

Kvantitativna strojna analiza razporeditve čustev na primeru *Visoške kronike*

DOI: 10.4312/Obdobja.44.59-66

Jaka Čibej, Filozofska fakulteta, Univerza v Ljubljani, Ljubljana; jaka.cibej@ff.uni-lj.si

V prispevku predstavljamo nekatere možnosti kvantitativne strojne analize razporeditve čustev na primeru zgodovinskega romana *Visoška kronika* Ivana Tavčarja. Besedilo romana razdelimo na odstavke in s pomočjo leksikona čustvenih asociacij SloEmoLex 1.0 vsakemu odstavku strojno pripišemo številsko razporeditev osmih čustev: jeza, strah, žalost, gnus, presenečenje, pričakovanje, zaupanje in sreča. Rezultate vizualiziramo in jih na kratko interpretiramo glede na vsebino romana ter nakažemo potencialne koristi tovrstne strojne obdelave in možnosti za prihodnje delo pri strojni obdelavi čustev v slovenski književnosti.

strojna analiza čustev, leksikon čustvenih asociacij, slovenska književnost, *Visoška kronika*, Ivan Tavčar

This article presents some options for a quantitative automatic analysis of the distribution of emotions using Ivan Tavčar's historical novel *Visoška kronika* (The Visoko Chronicle) as an example. The novel's text is divided into paragraphs, and the SloEmoLex 1.0 Lexicon of Word–Emotion Associations is used to assign a numerical distribution of eight emotions to each paragraph: anger, fear, sadness, disgust, surprise, anticipation, trust, or joy. The results are visualized and briefly interpreted in terms of the novel's content. In addition, several potential benefits of this automatic approach are presented, along with several possibilities for future work related to the automatic analysis of emotions in Slovenian literature.

automatic emotion analysis, word–emotion association lexicon, Slovenian literature, *Visoška kronika* (The Visoko Chronicle), Ivan Tavčar

1 Uvod

V slovenskem prostoru so tehnološko podprte metode analize besedil z razrastom korpusnega in računalniškega jezikoslovja v zadnjih desetletjih že dodobra uveljavljene. V tem času so bili zgrajeni številni korpusi, ki služijo kot gradivo za statistično podprte jezikoslovne analize različnih zvrsti besedil, npr. korpus pisne standardne slovenščine Gigafida (Krek idr. 2019), korpus akademske slovenščine KAS (Erjavec idr. 2017) in spremljevalni korpus Trendi (Kosem idr. 2021). Kljub temu napredku pa potencial, ki ga ponujajo jezikovnotehnološka orodja, v slovenščini še ni bil dovolj izkoriščen pri analizi književnih besedil,¹ čeprav se tovrstna praksa v tujini že nekaj časa uveljavlja –

1 Izjema je npr. stilometrična analiza pripovedne literature Janeza Ciglerja in Christopha Schmidta (Žejn 2020), pri kateri je bil uporabljen programski paket Stylo v okolju R. Na podoben način je bila opravljena tudi empirična analiza žanra kriminalke na primeru skandinavskih in slovenskih kriminalk (Koncilijska 2024).



celo kot ločena disciplina računalniške literarne teorije v okviru digitalne humanistike (pregled sorodnih del in možnosti pri strojni analizi čustev v književnosti sta opravila npr. Kim in Klinger (2022)).

Strojni in statistično podprti postopki lahko npr. pomagajo pri vizualizacijah in kvantitativni analizi določenih vidikov književnih besedil, izluščeni podatki pa so lahko dodatna empirična podlaga pri analizi strukture dramskega loka in dogajanja, povezanosti med liki, stilometriji ter razporeditvi čustev v besedilu. Mohammad (2011) npr. s pomočjo leksikona čustvenih asociacij z grafi vizualizira potek čustev skozi zgodbo v delih *Kakor vam drago* (William Shakespeare), *Hamlet* (William Shakespeare) in *Frankenstein* (Mary Shelley), kar med drugim razkrije čustveno bolj negativno ali pozitivno naravnane odseke besedila. Klinger, Kim in Padó (2020) raziščejo še naprednejše metode pripisovanja čustev s strojnim učenjem, ki razporeditve čustev ne analizirajo le globalno skozi celotno besedilo, temveč tudi v kontekstu posameznih dogodkov in likov v zgodbi, kar analizo čustvovanja prestavi z bolj robustnega nivoja celotnega dela na bolj granularno interpretacijo posameznih likov. Samothrakis in Fasli (2015) kot dodatno korist strojne analize čustvovanja v književnih delih izpostavita dejstvo, da kvantitativna analiza čustev v delih pomaga izboljšati točnost strojnega prepoznavanja različnih žanrov književnosti.

