

Benedikt Perak

# Emocije u korpusima: Konstrukcijska gramatika i graf-metode analize izražavanja emotivnih kategorija

---

Rad daje uvid u istraživanje jezičnih obrazaca konceptualizacije i izražavanja emocija koje se temelji na interdisciplinarnom pristupu kognitivnih znanosti, korpusne i računalne lingvistike. Opisuje se istraživačka metoda računalne obrade korpusnih podataka i konstrukcijske analize konceptualnih mreža (ConGraCNet, <http://emocnet.uniri.hr/congracnet/>) koja uključuje primjenu graf-algoritma za identifikaciju leksičke polisemičnosti emocionalnih kategorija i kvalitativno-kvantitativni opis semantičkih domena u različitim korpusima i jezicima na primjeru hrvatskih bliskoznačnih riječi *sram* i *stid* u korpusu hrWaC i određivanja njihove srodnosti s engleskim leksičkim pojmom *shame* u korpusu enTenTen13.

## Uvod

Jezično izražavanje emocija čini temelj osobnog identiteta, društvenih interakcija i komunikacije. Fenomen je to kojim se bave različite grane društveno-humanističkih istraživanja, od kognitivno utemeljenih pristupa emocijama (Lindquist *et al.* 2012; Barrett 2017a, 2017b; Sander *et al.* 2018; Shuman *et al.* 2017; Scherer 2009) i jeziku (Lüdtke 2015; Kotz i Paulmann 2011; Foolen *et al.* 2012; Kövecses 2012) do diskurzne analize (Gee 2004; Hart 2010; Edwards 1999; White 2010; Carbaugh 2007). Proučavanje komunikacije emocija ima sve značajniju ulogu u analizi procesiranja teksta (Jacobs 2016; Swain 2013; Lindquist i Gendron 2013; Schuller i Batliner 2013; Gmuer *et al.* 2015), afektivnom računalstvu (Ahmad 2011; Cambria 2016; Mohammad 2016), komunikacijsko-informacijskim znanostima (Al-Saaqa *et al.* 2018) i razvoju aplikacija umjetne inteligencije (Martinez-Miranda i Aldea 2005; Yonck 2017; Schuller 2018; Bartneck *et al.* 2017).

U ovom se radu<sup>1</sup> daje kraći prikaz istraživanja jezičnih obrazaca izražavanja emocija koje se temelji na interdisciplinarnim uvidima kognitivnih

---

1 Rad je realiziran u sklopu znanstvenog projekta EmocNet UNIRI-HUMAN-18-243 1408, financiranog od Sveučilišta u Rijeci i projekta Formal reasoning and semantics UIP-05-2017-9219 financiranog od Hrvatske zaklade za znanost.

znanosti te metodama korpusne i računalne lingvistike. Predstavljene istraživačke metode sastavni su dio projekta *Računalni resursi, metode identifikacije i ontološko modeliranje komunikacije psiholoških stanja EmoCNet* (<http://emocnet.uni-ri.hr/>) koji okuplja istraživače usredotočene na odgonetavanje pitanja poput ovih: 1) Kako se emocije, kao kategorije utjelovljenih subjektivnih doživljaja, kodiraju u konvencionalnim strukturama hrvatskoga jezika? 2) Kako se kategorije emocija konceptualiziraju? 3) Koje kategorije emocija uopće postoje? 4) Možemo li pomoću korpusnih metoda izmjeriti semantičku srodnost leksičkih kategorija? 5) Možemo li iskoristiti znanje o semantičkoj srodnosti u istraživanjima izražavanja emocija? 6) Koliko se mijenja pojmovni sadržaj emocionalnih kategorija s obzirom na različite korpusne? 7) Možemo li usporediti pojmovni sadržaj među jezicima?

Ovaj kratki rad, naravno, neće pružiti detaljne odgovore, već tek manji primjer istraživanja koji se usustavio razvojem korpusno-konstruktivne metode leksičke analize ConGraCNet. U prvom se dijelu raspravlja o epistemološkim problemima identifikacije emocija u komunikaciji, a u drugom o metodološkim aspektima prikupljanja korpusnih podataka i konstruktivnog sintaktičko-semantičkog pristupa. Treći se dio odnosi na opis primjene graf-algoritma za identifikaciju leksičke polisemičnosti emocionalnih kategorija i semantičkih domena blisko značajnih pojmova *stid* i *sram* te njihova prijevodnog ekvivalenta u engleskome *shame*. Četvrto poglavlje donosi zaključke o mogućnostima konstruktivno-korpusnog pristupa i aplikacije ConGraCNet pri istraživanju semantičkih domena emotivnih kategorija.

## Identifikacija emocija u komunikaciji

Epistemološki problem identifikacije emocija u ljudskoj interakciji vezan je uz pitanje subjektivnosti: nitko nikada nije doživio emocije drugoga. Taj nazgled trivijalan zaključak vodi nas do pitanja: kako onda znamo koju emociju doživljava onaj drugi?

U skladu s teorijom utjelovljene spoznaje znanje o subjektivnim emocionalnim kategorijama stječe se kroz procese društvene interakcije i učenja. Prije svega kroz kognitivne procese kategorizacije, imitacije, metonimije, analogije i metaforičke inferencije. Ispravno izražavanje vlastitih i interpretiranje tuđih afektivnih stanja kulturno je uvjetovan obrazac koji zahtijeva kognitivno procesiranje niza kategorija emocija, njihovih uzroka, bihevioralnih i psiholoških učinaka te poznavanje prikladnog jezičnog koda.

Temeljne su sastavnice jezičnog izražavanja afektivnih stanja leksičke kategorije i jezične konstrukcije. One čuvaju kulturno znanje o tome što je sve moguće doživjeti, što se sve može osjećati prema čemu i naposljetku kako ta subjektivna stanja prikladno i učinkovito izraziti.

Međutim što se sve pohranjuje u leksičke kategorije? Opće iskustvo ili pojedinačni ostvaraj? Je li veza s drugim pojmovima ontološki kongruentna ili je metaforična? Što uopće znači i na što se odnosi pojam *emocija*? Naravno,

definicije se mogu pronaći u različitim udžbenicima, glosarima, rječnicima. Primjerice na Hrvatskom jezičnom portalu *emocija* je definirana ovako: „*psih.* stanje duševne pobuđenosti obilježeno skupom subjektivnih osjećaja, obično praćenih fiziološkim promjenama, koje potiču osobu na reakciju (radost, gnjev, strah, ljubav itd.)“.

Pokriva li navedena definicija sva značenja riječi *emocija*? Očito ne, što se može vidjeti iščitavanjem drugih rječnika i definicija. Nadalje, problematično je i to kako se definiraju odnosi te riječi s drugim kategorijama. Naime kakav je odnos *emocije* s pojmovima: *misao, osjećaj, doživljaj, razum, strast* itd.? Sustavna umreženost afektivnih stanja u različite sastavnice ljudskih stanja, identiteta, interakcija i institucija čini višedimenzionalan istraživački fenomen čiji znanstveno formaliziran opis nužno mora dinamično supostaviti dimenzije biološkog determinizma sa složenošću kulturnog konstruktivizma (Perak i D'Alessio Puljar 2013; Lindquist *et al.* 2015; Scherer i Fontaine 2018), evolucijsku datost individualnih potreba s institucionalizacijom društvenih pregovora i/ili kontrole (Goldie 2017; Hogget 2015). I sve to iščitavati kroz slojeve diskurzne uspostave društvenog utjecaja i moći u određenoj zajednici putem kulturne distribucije pojmovnih obrazaca u jezičnim strukturama (Sharifian 2017a, 2017b).

## Korpusna metoda

Modeliranje jezične komunikacije afektivnih stanja interdisciplinarno zahvaća široko područje suvremenih istraživanja koja su dobila zamah razvojem empirijskih metoda prikupljanja tekstova, stvaranja morfosintaktički označenih i elektronički pretraživih korpusa (Weisser 2016; Crawford i Csomay 2016) koji omogućuju računalnu obradu podataka, obogaćivanje sastavnih struktura nizom metapodataka i agregacijom informacija, čime se jezični podaci pretvaraju u baze znanja i resurs za prepoznavanje društvene dinamike i reprezentacije kulturnih vrijednosti (Mautner 2016; Tissari 2017; González 2016; Tillman i Louwerse 2018).

U korpusnoj metodi istraživanja korpus predstavlja izvor podataka o obrascima opojmljivanja u nekoj zajednici govornika. To znači da se određena kategorija ili pojmovni sadržaj promatra u kontekstu vrste društvene interakcije. Svaki korpus prikazuje specifične rezultate o ponašanju riječi, uključujući čestotnost i/ili odnose s drugim riječima. Ta se posebnost može smatrati kulturološkom informacijom o uporabi te riječi u navedenoj govornoj zajednici. U tome smislu korpuse valja promatrati iz sociolingvističke perspektive koristeći se različitim tipovima interdisciplinarnih metoda prikladnih za kroskulturnu i intrakulturnu pojmovnu, jezičnu i društvenu analizu izražavanja *emocija* u različitim vrstama komunikacijskih fenomena.

