

SERBIAN LANGUAGE INSTITUTE OF SASA

LEXICOLOGY AND LEXICOGRAPHY
IN THE LIGHT OF CONTEMPORARY
APPROACHES

A collection of papers

Edited by:

Stana Ristić, PhD, scientific advisor

Ivana Lazić Konjik, PhD, research associate

Nenad Ivanović, PhD, research associate

Belgrade, 2016

ISBN 978-86-82873-??-?

ИНСТИТУТ ЗА СРПСКИ ЈЕЗИК САНУ

ЛЕКСИКОЛОГИЈА И ЛЕКСИКОГРАФИЈА У СВЕТЛУ САВРЕМЕНИХ ПРИСТУПА

Зборник научних радова

Уредништво:

др Стана Ристић, научни саветник
др Ивана Лазић Коњик, научни сарадник
др Ненад Ивановић, научни сарадник

Београд, 2016

Радослава М. ТРНАВАЦ
Филозофски факултет, Универзитет у Новом Саду
radoslava.trnavac@ff.uns.ac.rs

АНАЛИЗА СЕНТИМЕНТА: СИСТЕМ ОЦЕНЕ И ПРИРОДНИ СЕМАНТИЧКИ МЕТАЈЕЗИК*

У раду су приказана два лексичка приступа евалуацији и евалуативним придевима у тексту: системско-функционални систем оцене (Martin & White 2005) и природно семантички метајезик (Goddard et al., у штампи). Циљ рада је да опише класификацију евалуативних придева на примеру српског језика у наведена два система, као и да истакне потенцијал наведених система у смислу њихове хибридне употребе у сентимент анализи, која се бави аутоматском екстракцијом субјективног садржаја из дискурса (Pang & Lee 2008).

Кључне речи: систем оцене, природни семантички метајезик, сентимент, евалуација у тексту

Увод

Анализа сентимента или аутоматска екстракција субјективног садржаја/лексикона, као и језик евалуације¹, представљају у савременој рачунарској лингвистици активно подручје истраживања. Циљ сентимент

* Ово истраживање је спроведено у оквиру пројекта “The emotion-cognition ties” финансираног из средстава пројекта ‘Griffith University/Simon Fraser University Collaborative grant’.

¹ Језик евалуације је анализиран у оквиру многобројних праваца у лингвистици. Погледати опширније о афекту (Batson et al. 1992), субјективности (Traugott 1995, 2010; Langacker 1990), евиденцијалности (Aikhenvald 2004; Chafe & Nichols 1986), перспективи (Biber & Finegan 1988, 1989), модалности (Bybee & Fleischman 1995; Trnavac 2006) и оцени (Martin & White 2005). За детаљнији преглед литературе о језику евалуације погледати Trnavac & Taboada (20012).

мент анализе је да се одреди да ли је одређени текст² субјективан (видети Pang & Lee 2008), и, ако јесте, да ли изражава позитивну или негативну евалуацију (поларитет). Централну улогу у процени субјективности у тексту имају евалуативни придеви. Један од лингвистичких приступа који се често примењује у сентимент анализи, а користи евалуативне придеве, је систем оцене (the Appraisal Framework, Martin & White 2005), који је развијен под окриљем системско-функционалне лингвистике (Halliday 1985; Halliday & Matthiessen 2014). У оквиру семантичких истраживања, теорија природног семантичког метајезика такође се бави проучавањем лексикона у домену емоција и валуација (Wierzbicka 1999, 2009; Harkins & Wierzbicka 2001; Gladkova 2010, Goddard et al., у штампи). Овај рад има за циљ да опише резултате проучавања наведена два приступа у оквиру започетих истраживања у области евалуативних придева. Резултати тих истраживања, који су спроведени на енглеском језику, су у детаљнијем облику приказани у Годард и др. (у штампи). Рад такође презентира могуће перспективе за хибридну употребу система оцене и природног семантичког метајезика у домену сентимент анализе.