Precej besedil slovenske književnosti je že digitaliziranih. Nekatera dela so vključena tudi npr. v korpus pisne standardne slovenščine Gigafida in v korpus starejše slovenščine IMP (Erjavec 2014), največja količina književnih besedil v strojno berljivi (čeprav za strojno obdelavo ne povsem optimalno strukturirani) obliki pa je na voljo na portalu Wikivir,² ki se v okviru projekta Slovenska književnost na Wikiviru polni z deli v javni domeni iz Digitalne knjižnice Slovenije (Hladnik, Polajnar 2016). Tehničnih ovir za strojno obdelavo slovenskih književnih besedil torej ni. V prispevku zato predstavljamo prvi primer strojne analize razporeditve čustev v slovenskih književnih besedilih na primeru *Visoške kronike* Ivana Tavčarja, zgodovinskega romana iz prve četrtine 20. stoletja. Po kratkem opisu uporabljenih virov, metode in priprave podatkov (razdelek 2) predstavimo poglobitve izsledke strojne analize in vizualizacije, ki iz njih izhajajo (razdelek 3), v zaključku pa nakažemo nekaj možnosti za nadaljnje delo in izboljšave.

2 Metoda, viri in priprava podatkov

Metoda strojne analize razporeditve čustev, ki smo jo uporabili v raziskavi, je zelo robustna in temelji na leksikonu čustvenih asociacij SloEmoLex 1.0 (Brglez idr. 2024), ki je bil uporabljen npr. za analizo dehumanizacije prebežnikov v slovenskih medijih (Caporusso idr. 2024). Leksikon vsebuje slovenske lekseme ter njihove kvantificirane asociacije na osem primarnih čustev glede na klasifikacijo ameriškega psihologa Roberta Plutchika (gl. npr. Plutchik 1982): jeza, strah, žalost, gnus, presenečenje, pričakovanje, zaupanje in sreča. SloEmoLex 1.0 sicer ni bil izdelan posebej za slovenščino, temveč izhaja iz prevodov sorodnih leksikonov čustvenih asociacij za angleščino in

2 Slovenska leposlovna klasika na Wikiviru: https://sl.wikisource.org/wiki/Wikivir:Slovenska_leposlovna_klasika.

druge jezike, npr. LiLaH (Daelemans idr. 2020; Ljubešić idr. 2020) in NRC (Mohammad, Turney 2010). Čustvene asociacije so bile zbrane s pomočjo množenja, pri katerem so materni govorci angleščine ocenjevali, v kolikšni meri jih določena beseda asociira na določeno čustvo oz. ali jo pretežno zaznavajo kot negativno ali pozitivno (Mohammad, Turney 2013).

Vsak leksem v leksikonu je (med drugim) označen s čustvi, s katerimi je asociiran, pripisano pa ima tudi intenzivnost asociacije na posamezno čustvo (kot povprečje vsaj štirih mnenj). Poenostavljen prikaz nekaterih primerov kvantificiranih asociacij iz leksikona je predstavljen v Tabeli 1 – v zadnjem stolpcu so podane asociacije na čustva ter njihove intenzivnosti med 0 (asociacije ni) in 1 (asociacija je zelo močna). Leksem *katastrofa* denimo asociira na pet različnih čustev, najmočnejše na strah in presenečenje, najšibkeje pa na gnus.