Korpusna obrada diskurza u hrvatskoj je znanstvenoj zajednici (Tadić 1996) vezana uz razvoj resursa i alata koji omogućuju primjenu suvremenih metoda pri analizi jezične reprezentacije kulturnih fenomena. Prije svega to

su računalno pretraživi i morfosintaktički obilježeni korpusi kao što su Ri-znica (Brozović Rončević *et al.* 2018) i hrWaC (Ljubešić i Erjavec 2011), alati za automatiziranu računalnu jezičnu obradu (Samardžić *et al.* 2015) i na njima nastale metode istraživanja leksičkih obrazaca (Hudeček i Mihaljević 2017). Korpusni resursi i alati izvor su za suvremena istraživanja jezičnog izražavanja koja uopćeno prikazuju čestotnost pojmova i njihovih umreženja (Despot *et al.* 2019) na temelju kojih se mogu steći uvidi o prototipnom jezičnom opojmljivanju svijeta.

Računalni korpusi pohranjuju jezične uporabe u digitalnom obliku, što otvara brojne mogućnosti istraživanja teksta. Prije svega, moguće je pretraživanje golemih količina teksta. Na taj način lako stječemo uvid u jezičnu okolinu određenog pojma i njegovu ulogu u stvaranju značenja. Konkordancije su jedna od najosnovnijih korpusnih aplikacija (primjer 1).

1. *Ti nisi tvoje emocije, već si više od toga – duša koja ima tijelo, emocije i razum da joj služe, a ne da njome gospodare* (Croatian Web (hrWaC 2.2.: <http://www.zivotnaskola.hr/vrline/zahvalnost.html>)).

Povrh toga, digitalna formalizacija teksta omogućuje obogaćivanje metapodataka. Primjerice, uporabom prirodne obrade teksta, odnosno NLP-metode (*Natural language processing*) svaka se riječ može automatski označiti prema vrsti riječi ili morfosintaktičkim obilježjima. Metatagiranje korpusa omogućuje sintaktičko-semantičku analizu konstrukcija.

Jedan od najznačajnijih resursa za istraživanje jezičnih obrazaca u hrvatskome svakako je korpus hrWaC vršne domene .hr, dostupan na repozitoriju CLARIN.SI (Ljubešić i Erjavec 2011) te portalu (<https://the.sketchengine.co.uk>). hrWaC je, barem za sada, najveći dostupni pretraživi i morfosintaktički označeni korpus hrvatskoga jezika. S njime su usporedivi *web*-korpusi ostalih europskih jezika, poput engleskoga enTenTen, talijanskoga itTenTen i njemačkoga deTenTen.

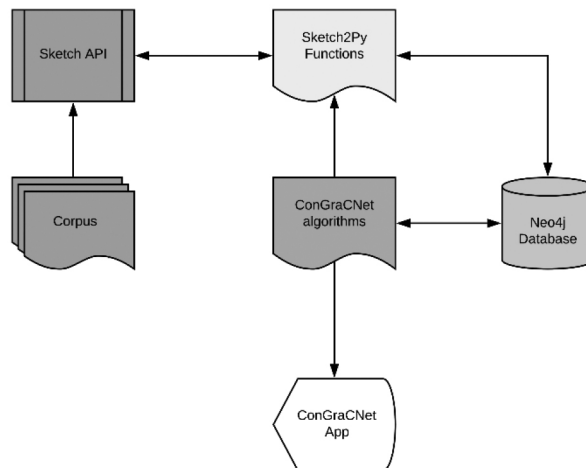
Na temelju korpusa hrWaC provedena su brojna istraživanja izravno povezana s analizom konceptualizacije i izražavanjem afektivnih stanja: strah (Perak 2014a, Perak 2015; Stanojević 2017), ljubomora (Perak 2014b), tuga i srodne kategorije (Perak 2017a), modeli višedimenzionalne konstrukcijske analize straha, ljutnje, ljubavi i ponosa (Perak 2017b), idiomatskog izražavanja (Parizoska i Petrović 2019). Nadalje, tomu se mogu dodati radovi u kojima se analizira neprimjeren jezik u hrvatskome (Perak, Damčević i Milošević 2018) koji mogu poslužiti za razvoj analize i algoritama detekcije govora mržnje.

## **Konstrukcijski model istraživanja sintaktičko-semantičkih odnosa aplikacijom ConGraCNet**

Digitalni korpusi iznimno su korisni za razvoj novih teorijskih i metodoloških razmatranja nadahnutih konstrukcijskim pristupima koji obuhvaćaju pitanja konceptualizacije i leksikografije (Hudeček i Mihaljević 2018), kogni-

tivnog uokvirivanja, uloge metonimije i metafore u konstruktivističkom pristupu opojmljivanju emocija (Perak 2018; Despot i dr. 2019), opojmljivanja društvenih kategorija metonimijskim i metaforičkim procesima koji uvjetuju procesiranje afektivnih stanja i vrijednosti (Perak 2014a, 2019a, 2019b), razvoja semantičkih ontologija (Brdar i dr. 2019, 2020), kao i razvoja novih metoda semantičkog modeliranja uporabom alata NLP (Karl i dr. 2018) i graf-algoritama za analizu sintaktičko-semantičkih konstrukcija (Perak 2020). Istraživanja pokazuju učinkovitost empirijskih korpusnih istraživanja pri formalizaciji pojmovnih mreža izražavanja emocija i mogućnosti za reprezentaciju konstrukcije značenja u višedimenzionalnoj semantičkoj mreži kulturno distribuiranoj u zajednici govornika.

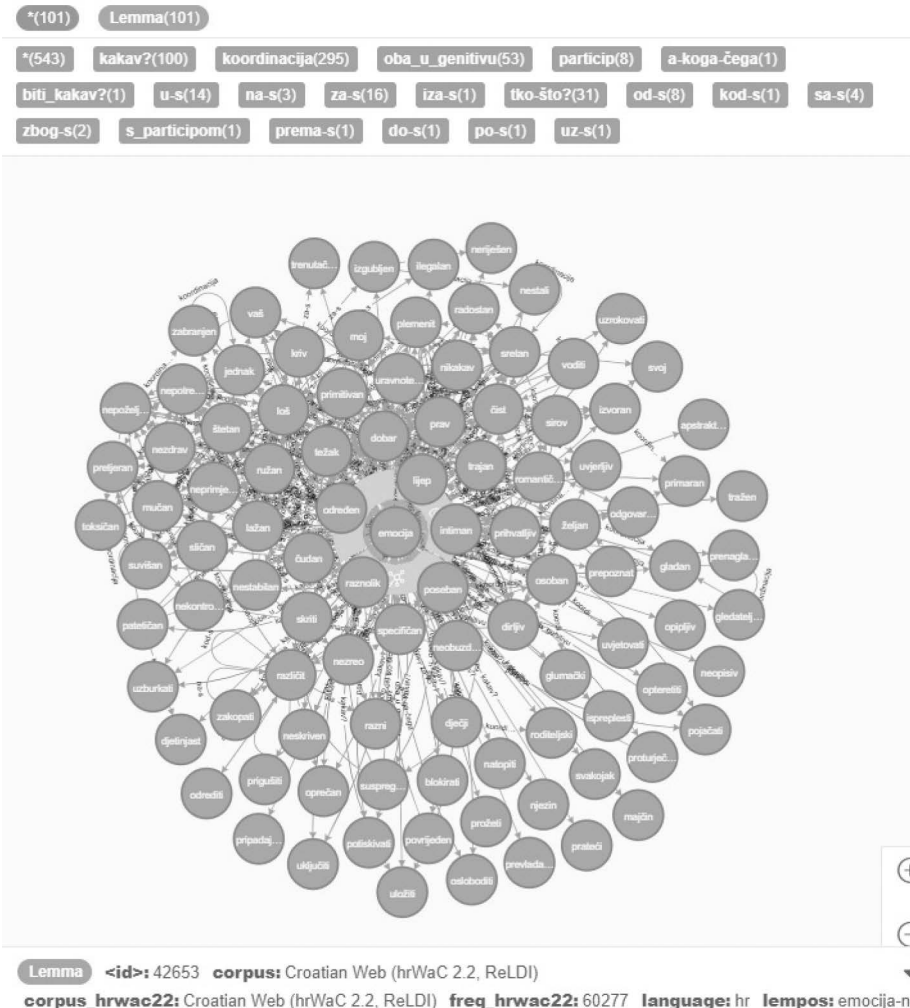
U sklopu projekta EmoCNet (<http://emocnet.uniri.hr/congracnet/>) razvijena je aplikacija ConGraCNet koja ima za cilj automatizirano stvaranje označenih korpusa, dohvaćanje podataka iz digitalnih korpusa, njihovo modeliranje, pohranu, analizu i vizualizaciju semantičko-sintaktičkih struktura. Jedna od značajki te aplikacije omogućuje dohvaćanje i algoritamsku obradu korpusnih podataka iz Sketch Enginea. Sketch Engine postao je najznačajniji portal digitalnih korpusa različitih jezika. Osobito je zanimljiva mogućnost dohvaćanja setova podataka o takozvanim skicama riječi. Skice riječi, odnosno Word Sketch predstavljaju sumarizirane kolokacije temeljnih konstrukcija odabranog leksičkog pojma uređene prema čestotnosti ili kojoj drugoj mjeri suodnosa. Aplikacija ConGraCNet Word Sketch prikuplja podatke i pohranjuje ih lokalno, umrežavajući riječi u graf-bazi na način koji je prikazan na ilustraciji 1.