У наредна два одељка биће представљени систем оцене и природни семантички метајезик, као и претходна истраживања унутар датих приступа везана за евалуативне придеве као главне носиоце сентимента у тексту (Martin & White 2005, Martin, у штампи, Goddard et al., у штампи). Након тога следи поглавље које дефинише основне појмове из сентимент анализе. Закључно поглавље је фокусирано на могућности заједничке употребе два приступа у оквиру сентимент анализе.

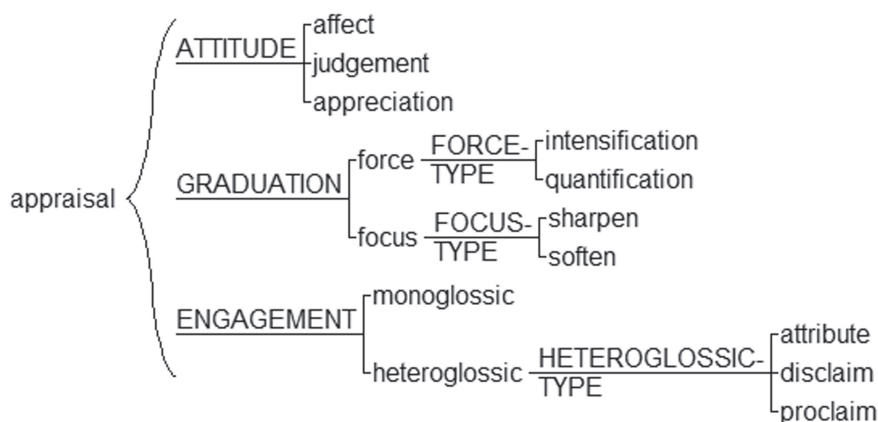
II. Систем оцене и евалуативни придеви

Систем оцене (Martin & White 2005), који је настао у Аустралији, и који је развијан током деведесетих и двехиљадитих година, представља скуп језичких средстава за изражавање емоција, судова и процена, као и средстава за амплификацију тих процена. Наведени приступ је примењен на политички дискурс (White 1998; Coffin & O'Halloran 2006), дискурс у оквиру образовања и развоја писмености (Macken-Horarik 2003; Page 2003), на област језичког усвајања код деце (Painter 2003), као и на анализу књижевних текстова (Love 2006) и свакодневних конверзација (Eggs & Slade 1997). Најновије примена овог система је у оквиру сентимент анализе, о којој ће бити више речи у одељку

² Под 'текстом' се подразумева било који лингвистички израз почев од индивидуалних речи, преко фраза и реченица било које врсте, у писаној или усменој комуникацији.

IV. Систем евалуације је базиран на принципима системско-функционалне лингвистике (Halliday 1985; Halliday & Matthiessen 2014), те је формиран од подсистема категорија и опозиција, а његови темељи су у почетку били засновани на граматичкој, а тек касније и на лексичкој анализи (погледати опширније Goddard et al., у штампи).

Систем оцене има три подсистема (слика 1): систем за изражавање става (Attitude), систем за изражавање ангажмана (Engagement), односно начина на који аутор позиционира себе у односу на став коју презентује (цитирање, усаглашавање, неслагање...) и систем градиције (Graduation) који изражава способност говорника да интензивира или редукује снагу свог мишљења.

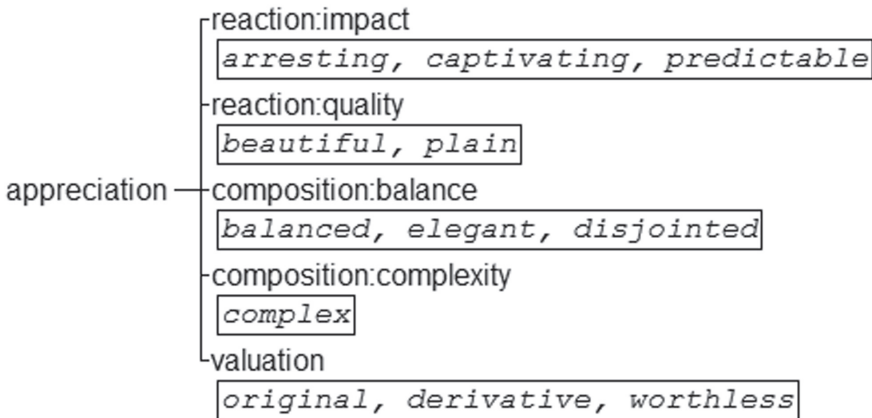


Слика 1. Систем оцене (адаптирано из Martin & White 2005: 38)

