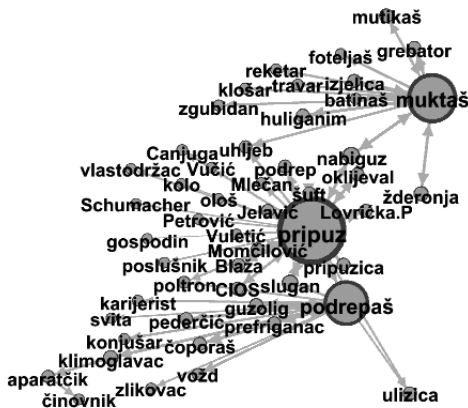
Pojmovi povezani centralnim leksimima *manipulator*, *lopov*, *bijednik*, *seljačina*, *debil* odnose se na identitetska i psihološka obilježja osoba koji uključuju širi raspon nepoželjnih osobina poput neznanja i kognitivnih poteškoća do moralno osuđenih i sankcioniranih značajki. Pojam *manipulator* centralni je čvor koji kroz svoje veze ukazuje na psihološko obilježje neiskrene i koristoljubive interakcije s drugim osobama koje naposljetku tvore pogrdni identitet manipulatora. U ovome se klasteru nalaze i manipulirane osobe. To su primjerice imenice *kreten* povezan sa srodim imenicama *debil*, *idiot* ali i *seljačina*. Tu se nalazi i leksim za ženski spolni organ *pizda* s očitim prijenosom značenja uz karakterne osobine koje primarno opisuju slabu osobu, plahu osobu, obilježenu i nazivom za pticu *kukavicu*. Osim toga se leksema za životinje pojavljuju *ovca* i *majmun* koji se povezuje s *debilom* i *pederom* iz drugog klastera.

Prikaz 11. Graf prikaz klastera *manipulator*, *lopov*, *bijednik*, *seljačina*, *debil* koordinirane konstrukcije.



Klaster s leksemima *pripuz*, *muktaš*, *podrepaš* odnosi se na nepoželjne načine pri-bavljanja egzistencije i radnih navika. Tu se nalaze pojmovi poput *uhljeb*, *šuft*, *gre-bator*, *foteljaš* ali i *karijerist* te *vlastodržac*.

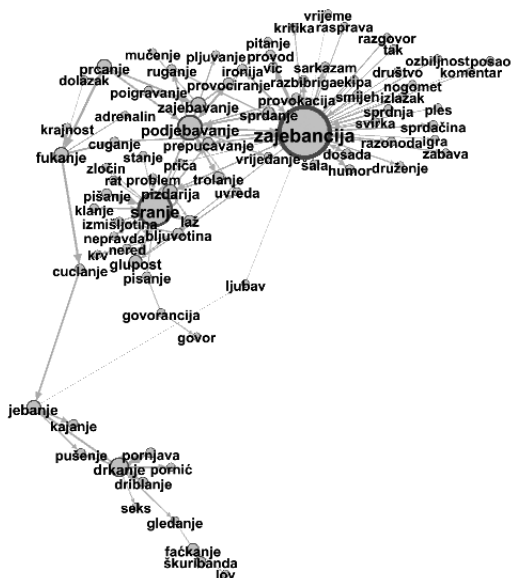
Prikaz 12. Graf prikaz klastera *pripuz*, *muktaš*, *podrepaš* koordinirane konstrukcije.



Klaster s centralnim pojmovima *zajebancija*, *sranje*, *jebanje* odnosi se na provođe-nje slobodnog vremena, nepoželjnu vrstu komunikacije te na aspekte seksualnih aktivnosti. *Zajebancija* označava interakciju koja je vezana uz *glupost*, *pizdariju*, *sprdnju*, *provod* i *razbibrigu*. *Sranje* se u ovome klasteru ne veže uz ishodišnu temu fekalija, već preneseno vezuje, kao i *pizdarija* uz *izmišljotinu* i *glupost*. Nepoželjnost

takvih interakcija i komunikacija određena je poveznicama sa leksemima *klanje*, *patnja*, *rat*, *zločin* i *nered*.

Prikaz 13. Graf prikaz klastera *zajebancija*, *sranje*, *jebanje* koordinirane konstrukcije.