**Prikaz 1.** Postupci obrade tekstova i stvaranja baze znanja aplikacije ConGraCNet Word Sketch (<http://emocnet.uniri.hr/congracnet/>)

Aplikacija ConGraCNet Word Sketch preuzima leksičke podatke iz specifičnih korpusa preko servisa Word Sketch. Zahtjev za skicama riječi šalje se

na programsko sučelje Word Sketch API. Podaci o svim skicama za određenu lemu (kanonski oblik riječi) pohranjuju se u graf-bazi Neo4j, pri čemu su leme pohranjene kao jedinstveni čvorovi sa specifičnim korpusnim svojstvima: korpus, jezik, čestotnost u korpusu, relativna čestotnost u korpusu i primjer. Različiti tipovi sintaktičkih odnosa pohranjuju se kao veze među čvorovima sa svojstvima: tip odnosa, broj supojavljivanja, mjera supojavljivanja (*score*). Jedna lema tako pohranjuje podatke o različitim odnosima u potencijalno različitim korpusima istog jezika.



**Prikaz 2.** Sučelje graf-baze Neo4j s prikazom nasumično odabranih 100 lema povezanih s ishodišnom lemom *emocija*

Sustavno umrežavanje podataka omogućuje identificiranje pojmovnih mreža emocionalnih kategorija, svojstava i procesa. Primjerice, za pojam *emocija* mogu se pronaći sljedeći sintaktički odnosi s drugim pojmovima:

kakav? 300, veznik 37, prijedlog-iza 31, koordinacija 476, prijedlog 21, koga-što 255, n-koga-čega 190, subjekt\_od 170, oba\_u\_genitivu 182, u\_genitivu-n 72, koga-čega 30, particip 39, a-koga-čega 12, biti\_kakav? 15, infinitiv 3, s\_prilogom 18, a-koga-što 6, komu-čemu 10, a-komu-čemu 4, subjekt\_je 10, u\_akuzativu 16, plus\_se 2, u\_genitivu 12, u\_dativu 2, tko-što? 109, subjekt+biti 7, u\_genitivu-a 6, u\_akuzativu-a 1, u\_dativu-a 1, s\_participom 7, za-s 3, o-s\_X 16, u-s\_X 17, o-s 5, za-s\_X 11, na-s\_X 32, bez-s 10, u-s 44, bez-s\_X 20, na-s 25, od-s\_X 11, sa-s\_X 41, do-s\_X 1, od-s 8, sa-s 20, do-s 1, nad-s 1, kod-s 6, po-s 1, iz-s 2, pod-s 1, kroz-s 1, pred-s 1, između-s 2, prema-s 7, umjesto-s 1, poput-s 6, zbog-s\_X 1, iz-s\_X 4, poput-s\_X 7, prema-s\_X 6, kroz-s\_X 1, kod-s\_X 4, uz-s\_X 2, nad-s\_X 1, pod-s\_X 2, po-s\_X 1, između-s\_X 3, umjesto-s\_X 1.

Broj koji stoji uz tip konstrukcije označava broj pohranjenih kolokacija koje su rangirane prema čestotnosti ili mjeri *score*. Primjerice, prvih 11 kolokacija prema mjeri *score* u konstrukciji koordinacija za lemu *emocija* jesu:

	source	friend	freq	score	gramRel
0	emocija-n	misao-n	502	9.2200	koordinacija
1	emocija-n	strast-n	212	8.4600	koordinacija
2	emocija-n	osjećaj-n	336	8.3500	koordinacija
3	emocija-n	razum-n	175	8.1500	koordinacija
4	emocija-n	raspoloženje-n	99	7.5000	koordinacija
5	emocija-n	um-n	127	7.4700	koordinacija
6	emocija-n	doživljaj-n	82	7.3600	koordinacija
7	emocija-n	sjećanje-n	86	7.2900	koordinacija
8	emocija-n	ponašanje-n	128	6.9800	koordinacija
9	emocija-n	psiha-n	51	6.9500	koordinacija
10	emocija-n	energija-n	157	6.8500	koordinacija

**Tabela 1.** Najistaknutije kolokacije za leksem *emocija* u sintaktičko-semantičkom odnosu koordinacije

Na temelju sistemske analize kolokacija u navedenim sintaktičkim odnosima može se mapirati njihova specifična semantička funkcija. Primjerice, jezikoslovna konstrukcija koordinacije [pojam x i/ili/ni/niti pojam y] prototipno reprezentira konvencionalno opojmljivanje ontološki srodnih pojmova, kao što je prikazano za zajednicu imenskog pojma osjećaj na temelju korpusa hrWaC (tabela 1). S druge strane, sintaktički odnos pod nazivom *kakav?* prikazuje kolokacije pridjeva kojima se ističu opojmljena obilježja leksema *emocija*.

	source	friend	freq	score	gramRel
0	emocija-n	negativan-a	2232	9.3700	kakav?
1	emocija-n	nabiti-a	380	9.0400	kakav?
2	emocija-n	potisnuti-a	233	8.3200	kakav?
3	emocija-n	pozitivan-a	1050	7.8200	kakav?
4	emocija-n	snažan-a	674	7.5200	kakav?
5	emocija-n	prepun-a	285	7.3900	kakav?
6	emocija-n	iskren-a	251	7.2900	kakav?
7	emocija-n	neugodan-a	248	7.1500	kakav?
8	emocija-n	nabijen-a	99	7.0700	kakav?
9	emocija-n	dubok-a	406	7.0400	kakav?

**Tabela 2.** Najistaknutije kolokacije za leksem *emocija* u sintaktičko-semantičkom odnosu *kakav?*

I tako redom, za svaki se tip konstrukcija može pripisati, odnosno mapirati određena semantičko-referencijalna funkcija koja reprezentira prototipno opojmljivanje pojma za određeni korpus. Detaljan opis mapiranja svih sintaktičkih konstrukcija i njihovih semantičkih funkcija izvan je dosega ovog rada, ali valja naglasiti da se međusobni odnos konstrukcija može sagledavati u hijerarhijsko-emergentnom odnosu (Perak 2014a, 2019). To znači da su određene konstrukcije ontološki i konceptualno jednostavnije i kao takve čine osnovu za hijerarhijsko umrežavanje i gradbu složenijih konstrukcija. Jednostavnije su konstrukcije, primjerice, koordinirana konstrukcija dviju imenica povezanih logičkim operatorom *i/ili*, a u složenije spadaju procesne konstrukcije u kojima se opojmljuje što se s imenicom kao objektom može raditi ili što opojmljeni objekt može raditi kao subjekt.

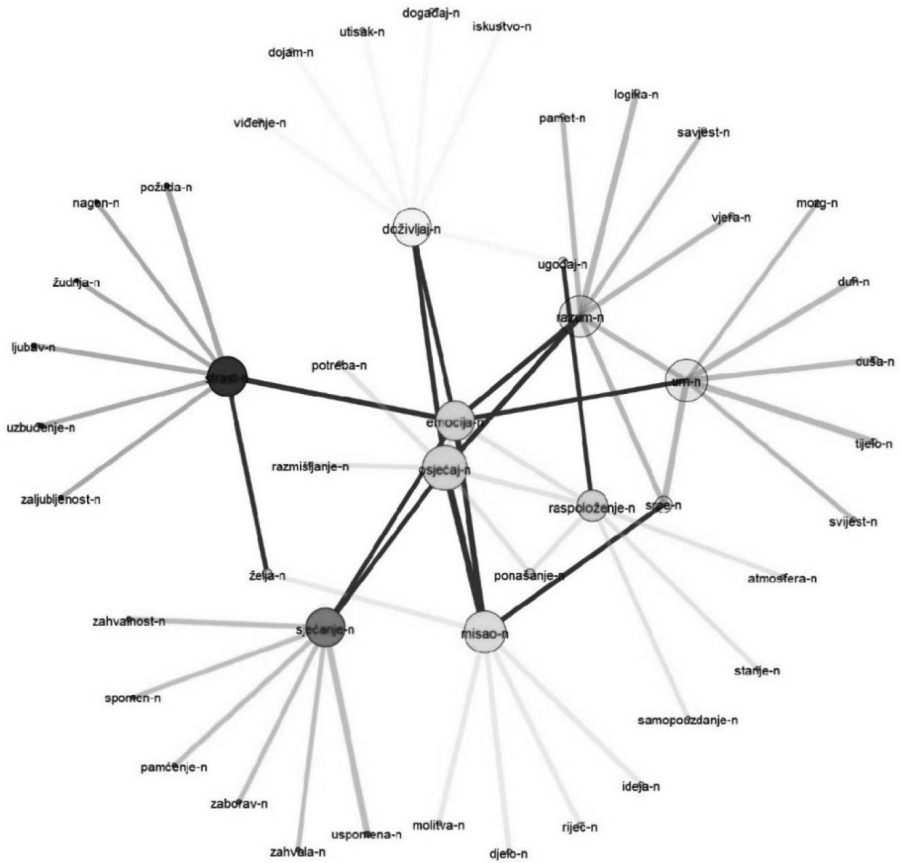
## Algoritamski pristup sintaktičko-semantičkim konstrukcijama

Sljedeći korak u uporabi podataka jest spajanje, filtriranje, mapiranje i umrežavanje podataka. Tim se postupcima otvaraju mogućnosti uporabe algoritama za pronalaženje svojstava mreža: algoritmi za identifikaciju zajednica čvorova (engl. *community detection algorithm*), mjere centralnosti (pageRank, Betweenness), izračun stupnja udaljenosti po određenoj sintaktičkoj dimenziji itd.