Евалуативни придеви који су предмет овог рада припадају првом подсистему, средствима којима говорник изражава свој став (Attitude) у погледу емоција, моралних вредности или естетске евалуације људи и ствари. Подсистем који изражава ангажман користи најчешће модална средства, кондиционалне или концесивне реченице. Ова средства исказују позицију говорника везану за мишљење које је изречено, док подсистем градиције, који најчешће укључује прилоге, интензивира евалуативне придеве или их тонира ка позитивном односно негативном поларитету (нпр. *веома њријатан, доста компликован*).

Као што је приказано на слици 1, подсистем који садржи средства за говорников став обележава емоције (affect), моралне (Judgement) и естетске (Appreciation) ставове. За сваки подсистем је развијен прототипски реченични фрејм (Taboada & Grieve 2004). Средства којима се изражавају емоције би била изражена фразом *Ја осећам X (сипрах)*,

моралне вредности би биле исказане сегментом *Он је X (iправедан)*, док се естетске вредности изражавају реченицом *То је X (леiо)*. Евалуативни придеви често припадају трећој категорији за изражавање става - естетским вредностима. У оквиру те поткатегорије издвојила би се следећа средства: средства за изражавање учинка (*имiресивно*) и квалитета (*дивно*) који има одређена реакција, средства за изражавање равнотеже (*елеганiно*) и комплексности нечега (*једносиiавно*), као и средства за изражавање валуације (*скуiо*), што је приказано на слици 2:



Слика 2 Категорије естетских вредности (Martin & White 2005: 56)

Приликом коришћења наведеног система евалуације треба имати у виду да су његове категорије хипотетички креиране и да доза субјективности у њима постоји, као и да је евалуативна вредност придева одређена и контекстом. Ипак, постоје предвиђена правила по којима се врши анотација текстова и категоризација евалуације у оквиру њих (погледати Taboada & Carretero 2012).

III. Природни семантички метајезик

У овом одељку ћемо приказати класификацију евалуативних придева која је обрађена у недавно завршеном раду у оквиру теорије природног семантичког метајезика (Goddard et al., у штампи). Творац семантичке анализе когнитивно-функционалног карактера базиране на природном семантичком метајезку је А. Вјежбицка (1972). Природни семантички метајезик је заснован на редукованом парафразирању у циљу експликације различитих језичких концепата на основу ограниченог броја речи-примитива чије је значење, према главној хипотези ове теорије, еквивалентно у свим језицима света (Wierzbicka 1996,

Gladkova 2010; Goddard 2011; Goddard & Wierzbicka 2014). У табели 1 је представљена најновија листа семантичких примитива која броји 65 концепата и која је преведена на српски језик а преузета из оригиналне листе на енглеском језику публиковане у књизи Годарда и Вјежбицке (2014):