Leksem	Zaznamovanost	Čustvene asociacije
<i>razočaran</i>	negativno	jeza (0,348) gnus (0,328) žalost (0,636)
<i>katastrofa</i>	negativno	jeza (0,721) gnus (0,508) strah (0,875) žalost (0,758) presenečenje (0,812)
<i>hvaležnost</i>	pozitivno	sreča (0,914)
<i>radost</i>	pozitivno	pričakovanje (0,633) sreča (0,765)

Tabela 1: Primeri kvantificiranih čustvenih asociacij iz leksikona SloEmoLex 1.0.

Leksemi v leksikonu so navedeni le v osnovnih oblikah, zato smo besedilo *Visoške kronike* najprej strojno lematizirali in oblikoskladenjsko označili z orodjem CLASSLA-Stanza (Ljubešić, Dobrovoljc 2019; različica 1.1.0), nato pa za vsako čustvo izračunali čustveno intenzivnost na nivoju posameznega odstavka – najprej preverimo, kateri leksemi se pojavijo v odstavku, izračunamo njihovo relativno pogostost znotraj odstavka ter jo pomnožimo s količnikom čustvene zaznamovanosti iz leksikona (tj. ali beseda asociira s čustvom – 1 – ali ne – 0) in s količnikom čustvene intenzivnosti (v kolikšni meri beseda asociira na to čustvo na lestvici med 0 in 1). Čustvena intenzivnost (I) odstavka P s čustvom E je torej definirana kot vsota vseh zmnožkov relativnih frekvenc n različnih leksemov (w_k), njenega količnika čustvene zaznamovanosti (q_e) in količnika čustvene intenzivnosti (q_i):

$$I(P, E) = \sum_{k=1}^n \frac{f_a(w_k) * q_e(w_k, E) * q_i(w_k, E)}{N}$$

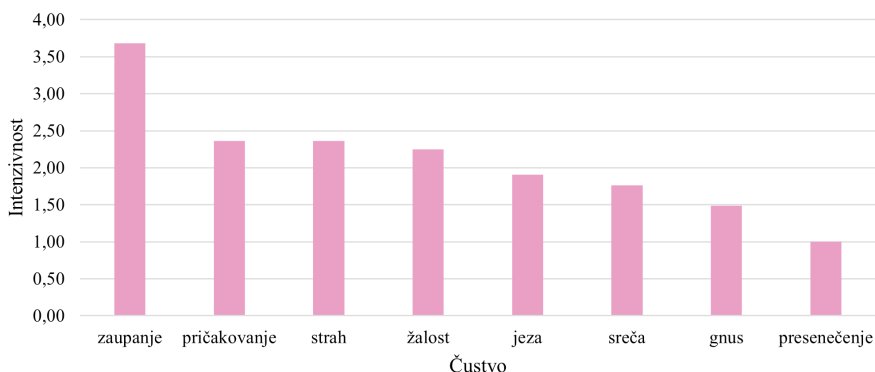
Za vsakega od približno 1.370 odstavkov v skupno 15 delih (14 poglavij in pripis) tako pridobimo številsko razporeditev vseh osmih čustev. Višje vrednosti čustvene

intenzivnosti nakazujejo, v katerih odsekih romana naletimo na zgoščene skupke besed, ki asociirajo na določeno čustvo. Rezultate in izsledke si podrobneje ogledamo v nadaljevanju.

3 Analiza

3.1 Globalna razporeditev čustev v romanu

Če čustveno intenzivnost izračunamo za celoten roman za vseh osem osnovnih čustev, dobimo razporeditev, kot jo prikazuje Graf 1 (vrednosti so normalizirane tako, da ima najmanj zastopano čustvo vrednost 1,0). Če enostaven pregled frekvenčnih seznamov besed, ki največ prispevajo k čustveni intenzivnosti, ponudi nekaj zanimivih vpogledov v poglobitve tematike romana.



Graf 1: Globalna razporeditev čustev v romanu *Visoška kronika*.