Primjerice, iskoristimo li semantičku značajku sintaktičke relacije *koordinacija* da kolocira ontološki srodne pojmove (*mačka i miš, ljubav i nježnost* itd.), možemo spajanjem podataka, filtriranjem kolokacija prema morfosintaktičkim obilježjima i mapiranjem njihovih sintaktičkih odnosa izraditi leksičku mrežu drugog stupnja, takozvanu friend-of-friend (FoF) mrežu, koja će za bilo



koju ishodišnu lemu prikazati prototipnu semantičku umreženost srodnih pojmovnih okvira i značenja. Na prikazu 3 predstavljena je takva sintaktičko-semantička mreža imenskog pojma emocija na temelju 8 kolokacija prvog stupnja i za svaku od tih 8 po još 8 kolokacija drugog stupnja.



**Prikaz 3.** Prototipna semantička umreženost imenskog leksema emocija u mreži 8 najistaknutijih kolokacija prvog stupnja (friend) u koordiniranoj konstrukciji + 8 najistaknutijih kolokacija drugog stupnja (friend-of-friend) svake kolokacije prvog stupnja

Mjera je za istaknutost veze LogDice score (logaritamski odnos čestotnosti supojavljenih čvorova i čestotnosti njihova supojavljivanja). U mreži se nalazi 48 lema raspoređenih u 6 identificiranih leksičkih zajednica:

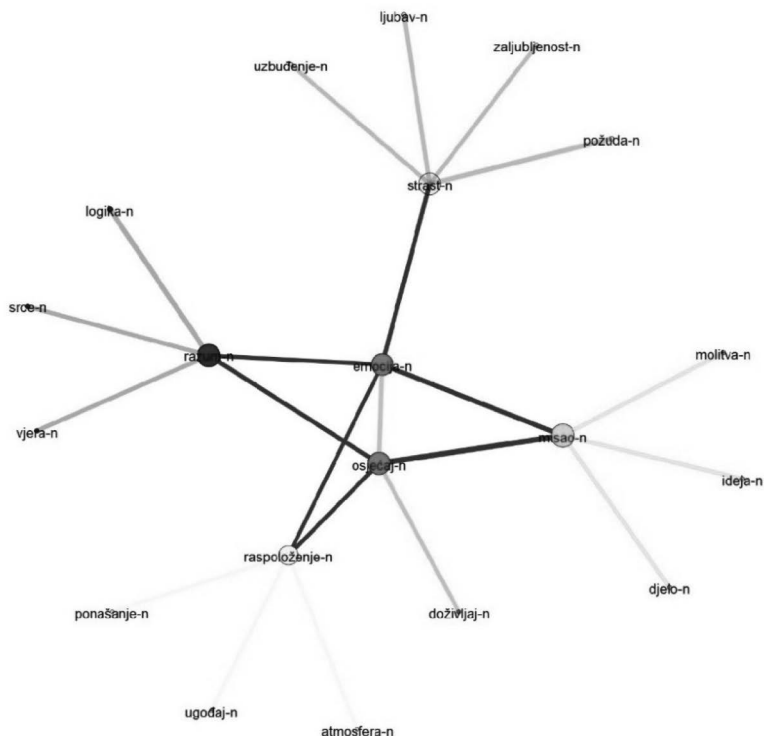
[1] razum-n, um-n, srce-n, logika-n, vjera-n, pamet-n, savjest-n, tijelo-n, duh-n, duša-n, svijest-n, mozg-n

[2] emocija-n, osjećaj-n, raspoloženje-n, razmišljanje-n, potreba-n, ponašanje-n, atmosfera-n, samopouzdanje-n, stanje-n

[3] strast-n, požuda-n, ljubav-n, uzbuđenje-n, zaljubljenost-n, nagon-n, žudnja-n

- [4] doživljaj-n, ugođaj-n, dojam-n, iskustvo-n, utisak-n, događaj-n, viđenje-n  
 [5] sjećanje-n, uspomena-n, zahvalnost-n, pamćenje-n, zaborav-n, spomen-n, zahvala-n  
 [6] misao-n, djelo-n, ideja-n, molitva-n, riječ-n, želja-n

Detaljnost mreže može se modificirati ako se u algoritmu promijeni broj kolokacija prvog i drugog stupnja. Primjerice, smanjenjem broja kolokacija na četiri najistaknutije veze ( $n=4$ ) dobiva se mreža kao na prikazu 4.



**Prikaz 4.** Leksička mreža izvornog pojma *emocija-n*

Na prikazu 4. leksička je mreža izvornog pojma *emocija-n* s 4 kolokacije drugog stupnja ( $n=4$ ): 20 elemenata i 5 klastera:

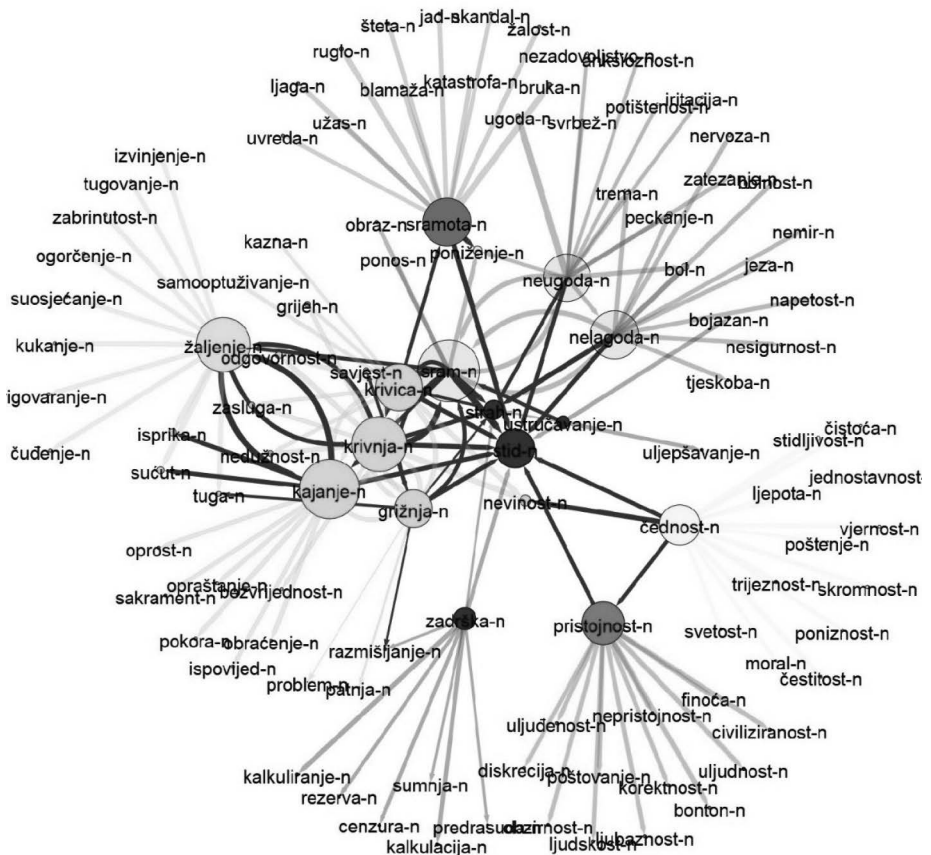
- [1] *strast-n*, *požuda-n*, *ljubav-n*, *uzbuđenje-n*, *zaljubljenost-n*
- [2] *misao-n*, *djelo-n*, *ideja-n*, *molitva-n*
- [3] *razum-n*, *logika-n*, *vjera-n*, *srce-n*
- [4] *raspoloženje-n*, *ponašanje-n*, *ugodaj-n*, *atmosfera-n*
- [5] *emocija-n*, *osjećaj-n*, *doživljaj-n*

Primjenom algoritma utvrđuje se da je leksem *emocija* najrodniji leksemima iz zajednice 5: *osjećaj*, *doživljaj*. Razlučivanjem ostalih zajednica dobiva se uvid u pojmovne okvire koji su istaknuto povezani s pojmom *emocija*. Tu

su uključena obilježja afektivnih stanja visoke pobuđenosti poput *strasti*, ali i dugotrajnija stanja i *raspoloženja*. Zajednice 2 i 3 govore u prilog nužnosti *razumskih* i *misaonih* komponenata za razumijevanje pojma emocija, bez obzira na to uklapa li se njihov odnos u kulturni model razdvajanja razumskih funkcija od emocionalnih sposobnosti, kao što je to navedeno u primjeru 1, ili se tumači modelom utjelovljene spoznaje koji zagovara nužnost emotivnih značajki pri procesiranju viših kognitivnih komponenti.

Na tom se primjeru očituje potencijal za dinamičnu pojmovnu, socijalnu i kulturološku analizu navedenog mrežnog pristupa pri konstruktivističkom razmatranju tvorbe značenja i značenjskih odnosa.

Razjasnimo još malo učinkovitost konstruktivističkog pristupa u razmatranju znanja o leksičkim kategorijama na primjeru dvaju srodnih pojmova: *sram* i *stid*. Bilo koji govornik hrvatskoga znat će da su te dvije riječi blisko značajne. Ali zna li prosječan govornik sadržajnu razliku? Zna li kad je primjerenije upotrijebiti jednu ili drugu riječ u kontekstu izražavanja vlastitog afektivnog stanja ili prepoznavanja psihološkog sadržaja izrečene emotivne kategorije?