Табела 1. Листа семантичких примитива

ЈА, ТИ, НЕКО, НЕШТО~СТВАР, ЉУДИ, ТЕЛО	супстантиви
ВРСТА, ДЕО	релацијски супстантиви
ОВАЈ (ТАЈ)~ОВАКО(ТАКО), ИСТИ, ДРУГИ	детерминатори
ЈЕДАН, ДВА, НЕКОЛИКО, СВЕ, МНОГО, МАЛО	квантификатори
ДОБРО, ЛОШЕ	евалуатори
ВЕЛИКИ, МАЛИ	дескриптори
ЗНАТИ, МИСЛИТИ, ХТЕТИ/ЖЕЛЕТИ, НЕ ЖЕЛЕТИ, ОСЕЋАТИ, ВИДЕТИ, ЧУТИ	ментални предикати
РЕЋИ, РЕЧИ, ИСТИНА/ИСТИНИТО	говор
РАДИТИ, ДЕСИТИ СЕ, КРЕТАТИ СЕ	активност, догађаји, покрет
БИТИ (НЕГДЕ), ПОСТОЈИ, БИТИ (НЕКО/НЕШТО)	локација, постојање, спецификација
(БИТИ) МОЈ	посесивност
ЖИВЕТИ, УМРЕТИ	живот и смрт
КАДА~ВРЕМЕ, САДА, ПРЕ, ПОСЛЕ, ДУГО, КРАТКО, НЕКО ВРЕМЕ, ТРЕНУТАК	време
ГДЕ~МЕСТО, ОВДЕ, ИЗНАД, ИСПОД, ДАЛЕКО, БЛИЗУ, ПО/НА СТРАНИ, УНУТАР, ДОДИРНУТИ/ДОДИРИВАТИ	простор
НЕ, МОЖДА, МОЖЕ, ЗБОГ, АКО	логички концепти
ВЕОМА, ВИШЕ	интензификатор, аугментатор
КАО	сличност

Семантички примитиви наведени у табели 1 су резултат дуготрајне емпиријске анализе спроведене на око 30 типолошки различитих језика. Систем такође користи између 60 и 80 речи које имају хипотетички универзални карактер, али представљају концепте који су сложенији од примитива, тзв. семантичке молекуле (Goddard & Wierzbicka 2014). Семантичка анализа језичких концепата према теорији природног се-

мантичког метајезика обухвата лингвистички поједностављене експликације засноване на субјективној експресији и утврђеним граматичким обрасцима у оквиру природног језика израженог у првом лицу. Наведеним експликацијама се саопштава оно шта говорник сматра да је значење одређене речи/концепта, лингвистичког израза или граматичке структуре.

У оквиру експликације која има за циљ да семантички анализира одређени језички концепт издвајају се три обавезна дела. Прво се наводи прототипска мисао субјекта, затим садржај који непосредно описује дати језички концепт, а затим следи консеквенца која изражава одређену емоцију (добру или лошу) и која је типична за описану врсту мисли. Сам садржај описаног концепта у оквиру експликације подсећа на поједностављену речничку дефиницију појма јер поседује ограничени број речи одређених лексичким језгром примитива.

У недавном раду о семантички евалуаторима (Goddard & et al, у штампи.) аутори на основу емпиријске и корпусне анализе издвајају четири основна типа евалуативних придева у оквиру теорије природног семантичког метајезика. Ми ћемо их у овој студији представити, али не у смислу полемике да ли наведене групације у потпуности одговарају реалности српског језика, с обзиром да се поменути рад тиче корпуса енглеског језика, већ у смислу применљивости теорије природног семантичког метајезика на сентимент анализу.

Први или А тип у оквиру теорије природног семантичког метајезика укључује придеве као што су *диван, сјајан, одличан*. Он семантички садржи „мисао која се јавља у првом лицу и која изазива неку емоцију“ (Goddard & et al., у штампи). За њих се везује изузетна субјективност. Придеви који припадају другом, типу В (*занимљив, њојресан*), не садрже толики степен субјективности, али ипак поседују у свом значењу и мисао и осећање, и анализирају се као “искуствени”. Наведени тип садржи две поткатегорије од којих В2 верзија укључује додатну семантичку компоненту која се дотиче потенцијалног утицаја емоције на људско тело. Тип С садржи вокабулар који имплицира неки дуготрајнији утицај на свест субјекта, типа *знамений, моћан*. Последњи тип, тип D, обухвата когнитивне евалуаторе, као што су *комплексан, бриљантијан* који су менталне природе и који у себи не садрже експлицитну емотивну компоненту.

У следећем одељку ћемо илустровати све типове евалуативних придева понудивши експликације у оквиру сваког од типова.

IV. Класификација евалуативних придева у оквиру природног семантичког метајезика

Као што је приказано у табели 2, тип А у уводној компоненти своје експликације поседује когнитивни елемент: „Ја мислим о овом X овако:..“ Након тога следи садржај који описује дати евалуативни придев, да би се експликација завршила следећом евалуацијом – „овај X је веома добар“ или „овај X је веома лош“.