Največjo čustveno težo doprinesejo besede, ki bralca asociirajo na zaupanje, med njimi npr. *oče, mati, brat, pravica, resnica, sodnik, vera, cerkev, nebeški, prijatelj* in *general*. Besede dobro odsevajo nekatere ključne motive v romanu – zapletene družinske in medčloveške odnose v času vojne in po njej, skrivnosti, ki pritiskajo na te odnose, ter iskanje pravice in resnice kot del pokore (tudi v religioznem kontekstu). Med negativnimi čustvi prevladujeta strah in žalost, npr. z besedami *kri, umreti, smrt, vojska, krvav, mrtev, morilec, hudič, greh, peklenški, grešnik* in *pokora*. Nabor nakazuje, da se odlomki, ki so najbolj čustveno negativni, vrtijo okrog za zgodbo ključnega Polikarpovega zločina in njegovih posledic za družino Khallan. Po drugi strani je v romanu mnogo manj srečnih trenutkov – poleg splošnih pozitivnih besed (*dober, lep, veselje*) k čustveni intenzivnosti sreče veliko prispevajo *ples* in *plesati* (prizor plesa po Agatinem prihodu na Visoko) ter *nevesta, ženin* in *zakon* (poroka Agate in Jurija ter Izidorja in Margarete na koncu romana).

3.2 Časovnica čustev skozi roman

Spremembe intenzivnosti določenega čustva skozi knjigo lahko vizualiziramo s črtnim grafom. Tovrstna analiza s ptičje perspektive razkrije, kateri odseki besedila so čustveno najbolj izstopajoči. Primer naraščanja in upadanja žalosti z drsečim povprečjem

prikazuje Graf 2, na kateri črtkana črta predstavlja globalno povprečje intenzivnosti, pikasta črta mediano, polna črta pa intenzivnost žalosti po posameznih odstavkih.



Graf 2: Naraščanje in padanje intenzivnosti žalosti skozi roman *Visoška kronika*. Na grafu so razdvoumljeni nekateri lokalni maksimumi, ki ponazarjajo ključne žalostne dogodke.

Maksimumi na grafu ponazarjajo tiste dele besedila, ki glede na leksikon čustvenih asociacij bralca najbolj asociirajo na žalost – takoj na začetku je npr. opis Polikarpovega okrutnega značaja, ki je grenilo življenje Khallanovih:

samo to sem občutil, da je bil Polikarp Khallan trd, temen in brezsrčen gospodar, kateremu ni nikdar posijalo sonce v dušo. Čez vse mere je delal grdo s svojo zakonsko ženo, katera mu je rodila mene in pet let mlajšega brata Jurija. Naša hiša ni bila hiša božjega blagoslova. Imeli smo več nego drugi, posedovali smo vsega dosti, a pod nobeno streho se ni tako malo molilo in tako obilo preklinjalo kakor ravno pod našo.

Z grafa je razvidno tudi, da lahko roman vsebinsko razdelimo na dva dela. Razmejitev med njima je dogodek, ki tudi s kvantitativnega vidika vzbuja največ žalosti: pogovor med Polikarpom in Izidorjem tik pred Polikarpovo smrtjo. V njem Polikarp prizna, da je umoril Jošta Schwarzkoblerja, in Izidorju naloži pokoro za svoj greh, kar je uvod v drugi del romana. Najbolj žalosten dogodek v drugi polovici zgodbe je sojenje Agati Schwarzkobler, njegovi razrešitvi pa sledi grenek zaključek *Visoške kronike*.

3.3 Korelacije med liki in čustvi

Orodje CLASSLA-Stanza v besedilu označi tudi imenske entitete, tj. eksplicitne omembe oseb, krajev ipd. Na podlagi tega lahko izračunamo, v katerih delih besedila se najpogosteje omenja določen lik, in analiziramo korelacije med omembami določenega lika in intenzivnostjo čustev. Tabela 2 prikazuje statistično pomembne Pearsonove koeficiente korelacije ρ (na lestevici od -1 (negativna linearna korelacija) do $+1$ (pozitivna linearna razmerje)) med intenzivnostjo določenega čustva in omembami določenega lika v romanu (ob upoštevanju le tistih odstavkov, v katerih je lik eksplicitno omenjen).