**Prikaz 5. Leksička mreža izvornog pojma *stid*-n**

Na prikazu 5. leksička je mreža izvornog pojma *stid*-n s 10 kolokacija drugog stupnja s usmjerenim vezama 107 lema i 7 klastera:

[1] sram-n, nelagoda-n, neugoda-n, poniženje-n, obraz-n, ponos-n, savjest-n, bol-n, nervoza-n, bolnost-n, trema-n, jeza-n, napetost-n, peckanje-n, tjeskoba-n, nesigurnost-n, nemir-n, ugoda-n, zatezanje-n, iritacija-n, svrbež-n, potištenost-n, anksioznost-n, nezadovoljstvo-n

[2] krivnja-n, krivica-n, kajanje-n, grižnja-n, nevinost-n, odgovornost-n, grijeh-n, nedužnost-n, zasluga-n, bezvrijednost-n, pokora-n, obraćenje-n, opraštanje-n, oprost-n, sakrament-n, ispovijed-n, samooptuživanje-n, kazna-n, patnja-n, problem-n

[3] *stid*-n, strah-n, razmišljanje-n, zadržka-n, kalkuliranje-n, kalkulacija-n, cenzura-n, rezerva-n, predrasuda-n, sumnja-n, ustručavanje-n, bojazan-n, uljepšavanje-n

[4] čednost-n, stidljivost-n, trijeznost-n, skromnost-n, čestitost-n, vjernost-n, čistoća-n, svetost-n, poniznost-n, jednostavnost-n, moral-n, poštenje-n, ljepota-n

[5] pristojnost-n, bonton-n, uljudnost-n, civiliziranost-n, uljuđenost-n, ljubaznost-n, korektnost-n, obzirnost-n, finoća-n, nepristojnost-n, ljudskost-n, diskrecija-n, poštovanje-n

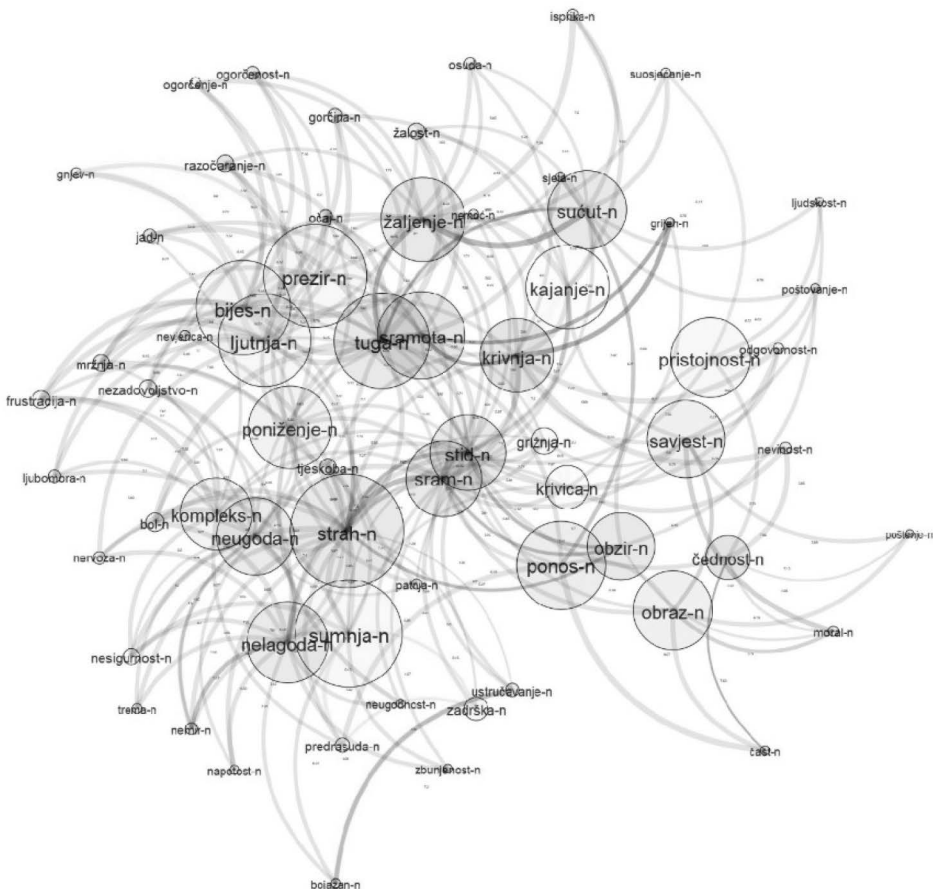
[6] žaljenje-n, sućut-n, isprika-n, tuga-n, zabrinutost-n, čuđenje-n, kukanje-n, prigovaranje-n, suosjećanje-n, izvinjenje-n, ogorčenje-n, tugovanje-n

[7] sramota-n, bruka-n, ruglo-n, blamaža-n, ljaga-n, žalost-n, katastrofa-n, užas-n, jad-n, šteta-n, uvreda-n, skandal-n.

Prema analizi mreže riječi *stid* s 10 najistaknutijih kolokacija drugog stupnja u gramatičkoj relaciji koordinacije uočljivo je da se *stid* značajnije supojavljuje s pojmovima *zadržka*, *ustručavanje* i *strah*, dok je *sram* umrežen s pojmovima *nelagoda*, *neugoda* i *poniženje*. Možda još važnije, *stid* se supostavlja antonimnim pojmovima *čednost* i *pristojnost*, dok je *sram* povezan s pojmovima *žaljenje*, *obraz* i *ponos*. Oba su pojma značajno povezana s *krivnjom*, *krivicom* i *kajanjem*. To bi pojednostavljeno moglo značiti da kad netko osjeća *stid*, ističe nelagodu zbog kršenja propisane kulturne norme *pristojnosti*, dok osjećaj *srama* istaknutije profilira negativnu hedoničku valenciju neugode i nezadovoljstva povezanu sa subjektivnim osjećajem prekoračenja osobnih i društvenih normi.

Gore prikazane mreže uspostavljene su s manjim brojem kolokacija prvog stupnja i s istim brojem kolokacija drugog stupnja (4, odnosno 10) zbog jasnoće prikaza. Međutim, moguće je konstruirati leksičku mrežu i s većim brojem kolokacija prvog i drugog stupnja, a zatim tehnikama obrezivanja (engl. *pruning*) slabije povezanih čvorova, koji su očito semantički udaljeniji od ishodišnog leksičkog pojma, dobiti potpuniju sliku semantičkih veza na temelju uključivanja većeg broja kolokacija prvog stupnja i njihove strukture umreženja putem kolokacija drugog stupnja. Aplikacija ConGraCNet omogućuje takav pristup semantičke analize kroz algoritme selekcije čvorova

na temelju određene graf-mjere centralnosti čvorova (*degree*, *betweenness* itd.) i prikaza subgrafa iz kojeg su izuzeti filtrirani čvorovi. Primjerice, na leksičkoj mreži pojma *stid* konstruiranoj iz 30 kolokacija prvog i drugog stupnja, a zatim obrezanoj na temelju parametra broja veza čvorova  $degree > 2$  (prikaz 6) jasnije se uočava razlika u strukturi pojmova *stid* i *sram*. Naime, *stid* je povezaniji s gornjim desnim dijelom semantičke mreže gdje su reprezentirani značenjski okviri savjesti, pristojnosti, žaljenja i krivnje, dok je *sram* istaknutije usmjeren na značenjske okvire straha, nelagode, neugodnosti, poniženja, frustracije, ponosa.



**Prikaz 6.** Leksička mreža izvornog pojma *stid-n* sa 65 čvorova i 237 veza

Na prikazu 6. leksička je mreža izvornog pojma *stid-n* sa 65 čvorova i 237 veza konstruirana na temelju mreže 30 kolokacija prvog i drugog stupnja (inicijalno 355 čvorova, 575 veza) obrezane na temelju parametra broja veza čvora s drugim čvorovima ( $degree > 2$ ). Rezultirajući subgraf podijeljen je u 41 klaster na temelju algoritma Leiden (CPM rezolucija = 0.81): [0] *stid-n*, *sram-n*, *strah-n*, *obzir-n* [1] *prezir-n*, *bijes-n*, *ljutnja-n*, *gnjev-n* [2] *sramota-n*, *krivnja-n*,

grijež-n [3] kajanje-n, krivica-n, grižnja-n [4] nelagoda-n, neugoda-n, nemir-n [5] zadržka-n, sumnja-n, predrasuda-n [6] obraz-n, savjest-n, moral-n [7] žaljenje-n, sućut-n, tuga-n [8] poniženje-n, neugodnost-n [9] čednost-n, čast-n [10] pristojnost-n, suosjećanje-n [11] ustručavanje-n, bojazan-n [12] kompleks-n, frustracija-n [13] ponos-n, sjeta-n [14] patnja-n [15] bol-n [16] nevjerica-n [17] nesigurnost-n [18] ljubomora-n [19] ogorčenje-n [20] odgovornost-n [21] zbu-njenost-n [22] nemoć-n [23] očaj-n [24] žalost-n [25] jad-n [26] nevinost-n [27] tjeskoba-n [28] osuda-n [29] isprika-n [30] gorčina-n [31] nervoza-n [32] trema-n [33] napetost-n [34] nezadovoljstvo-n [35] poštenje-n [36] ogorčenost-n [37] ra-zočaranje-n [38] ljudskost-n [39] poštovanje-n [40] mržnja-n.