Табела 2 Тип А (*чудесан, њредиван*)

Ја мислим о овом X овако:	МИСаО У ПРВОМ ЛИЦУ
„ --- --- “	МИСаО
Када ја мислим овако, ја осећам нешто веома добро/лоше због тога	ОСЕЋАЊЕ

Наведена група придева у оквиру система оцене би припадала подсистему афекта. Као пример, навешћемо експликацију која би илустроваала дати тип употребе кроз придев *њредиван*:

Табела 3 [1] *Предиван/њредивна* (сан, дан, успомена, итд.)

Ја мислим о овом X овако:	МИСаО У ПРВОМ ЛИЦУ
„Овај X је веома веома добар Ако људи знају о овоме, они осећају нешто веома добро због тога“	МИСаО
Када ја мислим овако, ја осећам нешто веома добро због тога	ОСЕЋАЊЕ

Следећи тип са подгрупама В1 и В2 обухвата искуствене евалуативне изразе који у много мањем степену садрже компоненту емоционалности него тип А. За разлику од експликација који су саставни део типа А и које започињу примитивима „Ја мислим о овом X овако“, уводни део експликација типа В има следећу компоненту: „Неко може да мисли овако (о овоме)“. Ови евалуатори укључују постојање хипотетичке особе („некога“), као и емоције које се везују за ту хипотетичку особу. С обзиром да експликација датог подтипа садржи компоненту којом се наводи мишљење говорника, наведени подтип евалуативних придева се зближава са подсистемом ангажмана у оквиру система оцене. Наведени придеви се повезују са оним како неко размишља и како се неко

осећа док присуствује описиваном догађају. Веома често, ови придеви се употребљавају у дуративним контекстима јер се размишљање о датом догађају може учестало понављати, па ће семантичка експликација укључивати следећу компоненту: „у току овог времена, тај неко може мислити овако много пута“. Уколико се не ради о дуративном контексту, искуствена евалуација може да буде карактеризована као мисао која се појављује једном („када се нешто деси некоме, када неко види или чује нешто једном, онда тај неко може да мисли о томе овако...“).

Типска експликација за подтип В1, са придевима као што су то *занимљив*, *интересантан*, *досадан*, *предвидљив* је представљена у следећој табели, која укључује компоненту дуративног контекста:

Табела 4 Тип В (*интересантан*)

Током ³ овог времена (на пример, када неко гледа филм)	ДУРАТИВНИ ФРЕЈМЛ
тај неко може да мисли овако много пута:	ПОТЕНЦИЈАЛНА МИСаО
” --- --- “	МИСаО
када неко овакав мисли тако, он/она, могу да осећају нешто (веома) добро/лоше због тога	ОСЕЋАЊЕ

Као пример, у оквиру подтипа В1 илустроваћемо следећу експликацију за придев *занимљив*:

Табела 5 [2] Занимљив/занимљиво (филм, вече)

Током овог времена (на пример, када неко гледа филм)
тај неко може да мисли овако много пута:
„неке ствари се дешавају сада зато што неко хоће да се људи осећају добро као када се људи осећају док хоће да се смеју [м]“ ⁴
када неко овакав мисли тако, он/она могу да осете нешто добро због тога

Подтип В2 има сличан образац експликације као В1, али постоји додатна компонента која укључује неку потенцијалну телесну реакцију. Експликација за придев *интересан* који припада наведеном подтипу укључује дуративну заједно са још неким компонентама:

³ Прилошка одредба *током* не представља семантички примитив, већ замењује у овом контексту примитив *када*.

⁴ *Смејају се* је семантичка молекула, појам који хипотетички постоји у свим језицима света, али се може семантички разложити на једноставније компоненте (Goddard & Wierzbicka, у штампи).