Lik	Čustvo	ρ	p
Jošt	gnus	0,64	$\leq 0,01$
Jošt	strah	0,66	$\leq 0,001$
Jošt	žalost	0,53	$\leq 0,05$
Agata	zaupanje	-0,19	$\leq 0,05$
Jurij	presenečenje	0,51	$\leq 0,001$
Margareta	zaupanje	0,44	$\leq 0,05$

Tabela 2: Pearsonovi koeficienti korelacije med relativnimi frekvenca mi omemb likov in intenzivnostjo čustev (na ravni odstavka).

Jošt Schwarzkobler se kot sodelujoči v vojni in žrtev Polikarpovega zločina omenja predvsem v kontekstih, v katerih prevladujejo strah, gnus in žalost. Agata se po drugi strani pogosteje pojavlja v odsekih, ki ne vzbujajo zaupanja, tj. v kontekstu obtožb, da je čarovnica, in sojenja, ki sledi. Jurij kot Izidorjev aktivnejši brat korelira s presenečenjem v kontekstu njegovega posredovanja v Agatini preizkušnji, ko jo reši iz vode. Margareta, ki na koncu romana skrbi za ranjenega Izidorja in se poroči z njim, še pred tem pa na sojenju priča v prid Agati, korelira z zaupanjem.

4 Zaključek

V prispevku smo na kratko predstavili nekatere možnosti kvantitativne strojne analize razporeditve čustev s pomočjo leksikona čustvenih asociacij na primeru romana *Visoška kronika*. Metoda kljub robustnemu pristopu, ki se na prvi pogled morda zdi preveč površinski, ponuja pričakovane izsledke – da v ključnih delih romana prevladujejo negativna čustva in da so čustveno najbolj intenzivni tisti dogodki, ki so za potek romana najbolj ključni. Izsledke je mogoče uporabiti kot podlago za didaktične pripomočke, ki lahko bolj plastično predstavijo časovnico in dogodke ter njihov morebitni vpliv na bralčevo čustveno dožemanje zgodbe in oceno literarnih oseb. Strojna analiza čustev je uporabna tudi kot predstopnja za poglobljeno kvalitativno ročno analizo, saj pomaga pri identifikaciji relevantnih delov besedila in prihrani ročno delo. Z metodo je mogoče označiti tudi večji korpus slovenske književnosti (kar bi npr. omogočilo iskanje po odlomkih, ki močneje asociirajo na določeno čustvo) ter opraviti primerjalno analizo čustvovanja po različnih delih in avtorjih, kar ponuja dodaten kvantitativni vidik k tradicionalno uporabljenim kvalitativnim metodam in olajša interpretacijo dogajanja.

Omeniti je treba, da ima metoda z uporabo čustvenega leksikona nekaj pomanjkljivosti, ki jih je treba upoštevati pri interpretaciji rezultatov. Postopek je zelo robusten in deluje le na podlagi besednih asociacij površinskih oblik, ne upošteva pa npr. večbesednih enot (*umreti od smeha*) ali zanikanja (»prejšnje veselosti ni več imela«). Primerna je predvsem za daljša besedila oz. vsaj za robustno analizo večjih količin krajših besedil (npr. pesmi). Leksikon SloEmoLex 1.0 je gotovo še pomanjkljiv, saj je njegova osnova angleški leksikon, ki je bil zgolj preveden, zato v njem manjkajo številne slovenske besede (in sopomenke že vključenih besed), s katerimi ga je mogoče še dopolniti.

V prihodnje bi bilo smiselno raziskavo ponoviti na večji količini besedil ter pri analizi upoštevati morebitna dodatna orodja, kot je npr. označevalnik koreferenc, ki bi v besedilu označil tudi implicitne omembe likov (z zaimki in drugimi nanašanji) in tako omogočil natančnejši pregled določene entitete skozi besedilo v povezavi s potekom čustev. Kot alternativno in naprednejšo metodo bi bilo dobro preizkusiti tudi čustveno analizo s pomočjo velikih jezikovnih modelov ter preveriti, ali rezultati, pridobljeni z modeli, natančneje določajo prevladujoča čustva v besedilu kot leksikonski pristop.