Na Hrvatskom jezičnom portalu pronalazimo sljedeće definicije za navedene lekseme:

### STID

1. *psih.* osjećaj neugode praćen rumenilom, nemirom, zbu-njenošću kao refleks na povredu običaja i pristojnosti ili na ugroženost dostojanstva; sram
2. sramota

### SRAM

1. *psih.* a. mučan ili neugodan osjećaj prouzročen vlastitom pogreškom, povrijeđenom čašću, osjećajem manje vrijednosti; stid b. emocija koja je posljedica neuspjeha da se živi prema očekivanjima same osobe; krivnja
2. sramežljivost, zbu-njenost
3. ono što sramoti, sramota.

Naša kratka kvalitativno-kuantitativna mrežna analiza prilično se podudara s leksikografskim definicijama jezikoslovnih stručnjaka zabilježenima u rječniku. To je, posredno, i cilj digitalne aplikacije ConGraCNet: učiniti transparentnijim procese jezikoslovnih analiza i približiti ih što većem broju istraživača. U rječničkoj definiciji spominju se procesi koji prate dotične osjećaje: crveniti itd... Postoji li način da se to iščita iz korpusnog konstrukcijskog pristupa?

Aplikacija ConGraCNet omogućuje stvaranje mreža drugog stupnja za bilo koju od gore navedenih gramatičkih konstrukcija koje se odabiru u sučelju. Pravilnim odabirom konstrukcije koja mapira semantičko pitanje: koji su kauzalni procesi vezani uz entitet *x*? moguće je dobiti mrežu prototipnih kolokacija koje opisuju ponašanje pojma u zadanoj dimenziji. U hrvatskome to bi mogla biti prijedložna konstrukcija [proces\_*x* od entitet\_*x*]. Izgled sučelja vidljiv je u prikazu 7 s rezultatima umreženja leksema *stid* u prijedložnoj konstrukciji *od*.

**Construction parameters**

Choose grammatical relation

od-s\_X

Filter the pos value of the friend lexeme

v

**Visualization parameters**

Vertex size by

weighted\_degree

Vertex size

1.00

0.00 2.00

Label size by

weighted\_degree

Label size

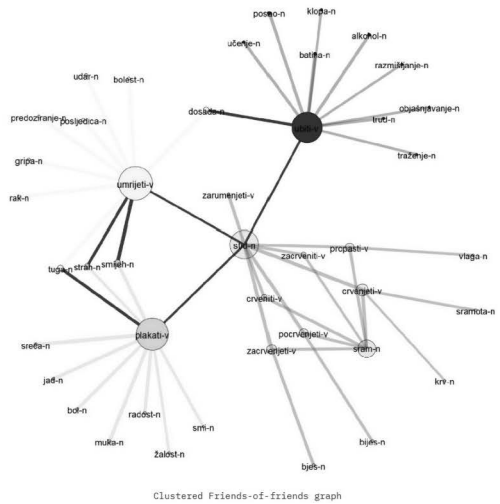
24.79

10.00 40.00

Layout type

kk

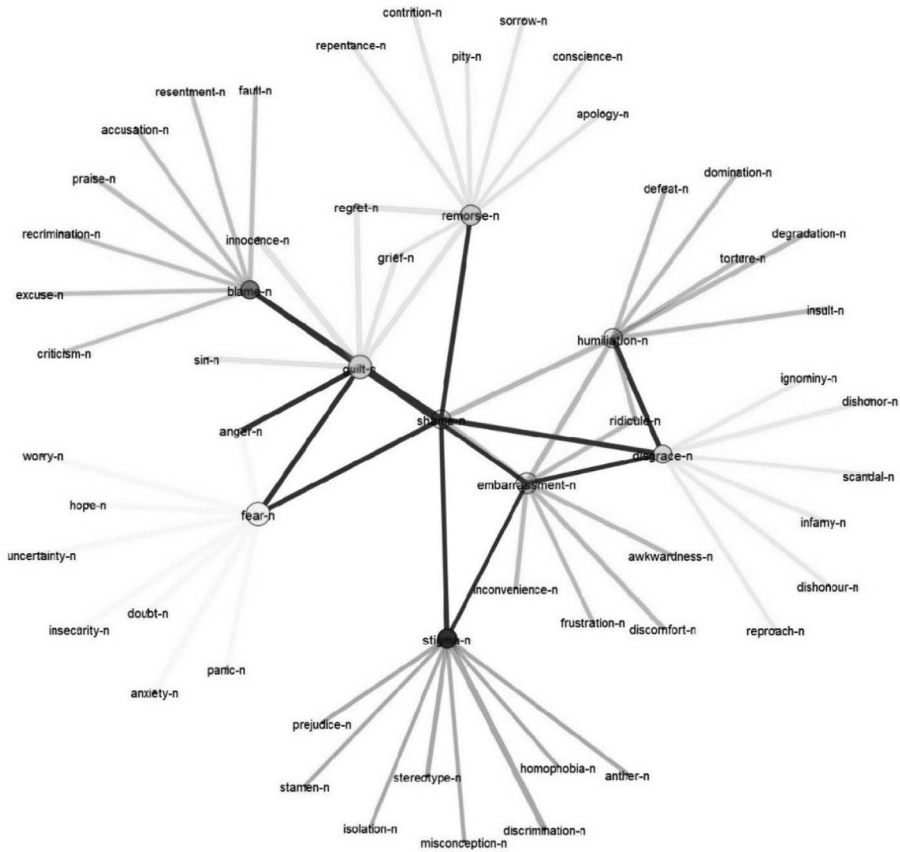
Cluster visualization



**Prikaz 7.** Sučelje aplikacije ConGraCNet i leksička mreža izvornog pojma *stid-n* s do 10 kolokacija drugog stupnja u gramatičkoj relaciji *od-s\_X*

Za sintaktičko-semantičku mrežu *stida* u prijedložnoj konstrukciji moguće je reći da prikazuje metonimijsko proširenje prostornog značenja [(udaljiti|kretati se) *od* entitet\_x] prema kauzalnom značenju [proces\_X\_POSLJEDICA *od* entitet\_x\_UZROK] prijedloga *od*, otkrivajući što je tipični kauzalni proces doživljaja *stida* opojmljenog kao entiteta. Istaknuti procesi koji se pojavljuju u prvom krugu kolokacija vezani su uz *crvenilo* te neugodne procese *plakati* ili čak *umrijeti*, što je podudarno s leksičkim opisom u rječnicima. Zanimljivo je primijetiti da se u krugu leksema drugog reda (friend-of-friend) nalaze druge prototipne kolokacije s potencijalom za kauzalnu analogiju. Primjerice, može se reći da neko *plače od stida* kao što se može reći da *plače od straha*.

Za kraj navedimo da je konstrukcijski pristup moguće primijeniti i na druge jezike, odnosno korpus. S obzirom na prikupljanje podataka sa servisa SketchEngine to znači da su nam na raspolaganju uvidi o leksičkom ponašanju u stotinama korpusa brojnih europskih i svjetskih jezika. Primjerice, istražimo li sumjerljivost hrvatskih pojmova *sram* ili *stid* riječi *shame* u engleskome korpusu enTenTen13, usporedbom sumjerljivih konstrukcija dobivamo leksičku mrežu koja je prikazana na ilustraciji 8.



**Prikaz 8.** Leksička mreža izvornog pojma *shame-n*

Na prikazu 8 leksička je mreža izvornog pojma *shame-n* s do 10 kolokacija drugog stupnja u gramatičkoj relaciji „%w“ and/or... s 58 leksema i 6 leksičkih zajednica:

[1] *shame-n*, *embarrassment-n*, *humiliation-n*, *discomfort-n*, *awkwardness-n*, *inconvenience-n*, *ridicule-n*, *frustration-n*, *degradation-n*, *torture-n*, *insult-n*, *domination-n*, *defeat-n*

[2] *guilt-n*, *remorse-n*, *innocence-n*, *regret-n*, *sin-n*, *grief-n*, *contrition-n*, *repentance-n*, *pity-n*, *apology-n*, *sorrow-n*, *conscience-n*

[3] *stigma-n*, *discrimination-n*, *stereotype-n*, *prejudice-n*, *anther-n*, *homophobia-n*, *stamen-n*, *misconception-n*, *isolation-n*

[4] *fear-n*, *anger-n*, *anxiety-n*, *doubt-n*, *worry-n*, *hope-n*, *uncertainty-n*, *insecurity-n*, *panic-n*

[5] *blame-n*, *praise-n*, *fault-n*, *accusation-n*, *resentment-n*, *recrimination-n*, *excuse-n*, *criticism-n*

[6] *disgrace-n*, *dishonor-n*, *infamy-n*, *ignominy-n*, *dishonour-n*, *reproach-n*, *scandal-n*



Razlučivanjem leksičkih zajednica istaknuta je razlika između imenice *shame*, koja je vezana uz zajednicu *embarrassment* i *humiliation*, i imenice *guilt*, koja je vezana uz pojmove *remorse*, *regret* i *sin*. Na temelju usporednih analiza dvaju različitih korpusa mogli bismo reći da uporabna semantička domena leksema *shame* više odgovara semantičkoj domeni hrvatskog leksema *sram* nego domeni leksema *stid*.