Табела 6 [3] Потресан/потресна (филм, сцена)

Током овог времена (на пример, када неко гледа филм)
овај неко може да мисли овако много пута:
„нешто лоше се дешава сада некоме ко је лош тај неко осећа нешто лоше због тога“
када неко мисли тако, тај неко не може да не осећа нешто због тога
у исто време тај неко може да осећа нешто у телу као људи који осећају да хоће да плачу [м]

Подтип С садржи евалуативне придеве, као што су *инсџирајиван*, *дејресиван*, *узнемиравајући*, који имају у себи компоненту ефекта који траје након самог догађаја, а не у току догађаја. Осим тога, ови придеви такође садрже и социјалну евалуацију догађаја који се описује:

Табела 7 Подтип С (*инсџирајиван*, *дејресиван*)

Када неко ради нешто овако неко време (на пример, када неко гледа филм) нешто се дешава некоме због тога	ЕЕФЕКАТ
због тога, неко време после тога то изгледа овако:	ПОСЛЕ ЕФЕКТА
људи могу мислити о томе овако: „то је добро/лоше“	СОЦИЈАЛНА ЕВАЛУАЦИЈА

Експликација придева *дејресиван* који припада овој подгрупи је понуђена у следећој табели:

Табела 8 [4] *Дејресиван/дејресивно* (човек, стање)

Када неко ради нешто овако неко време нешто се дешава некоме због тога
због тога, неко време после тога то изгледа овако:
тај неко мисли овако понекад: „веома лоше ствари се дешавају људима много пута људи неће да то буде тако, у исто време они не могу да ураде ништа због тога“ када неко мисли тако, он/она не могу да не осете нешто лоше због тога“
људи могу да мисле о томе овако: „То је лоше“

Последњи подтип илуструје свеобухватну когнитивну евалуацију нечега без присуства емоције у значењу евалуативног придева. Ова група придева највише одговара категорији која обухвата моралне и естетске ставове у оквиру система оцена. Образац експликација за наведену врсту придева се може приказати на следећи начин:

Табела 9 Подтип D (*комплексан, одличан, импресиван*)

Ако неко зна како је ово X,	ЗНАЊЕ
он/она могу да мисле о томе (или о некоме) овако:	ПОТЕНЦИЈАЛНА МИСАО
” --- --- “	МИСАО

Наведени подтип придева се поклапа у великој мери са естетским оценама која су посебна подгрупа у оквиру система оцена. Следећа експликација илуструје дати подтип:

Табела 10 [5] Сложена (*књига, мисао, фигура*)

Ако неко зна како изгледа ово X,
он/она могу да мисле о X овако:
„то нешто може да има многе делове многи од тих делова нису као остали због тога, ако неко не зна много ствари о тим деловима, тај неко не може да зна шта је то нешто добро је ако тај неко може да зна шта је то нешто“

На основу горе представљених подгрупа и експликација, могу се извести следећи генерални закључци у смислу упоредивости два система евалуације:

Све четири подгрупе евалуативних придева које припадају природном семантичком метајезику одговарају категорији става или категорији ангажмана у оквиру система оцене.

Подгрупа А је упоредива са подгрупом афекта у оквиру категорије става.

Подгрупе В1 и В2 су према својим параметрима сличне категорији ангажмана.

Подгрупа С је најразноврснија у смислу састава евалуативних придева и одговара свим трима групама из категорије става.

Подгрупа D се рефлектује највише у оквиру категорије моралних и естетских вредности.

Још једна разлика, која није непосредно уочљива из наведене упоредне анализе експликација и класификације оцене, је разлика у начинима степеновања или градијације. Док у оквиру система оцене постоји тростепена подела на квантификаторе *мало*, *средње* или *јуно*, која се тиче интензитета придева, природни семантички метајезик у исте сврхе користи следећа средства унутар експликација: (1) компоненте које упућују на различите квалитативне разлике ‘доброг’ или ‘лошег’, као што су ‘добар’ ‘веома добар’ ‘веома, веома добар’, (2) укљученост или неукљученост у експликацију потенцијалних телесних реакција, (3) компоненте које подвлаче ‘посебност’ нечега, или (4) да ли је опи-

сана реакција оно о чему/шта ‘неко може да мисли/осећа’ или оно о чему/шта ‘неко не може да мисли/осећа’.

У следећем поглављу ћемо представити основне принципе сентимент анализе, као и употребу система оцене у оквиру наведене анализе.