Zajednica pojmova umreženih koordiniranom konstrukcijom oko centralih leksema *kurac*, *pica*, *guzica* vezan je uz temu seksualnih organa i tijela, ali zbog istopisnice *pica* privlači i lekseme vezane uz hranu.

Prikaz 14. Graf prikaz klastera *guzica*, *pica*, *kurac* koordinirane konstrukcije



Grafovi prikazuju povezanost pojmovnog sadržaja koji se može upotrijebiti za semantičku i leksičku analizu neprimjerenih riječi, kao i ostalih leksema koji povezivanjem s tabuiziranim, vulgarnim ili derogatornim izrazima asocijativno poprimaju nova neprimjerena značenja. Ovakav pristup detaljnije opisuje ishodišne teme neprimjerenog jezika od pojednostavljene kategorizacije koju predlaže Ljung (2011), te se može upotrijebiti za kulturološku analizu neprimjerenih domena. Također, s obzirom na uočenu važnost povezanosti oprečnih pojmova, smatramo da bi se analiza neprimjerenih riječi i pojmova uspješno uklopila u empirijske kulturološke analize modela primjerene komunikacije u određenoj govornoj zajednici, obrazaca društvene interakcije i procese gradbe pojedinačnih ili grupnih identiteta (Hall 2000; Lotman 1990).

6. Zaključak

Rad predstavlja empirijski pristup analizi tabuiziranih, vulgarnih i derogatornih imenica u hrvatskome jeziku. Primjenom leksikografske analize identificirano je 257 imenica te je zatim korpusnom analizom provjerena njihova čestotnost i ekstrahirani podaci o njihovoj uporabi u koordiniranoj konstrukciji [*x* i/ili *y*]. Na temelju graf analize pojmova umreženih koordiniranom konstrukcijom istaknuta su semantička i ontološka obilježja neprimjerenih imenskih leksema kao i povezanost sa drugim leksemima tvoreći tako značenjski sustav od 1065 umreženih pojmova. Pomoću algoritma za pronalaženje modularnosti identificirani su pojmovni klasteri za koje se može tvrditi da predstavljaju ontološko-semantičke domene i prikazuju sustave prototipnih sintaktičko-semantičkih umreženja. Korpusni pristup i metodološki postupci tvore čvrsti temelj za daljnja empirijska istraživanja neprimjerenog jezika te određenja specifičnih komunikacijskih stilova i kulturalnih modela unutar hrvatskog sociokulturalnog konteksta. Istraživanja je svakako potrebno proširiti na kroz-/inter-kulturalnoj razini te stvoriti sistemsko objašnjenje bio-psiho-socijalnih modela koji mogu pomoći pri interpretaciji i shvaćanju neprimjerenosti kao specifičnog komunikacijskog fenomena te njegovih funkcionalnih svojstava u interakciji i procesima izgradnje identiteta.

Literatura

- Allan, K.; Burrige, K. 2006. *Forbidden Words. Taboo and the Censoring of Language*. New York: Cambridge University Press.
- Anić, V. (1998). *Rječnik hrvatskoga jezika*. Novi liber.
- Badurina, L; Pranjković, I. 2016. Jezična i pragmatična obilježja psovke. *Romanoslavica*. 52, 2; 227-235.