Prispevek je nastal v okviru raziskovalnega programa Jezikovni viri in tehnologije za slovenski jezik (P6-0411) in raziskovalnega projekta Veliki jezikovni modeli za digitalno humanistiko (LLM4DH; GC-0002), ki ju financira Javna agencija za znanstvenoraziskovalno in inovacijsko dejavnost Republike Slovenije (ARIS). Iskrena hvala tudi anonimnima recenzentoma za konstruktivne pripombe.

Literatura

- BRGLEZ, Mojca, CAPORUSSO, Jaya, HOOGLAND, Damar, KOLOSKI, Boshko, POLLAK, Senja, PURVER, Matthew, 2024: *Slovenian Emotion Dimension and Emotion Association Lexicon SloEmoLex 1.0*. Slovenian language resource repository CLARIN.SI. <http://hdl.handle.net/11356/1875>
- CAPORUSSO, Jaya, HOOGLAND, Damar, BRGLEZ, Mojca, KOLOSKI, Boshko, PURVER, Matthew, POLLAK, Senja, 2024: A Computational Analysis of the Dehumanisation of Migrants from Syria and Ukraine in Slovene News Media. *Proceedings of the 2024 Joint International Conference on Computational Linguistics, Language Resources and Evaluation (LREC-COLING 2024)*. Torino: ELRA/ICCL. 199–210.
- DAELEMANS, Walter, FIŠER, Fišer, FRANZA, Jasmin, KRANJČIĆ, Denis, LEMMENS, Jens, LJUBEŠIĆ, Nikola, MARKOV, Ilia, POPIČ, Damjan, 2020: *The LiLaH Emotion Lexicon of Croatian, Dutch and Slovene*. Slovenian language resource repository CLARIN.SI. <http://hdl.handle.net/11356/1318>
- ERJAVEC, Tomaž, 2014: *Digital Library and Corpus of Historical Slovene IMP 1.1*. Slovenian language resource repository CLARIN.SI. <http://hdl.handle.net/11356/1031>
- HLADNIK, Miran, POLAJNAR, Jernej, 2016: Wikiji v izobraževanju: Po desetih letih izkušenj. *Andragoška spoznanja* XXII/4. 73–83.
- KIM, Evgeny, KLINGER, Roman, (2019) 2022: A Survey on Sentiment and Emotion Analysis for Computational Literary Studies. *Zeitschrift für digitale Geisteswissenschaften*. <https://doi.org/10.48550/arXiv.1808.03137>
- KLINGER, Roman, KIM, Evgeny, PADÓ, Sebastian, 2020: Emotion Analysis for Literary Studies: Corpus Creation and Computational Modelling. Nils Reiter, Axel Pichler, Jonas Kuhn (ur.): *Reflektierte algorithmische Textanalyse: Interdisziplinäre(s) Arbeiten in der CRETA-Werkstatt*. Berlin, Boston: De Gruyter. 237–268. <https://doi.org/10.1515/9783110693973-011>
- KONCILIJA, Ema, 2024: *Empirična analiza žanra kriminalke na primeru skandinavskih in slovenskih kriminalk*. Magistrsko delo. Ljubljana: Filozofska fakulteta. <https://repozitorij.uni-lj.si/IzpisGradiva.php?lang=slv&id=161808>
- KOSEM, Iztok, ČIBEJ, Jaka, DOBROVOLJC, Kaja, ERJAVEC, Tomaž, LJUBEŠIĆ, Nikola, PONIČVAR, Primož, ŠINKEC, Mihael, KREK, Simon, 2025: *Monitor Corpus of Slovene Trendi 2025-01*. Slovenian language resource repository CLARIN.SI. <http://hdl.handle.net/11356/2013>