## Zaključak

U ovom su radu prikazane mogućnosti korpusnih istraživanja s ciljem empirijske obrade izražavanja leksičkih obrazaca. Korpusni pristup podrazumijeva digitalizaciju resursa, pohranu korpusa tekstova i metapodataka, morfosintaktičko obilježavanje i sintaktičko-semantičku analizu teksta. Korpusno istraživanje izražavanja emocija uključuje identifikaciju izražavanja emocija i afektivnih stanja na temelju sintaktičko-semantičkih obrazaca zabilježenih u korpusima suvremenog hrvatskog jezika, stvaranje baze podataka o jezičnim i kulturološkim obrascima izražavanja afektivnih stanja te obogaćivanje podataka o konvencionaliziranim jezičnim strukturama podacima iz emocionalnih ontologija i rječničkih baza.

Korpusne metode konstrukcijske gramatike omogućuju: empirijski uvid u strukturiranje značenjske matrice, izbor sintaktičko-semantičkih konstrukcija bitnih za opojmljivanje prototipnih odnosa entiteta, njihovih svojstava i procesa, računalne metode analize pojmovne udaljenosti/srodnosti na temelju kognitivne hijerarhije kategorizacije, opis polisemičkih okvira značenjskih matrica, pragmalingvističke i sociolingvističke usporedbe značenjskih matrica unutar korpusa/između korpusa, podatke za kvalitativnu procjenu specifičnih ostvarenja pojmovnog uokvirivanja koja se može uklopiti u diskurznu kulturološku analizu.

Aplikacija ConGraCNet (<http://emocnet.uniri.hr/congracnet/>) omogućuje istraživanje modela izražavanja emocija i afektivnih stanja s ciljem identifikacije emotivnih kategorija, afektivnih stanja i semantičke obilježenosti u tekstovima. Ideja je aplikacije ponuditi istraživačke alate za identifikaciju izražavanja emocija i afektivnih stanja na temelju sintaktičko-semantičkih obrazaca zabilježenih u postojećim i novostvorenim korpusima suvremenog hrvatskog jezika te na temelju usporedbe s drugim jezicima.

## Literatura

- Ahmad, Khurshid (ur.). 2011. *Affective Computing and Sentiment Analysis: Emotion, Metaphor and Terminology*. Dordrecht: Springer Science & Business Media.
- Al-Saaqa, Sammar; Heba Abdel-Nabi; Arafat Awajan. 2018. A survey of textual emotion detection. *CSIT 2018: 8th International Conference on Computer Science and Information Technology*. IEEE: 136–142.
- Barrett, Lisa Feldman. 2017a. *How Emotions Are Made: The Secret Life of the Brain*. Boston: Houghton Mifflin Harcourt.
- Barrett, Lisa Feldman. 2017b. The theory of constructed emotion: an active inference account of interoception and categorization. *Social Cognitive and Affective Neuroscience* 12 (1): 1–23.
- Bartneck, Christoph; Michael J. Lyons; Martin Saerbeck. 2017. The relationship between emotion models and artificial intelligence. Osaka, Japan: *SAB2008 Workshop on The Role of Emotion in Adaptive Behavior and Cognitive Robotics*, 2008. CoRR. abs/1706.09554.
- Brdar, Mario. 2015. Metonymic chains and synonymy. *Fluminensia: časopis za filološka istraživanja* 27 (2): 83–101.
- Brdar, Mario; Rita Brdar-Szabó; Benedikt Perak. 2019. Metaphor repositories and cross-linguistic comparison: ontological eggs and chickens. U: *Metaphor and Metonymy in the Digital Age: Theory and Methods for Building Repositories of Figurative Language* (ur. Marianna Bolognesi, Kristina Despot i Mario Brdar): 225–252. Amsterdam: John Benjamins Publishing Company.
- Brdar, Mario; Rita Brdar-Szabó; Benedikt Perak. 2020. Separating (non-)figurative weeds from wheat. U: *Figurative Meaning Construction in Thought and Language* (ur. Annalisa Baicchi): 46–70. Amsterdam: John Benjamins Publishing Company.
- Brozović Rončević, Dunja; Damir Ćavar; Małgorzata Ćavar; Tomislav Stojanov; Kristina Štrkalj Despot; Nikola Ljubešić; Tomaž Erjavec. 2018. *Croatian language corpus Riznica 0.1*. Zagreb: Institut za hrvatski jezik i jezikoslovlje.
- Cambria, Erik. 2016. Affective computing and sentiment analysis. *IEEE Intelligent Systems* 31 (2): 102–107.
- Carbaugh, Donal. 2007. Cultural discourse analysis: Communication practices and intercultural encounters. *Journal of Intercultural Communication Research* 36 (3): 167–182.
- Crawford, William J.; Eniko Csomay. 2016. *Doing Corpus Linguistics*. New York – Abingdon: Routledge.
- Despot, Kristina; Mirjana Tonković; Mario Brdar; Mario Essert; Benedikt Perak; Ana Ostroški Anić; Bruno Nahod; Ivan Pandžić. 2019. MetaNet. HR: Croatian Metaphor Repository. U: *Metaphor and Metonymy in the Digital*

- Age (ur. Marianna Bolognesi, Mario Brdar i Kristina Despot): 123–146. Amsterdam: John Benjamins Publishing Company.
- Eckart de Castilho, Richard; Éva Mújdricza-Maydt; Seid Muhie Yimam; Silvana Hartmann; Iryna Gurevych; Anette Frank; Chris Biemann. 2016. A web-based tool for the integrated annotation of semantic and syntactic structures. U: *Proceedings of the Workshop on Language Technology Resources and Tools for Digital Humanities (LT4DH)*: 76–84.
- Edwards, Derek. 1999. Emotion discourse. *Culture & Psychology* 5 (3): 271–291.
- Foolen, Ad; Ulrike M. Lüdtke; Timothy P. Racine; Jordan Zlatev (ur.). 2012. *Moving Ourselves, Moving Others: Motion and Emotion in Intersubjectivity, Consciousness and Language*. Amsterdam: John Benjamins Publishing Company.
- Gee, James Paul. 2004. *An Introduction to Discourse Analysis: Theory and Method*. New York: Routledge.
- Gmuer, Angelina; Jeannette Nuessli Guth; Michael Siegrist; Maren Runte. 2015. From emotion to language: Application of a systematic, linguistic-based approach to design a food-associated emotion lexicon. *Food Quality and Preference* 40: 77–86.
- Goldie, Peter (ur.). 2017. *Understanding Emotions: Mind and Morals*. London: Routledge.
- González, Ana Marta. 2016. *The Emotions and Cultural Analysis*. Abingdon – New York: Routledge.
- Grgurić, Diana; Benedikt Perak. 2020. Tekstovi popularnih pjesama u konstrukciji nacionalnog sjećanja. *Fluminensia: časopis za filološka istraživanja* 31 (2): 297–322.
- Guia, José; Valéria Gonçalves Soares; Jorge Bernardino. 2017. Graph Databases: Neo4j Analysis. *Proceedings of the 19th International Conference on Enterprise Information Systems (ICEIS 2017)* 1: 351–356.
- Hart, Christopher. 2010. *Critical Discourse Analysis and Cognitive Science: New Perspectives on Immigration Discourse*. Basingstoke: Palgrave Macmillan.
- Hoggett, Paul. 2015. *Politics, Identity and Emotion*. London – New York: Routledge.
- Hudeček, Lana; Milica Mihaljević. 2017. Hrvatski mrežni rječnik – Mrežnik. *Hrvatski jezik: znanstveno-popularni časopis za kulturu hrvatskoga jezika* 4 (4): 1–7.
- Jacobs, Arthur M. 2016. Affective and aesthetic processes in literary reading. U: *Cognitive Literary Science: Dialogues between Literature and Cognition* (ur. Emily T. Troscianko i Michael Burke): 303–325. Oxford – New York: Oxford University Press.
- Karl, Dario; Božo Bekavac; Ida Raffaelli. 2018. A construction grammar approach in the NooJ Framework: Semantic analysis of lexemes describing emotions in Croatian. U: *Formalizing Natural Languages with NooJ 2018 and*