- Blondel, V. D., Guillaume, J. L., Lambiotte, R., & Lefebvre, E. (2008). Fast unfolding of communities in large networks. *Journal of statistical mechanics: theory and experiment*, 2008(10), 1000-8.
- Bujas, Ž. (2001). *Veliki hrvatsko-engleski rječnik: Croatian-English dictionary*. Nakladni zavod Globus.
- Bujas, Ž. (2008). *Veliki hrvatsko-engleski riječnik: najveći i najsuvremeniji hrvatsko-engleski rječnik*. Nakladni zavod Globus.
- Douglas, M. 2002. [1966] *Purity and Danger: An Analysis of the Concepts of Pollution and Taboo*. London and New York: Routledge.
- Gavran, I. 1962. *Bludna psovka: povijesno-psihološka studija*. Sarajevo: Udruženje katoličkih svećenika NR BiH.
- Hall, S. 2000. Who needs identity? U: Gay, P; Evans, J; Redman, P. (ur.), *Identity: A Reader*. London: Sage Publications, 15-30.
- Jay, T. 2000. *Why we curse: a neuro-psycho-social theory of speech*. Philadelphia - Amsterdam: John Benjamins Publishing Company.
- Jojić, L. (ur.). (2015). *Veliki rječnik hrvatskoga standardnoga jezika*. Zagreb: Školska knjiga.
- Lotman, M. Yuri 1990. *Universe of the Mind: A Semiotic Theory of Culture*. London: I. B. Tauris
- Lučić, R. 2015. Kako to tamo psuju? Vulgarizmi u hrvatskom i nizozemskom. U: S. L. Udier, K. C. Kovačević (ur.), *Višejezičnost kao predmet multidisciplinarnih istraživanja*, Zbornik radova s međunarodnoga znanstvenog skupa Hrvatskoga društva za primijenjenu lingvistiku održanoga od 25. do 27. travnja 2014. u Zagrebu. Zagreb. Srednja Europa, Hrvatsko društvo za primijenjenu lingvistiku, 583-596.
- Ljung, M. 2011. *Swearing. A Cross-Cultural Linguistic Study*. Palgrave Macmillan.
- Memišević, A; Matešić, M. 2014 'Pardon my English': usage and meanings of select English words and phrases in Croatian. *Višejezičnost kao predmet multidisciplinarnih istraživanja, XXVIII. međunarodni znanstveni skup Hrvatskoga društva za primijenjenu lingvistiku*, Zagreb, Hrvatska, 2014. (predavanje, neobjavljeni rad, cjeloviti rad (in extenso), znanstveni).
- Mikić, P; Pehar, M; Mikić, M. 1999. *Psovka u hrvatskome i njemačkome jeziku*. Mostar: Ziral.
- Perak, B. 2017. „Pojmovne mreže leksema neugodnih i nepobuđenih emocija. Usporedba pojmova tuga, jad, žalost, depresija i sjeta u koordiniranoj konstrukciji“. U: *Jezik kao predmet proučavanja i jezik kao predmet poučavanja. Zbornik radova s međunarodnoga znanstvenog skupa Hrvatskoga društva za primijenjenu lingvistiku održanoga od 12. do 14. svibnja 2016. godine u Rijeci*. Stolac, Diana i Anastazija Vlastelić (ur.). Zagreb: Srednja Europa.

- Pilch, P. 2011. *Psovka u hrvatskome i češkome jeziku*. Završni rad. Brno: Masarykovo Sveučilište u Brnu.
- Užarević, J. 2012. Fenomenologija psovke. U: J. Užarević, *Književni minimalizam*, 167–184. Zagreb: Disput.
- Vingerhoets, Ad; Bylsma M. Lauren; Vlam, de Cornelis 2013. Swearing. A Biopsychosocial Perspective. In: *Psychological Topics*, Vol. 22, No. 2, 287-304.
- Hrvatski jezični portal <http://hjp.znanje.hr> (<http://hjp.znanje.hr/index.php?s-how=baza>)

About shit, and other (in)appropriate things:

A corpus based lexical analysis and the graph construction of semantic domains of taboo, vulgar and derogatory words in the Croatian language

By means of applying a lexicographic and corpus methodology the current paper approaches the process of identification of the semantic domains of tabooed, vulgar, and derogatory words and classifies them based on the syntactic-semantic cross-link in a coordinated construction (*x i/ili y*). Inappropriateness is ontologically explicated in terms of being a specific part of a cultural system of communication strategies in social interaction between interlocutors with the aim of the cognizable activation of cognitive-affective processes. Within the dynamic psychosocio-cultural system particular accounts of conceptualization of social interaction are observed as inappropriate and consequently deemed as tabooed. A cognitive-linguistic corpus analysis of swearwords in the Croatian language is carried out by using the hrWac corpus. Based on the graph analysis of notions cross-linked in the coordinated construction the emphasis lies on the semantic and ontological characteristics of inappropriate noun lexemes as well as the connectedness with other lexemes, which subsequently forms a semantic system of 1065 cross-linked notions. The paper contributes to the further development of the foundation for the qualitative analysis of swearwords based on selected lemmas in a coordinated construction, and indicates towards a need for the development of an overarching methodology for the intercultural and cross-cultural research of inappropriate language.

Keywords: swearwords; inappropriate language; taboo; coordinated construction; graph theory