- KREK, Simon, ERJAVEC, Tomaž, REPAR, Andraž, ČIBEJ, Jaka, ARHAR HOLDT, Špela, GANTAR, Polona, KOSEM, Iztok, ROBNIK-ŠIKONJA, Marko, LJUBEŠIČ, Nikola, DOBROVOLJC, Kaja, LASKOWSKI, Cyprian, GRČAR, Miha, HOLOZAN, Peter, ŠUSTER, Simon, GORJANC, Vojko, STABEJ, Marko, LOGAR, Nataša, 2019: *Corpus of Written Standard Slovene Gigafida 2.0*. Slovenian language resource repository CLARIN.SI. <http://hdl.handle.net/11356/1320>
- LJUBEŠIČ, Nikola, DOBROVOLJC, Kaja, 2019: What does Neural Bring? Analysing Improvements in Morphosyntactic Annotation and Lemmatisation of Slovenian, Croatian and Serbian. Tomaž Erjavec, Michał Marcińczuk, Preslav Nakov, Jakub Piskorski, Lidia Pivovarov, Jan Šnajder, Josef Steinberger, Roman Yangarber (ur.): *Proceedings of the 7th Workshop on Balto-Slavic Natural Language Processing*. Firenze: Association for Computational Linguistics. 29–34.
- LJUBEŠIČ, Nikola, MARKOV, Ilia, FIŠER, Darja, DAELEMANS, Walter, 2020: The LiLaH Emotion Lexicon of Croatian, Dutch and Slovene. Malvina Nissim, Viviana Patti, Barbara Plank, Esin Durmus (ur.): *Proceedings of the Third Workshop on Computational Modeling of People's Opinions, Personality, and Emotion's in Social Media*. Barcelona: Association for Computational Linguistics. 153–157.
- MOHAMMAD, Saif, 2011: From Once Upon a Time to Happily Ever After: Tracking Emotions in Novels and Fairy Tales. Kalliopi Zervanou, Piroska Lendvai (ur.): *Proceedings of the 5th ACL-HLT Workshop on Language Technology for Cultural Heritage, Social Sciences, and Humanities*. Portland: Association for Computational Linguistics. 105–114.
<https://aclanthology.org/W11-1514.pdf>
- MOHAMMAD, Saif, TURNEY, Peter, 2010: Emotions Evoked by Common Words and Phrases: Using Mechanical Turk to Create an Emotion Lexicon. Diana Inkpen, Carlo Strapparava (ur.): *Proceedings of the NAACL-HLT 2010 Workshop on Computational Approaches to Analysis and Generation of Emotion in Text*. Los Angeles: Association for Computational Linguistics.
- MOHAMMAD, Saif, TURNEY, Peter, 2013: Crowdsourcing a Word-Emotion Association Lexicon. *Computational Intelligence XXIX/3*. 436–465.
- PLUTCHIK, Robert, 1982: A Psychoevolutionary Theory of Emotions. *Social Science Information XXI/4–5*. 529–553. <https://doi.org/10.1177/053901882021004003>
- SAMOTHRAKIS, Spyridon, FASLI, Maria, 2015: Emotional Sentence Annotation Helps Predict Fiction Genre. *PLoS ONE X/11*. <https://doi.org/10.1371/journal.pone.0141922>
- ŽAGAR, Aleš, KAVAŠ, Matic, ROBNIK-ŠIKONJA, Marko, ERJAVEC, Tomaž, FIŠER, Darja, LJUBEŠIČ, Nikola, FERME, Marko, BOROVIC, Mladen, BOŠKOVIČ, Borko, OJSTERŠEK, Milan, HROVAT, Goran, 2022: *Corpus of Academic Slovene KAS 2.0*. Slovenian language resource repository CLARIN.SI. <http://hdl.handle.net/11356/1448>
- ŽEJN, Andrejka, 2020: Računalniško podprta stilometrična analiza pripovedne literature Janeza Ciglerja in Christopha Schmida v slovenščini. *Fluminensia XXXII/2*. 137–158.
<https://hrcak.srce.hr/file/362519>