- Its Natural Language Processing Applications* (ur. Ignazio Mauro Mirto, Mario Monteleone i Max Silberstein): 123–114. Cham: Springer.
- Kong, Shuyu. 2014. *Popular Media, Social Emotion and Public Discourse in Contemporary China*. New York – Abingdon: Routledge.
- Kotz, Sonja A.; Silke Paulmann. 2011. Emotion, language, and the brain. *Language and Linguistics Compass* 5 (3): 108–125.
- Kövecses, Zoltán. 2012. *Emotion Concepts*. New York: Springer Science & Business Media.
- Lindquist, Kristen A.; Maria Gendron. 2013. What's in a word? Language constructs emotion perception. *Emotion Review* 5 (1): 66–71.
- Lindquist, Kristen A.; Ajay Satpute; Maria Gendron. 2015. Does language do more than communicate emotion? *Current Directions in Psychological Science* 24 (2): 99–108.
- Lindquist, Kristen A.; Tor D. Wager; Hedy Kober; Eliza Bliss-Moreau; Lisa Feldman Barrett. 2012. The brain basis of emotion: A meta-analytic review. *Behavioral and Brain Sciences* 35 (3): 121–143.
- Lüdtke, Ulrike M. (ur.). 2015. *Emotion in Language: Theory – Research – Application*. Amsterdam: John Benjamins Publishing Company.
- Ljubešić, Nikola; Tomaž Erjavec. 2011. hrWaC and slWaC: Compiling web corpora for Croatian and Slovene. U: *Text, Speech and Dialogue, Lecture Notes in Computer Science* (ur. Ivan Habernal i Vaclav Matousek): 395–402. Berlin – Heidelberg: Springer.
- Martinez-Miranda, Juan; Arantza Aldea. 2005. Emotions in human and artificial intelligence. *Computers in Human Behavior* 21 (2): 323–341.
- Mautner, Gerlinde. 2016. Checks and balances: How corpus linguistics can contribute to CDA. U: *Methods of Critical Discourse Analysis* (ur. Ruth Wodak i Michael Meyer): 154–179. London: Sage Publications.
- Mohammad, Saif M. 2016. Sentiment analysis: Detecting valence, emotions, and other affectual states from text. U: *Emotion Measurement* (ur. Herbert Meiselman): 201–237. Sawston: Woodhead Publishing.
- Parizoska, Jelena; Ivana Filipović Petrović. 2019. The conceptual organization of idioms dictionaries in e-lexicography. *Filologija* 73: 27–45.
- Pavlaković, Vjieran; Benedikt Perak. 2017. How does this monument make you feel? Measuring emotional responses to war memorials in Croatia. U: *The Twentieth Century in European Memory* (ur. Tea Sindbæk Andersen i Barbara Törnquist-Plewa): 268–304. Nijhoff: Brill.
- Perak, Benedikt; Sanja Puljar D'Alessio. 2013. Kultura kao emergentno svojstvo otjelovljene spoznaje. U: *Avanture kulture: kulturalni studiji u lokalnom kontekstu* (ur. Nenad Fanuko i Sanja Puljar D'Alessio): 77–108. Zagreb: Jesenski i Turk.

- Perak, Benedikt; Katarina Damčević; Jana Milošević. 2018. O sranju i drugim neprimjerenim stvarima: kognitivno-lingvistička analiza psovki u hrvatskome. U: *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* (ur. Diana Stolac i Anastazija Vlastelić): 245–270. Zagreb: Srednja Europa.
- Perak, Benedikt. 2014a. Opojmljivanje leksema strah u hrvatskome: sintaktičko-semantička analiza (doktorska disertacija). Zagreb: Filozofski fakultet Sveučilišta u Zagrebu.
- Perak, Benedikt. 2014b. Razlike metonimijskog i metaforičkog opojmljavanja kategorija straha i ljubomore. U: *Riječki filološki dani. Zbornik radova s međunarodnoga znanstvenog skupa Riječki filološki dani, održanoga u Rijeci od 22. do 24. studenoga 2012.* (ur. Diana Stolac): 595–612. Rijeka: Sveučilište u Rijeci.
- Perak, Benedikt. 2015. Tematske i agentivne konstrukcije „straha“ izvedene iz opojmljivanja unutar mjesnoga prostornog odnosa kodiranog prijedlogom „u“. *Poznańskie Studia Slawistyczne* 9: 333–352.
- Perak, Benedikt. 2017a. 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* (ur. Diana Stolac i Anastazija Vlastelić): 297–317. Zagreb: Srednja Europa.
- Perak, Benedikt. 2017b. Conceptualisation of the emotion terms: Structuring, categorization, metonymic and metaphoric processes within multi-layered graph representation of the syntactic and semantic analysis of corpus data. U: *Cognitive Modelling in Language and Discourse across Cultures* (ur. Annalisa Baicchi i Erica Pinelli): 299–319. Newcastle upon Tyne: Cambridge Scholars Publishing.
- Perak, Benedikt. 2018. The role of metonymy in the constructionist approach to the conceptualization of emotions. U: *Conceptual Metonymy: Methodological, Theoretical, and Descriptive Issues* (ur. Antonio Barcelona i Olga Blanco): 256–258. Amsterdam: John Benjamins Publishing Company.
- Perak, Benedikt. 2019a. The role of metonymy and metaphor in the conceptualization of the nation. An emergent ontological analysis of syntactic-semantic constructions. U: *Metaphor, Nation and Discourse* (ur. Ljiljana Šarić i Mateuzs-Milan Stanojević): 227–258. Amsterdam: John Benjamins Publishing Company.
- Perak, Benedikt. 2019b. Ontological and constructional approach to the discourse analysis of the commemorative speeches. U: *Political Rituals and Cultural Memory of the Twentieth-Century Traumas in Croatia* (ur. Vjeron Pavlaković i Davor Pauković): 63–100. London: Routledge.

- Perak, Benedikt. 2020. Developing the ontological model for research and representation of commemoration speeches in Croatia using a graph property database. U: *Digital Humanities: Empowering Visibility of Croatian Cultural Heritage: Selected Papers from the 2nd International Symposium Held in November 2017 in Zadar, Croatia* (ur. Marijana Tomić, Mirna Willer i Nives Tomašević): 88–110. Newcastle upon Tyne: Cambridge Scholars Publishing.
- Samardžić, Tanja; Nikola Ljubešić; Maja Miličević. 2015. Regional Linguistic Data Initiative (ReLDI). *Proceedings of the 5th Workshop on Balto-Slavic Natural Language Processing*: 40–42.
- Sander, David; Didier Grandjean; Klaus R. Scherer. 2018. Brain networks, emotion components, and appraised relevance. *Emotion Review* 10 (3): 238–241.
- Scherer, Klaus R. 2009. Emotions are emergent processes: they require a dynamic computational architecture. *Philosophical Transactions of the Royal Society B* 364: 3459–3474.
- Scherer, Klaus R.; Johnny R. J. Fontaine. 2018. The semantic structure of emotion words across languages is consistent with componential appraisal models of emotion. *Cognition and Emotion* 33 (4): 673–682.
- Schuller, Björn; Anton Batliner. 2013. *Computational Paralinguistics: Emotion, Affect and Personality in Speech and Language Processing*. Chichester: Wiley-Blackwell.
- Schuller, Dagmar; Björn W. Schuller. 2018. The age of artificial emotional intelligence. *Computer* 51 (9): 38–46.
- Schwarz-Friesel, Monika. 2015. Language and emotion. The cognitive linguistic perspective. U: *Emotion in Language. Theory – Research – Application* (ur. Ulrike M. Lüdtke): 157–174. Amsterdam: John Benjamins Publishing Company.
- Sharifian, Farzad. 2017a. *Cultural Linguistics: Cultural Conceptualisations and Language*. Amsterdam: John Benjamins Publishing Company.
- Sharifian, Farzad (ur.). 2017b. *Advances in Cultural Linguistics*. New York – London – Singapore: Springer.
- Shuman, Vera; Elizabeth Clark-Polner; Ben Meuleman; David Sander; Klaus R. Scherer. 2017. Emotion perception from a componential perspective. *Cognition and Emotion* 31 (1): 47–56.
- Stanojević, Mateusz-Milan. 2017. *Hladne glave i vruća srca o tijelu i emocijama u hrvatskom*. U: *Tijelo u hrvatskome jeziku, književnosti i kulturi. Zbornik radova 45. seminara Zagrebačke slavističke škole* (ur. Ivana Brković i Tatjana Pišković): 175–198. Zagreb: FF press i Zagrebačka slavistička škola.
- Swain, Merrill. 2013. The inseparability of cognition and emotion in second language learning. *Language Teaching* 46 (2): 195–207.

- Tadić, Marko. 1996. Računalna obradba hrvatskoga i nacionalni korpus. *Suvremena lingvistika* 41/42 (1-2): 603–611.
- Tillman, Richard; Max Louwerse. 2018. Estimating emotions through language statistics and embodied cognition. *Journal of Psycholinguistic Research* 47 (1): 159–167.
- Tissari, Heli. 2017. Current emotion research in English linguistics: words for emotions in the history of English. *Emotion Review* 9 (1): 86–94.
- Weisser, Martin. 2016. *Practical Corpus Linguistics: An Introduction to Corpus-Based Language Analysis*. Chichester: Wiley Blackwell.
- White, Geoffrey M. 2010. Moral discourse and the rhetoric of emotion. U: *Blackwell Anthologies in Social and Cultural Anthropology. Psychological Anthropology: A Reader on Self in Culture*: 68–82. Chichester: Wiley Blackwell.
- Yonck, Richard. 2017. *Heart of the Machine: Our Future in a World of Artificial Emotional Intelligence*. New York: Arcade Publishing.
- Zheng, Quan; David Skillicorn. 2017. *Social Networks with Rich Edge Semantics*. Boca Raton: CRC Press.

## Emotions in Text Corpora: The Application of Construction Grammar and Graph methods for analyzing the Semantic Structure of Emotional Categories

### Summary

The paper provides an insight into the research on the conceptualization and expression of emotions, based on interdisciplinary approach of the cognitive sciences, corpus and computational linguistics. The method of computational corpus data analysis and construction grammar conceptual network (ConGraCNet, <http://emocnet.uniri.hr/congracnet/>) approach is described, including the application of a graph algorithms to identify the lexical polysemy of emotional categories and to quantitative-qualitative description of semantic domains in different corpora and languages. The method is exemplified on the analysis of the Croatian lexemes *sram*, *stid* in hrWac corpus and their translational lexical equivalent *shame* in English enTenTen13 corpora.

**Ključne riječi:** konstrukcijska gramatika, računalna lingvistika, graf-analiza, sintaktičko-semantički odnosi, emocije, *sram*, *stid*

**Keywords:** construction grammar, computational linguistics, graph analysis, syntactic-semantic relations, emotion, shame