V. Сентимент анализа и систем оцене

Сентимент анализа представља интердисциплинарно поље изучавања које обједињује информације из области лингвистике и информатике, а има за циљ да аутоматски издвоји субјективан садржај из текста. Она одређује позитивну или негативну евалуацију у тексту на основу позитивних и негативних речи, на основу контекста истих речи, као и на бази лингвистичке структуре текста (Taboada, 2016). Најчешће апликације сентимент анализе обухватају аутоматску детерминацију субјективног садржаја из приказа књига, филмова, производа на интернету, као и анализу социјалних медија које врше компаније и политички аналитичари. Један од основних приступа проблему аутоматске екстракције сентимента је лексички приступ обрачунавања семантичке оријентације документа на основу семантичке оријентације речи или фраза у тексту (Turney 2002). На пример, Табоада и др. (2011) описује SO-CAL софтвер, калкулатор за семантичку оријентацију, који служи за аутоматску анализу евалуације у тексту. Поменути софтвер укључује речник од 3 306 придева, 1 068 именица, 701 глагола, 587 прилога рангираних на скали од -5 до + 5, као и различите типове модификатора и негација (локалног и реченичног карактера). Вредност пунозначних речи је изабрана на основу баланса између различитих интерпретација које реч може да има. Осим коришћења датог софтвера, Табоада и Гриве (2004: 159) користе за сентимент анализу и систем оцене ради категоризације субјективности у тексту, да би се утврдило да ли се сентимент у тексту односи на објекте, емоције или нека понашања. Такође, користећи подсистеме за амплификацију и ангажман, аутори анализирају пишчеву или говорникову посвећеност израженом субјективном мишљењу. У почетку анализе аутори, користећи SO-CAL софтвер, додељују вредности семантичке оријентације (позитивне или негативне), а затим да би оценили потенцијалну евалуативну вредност текста, додељују субјективним речима и фразама једну или више карактеристика из система оцене (на пример, једну од карактеристика из подсистема за изражавање става и једну од карактеристика из система градације). Када је текст на основу датих калкулација класификован као позитиван или негативан, следећи корак се састоји у дефинисању чињенице колико приказ изражава афект, суд или естетску оцену. Различити текстови

изражавају различиту фреквентност ових вредности. Док ће, на пример, онлајн опис лека-антидепресанта највише поседовати вредности афекта, књижевна критика би имала највише евалуатора из подсистема за естетске вредности. Такође, аутори на основу посебних формула⁵, утврђују потенцијалну евалуативну вредност придева - вероватноћу која би показивала базичну способност придева да изражава афект, моралну или естетску оцену у субјективно израженом дискурсу.

VI. Систем оцене и природни семантички језик у анализи сентимента

У претходним одељцима је описано на који начин оба лингвистичка система учествују у анализи евалуативног лексикона. Једна од могућих заједничких примена обе теорије би била у оквиру хибридног система где би аотације природног семантичког метајезика биле додате класификацији система оцене, те би сви називи за подсистеме, укључујући афект, градацију и ангажман били изражени експликацијама у оквиру природног семантичког метајезика. Сама класификација евалуативних придева у оквиру природног семантичког метајезика звучи далеко интуитивније за носиоце било ког природног језика, него што је то евалуативни систем базиран на систему оцене који употребљава уско специјализоване термине, као што су то афект и ангажман.

Радећи на пројекту дескрипције евалуативних придева у енглеском језику у оквиру природног семантичког метајезика, аутор овог текста се уверила да експликације могу бити веома корисне у различитим фазама сентимент анализе. При креирању речника са евалуативним лексиконом или при аотацији текстова у смислу субјективности, детаљан опис експликација и њихова веза са категоријама би били корисни у одређењу поларитета и снаге сентимента. Што се тиче снаге сентимента, за разлику од система оцене који има троструку класификацију на придеве слабе, средње и велике снаге, природни семантички метајезик у својим експликацијама може бити делотворнији у опису тог сегмента уз коришћење примитива 'добар' или 'веома добар', 'може', или 'не може'. Класификација система оцене, као и експликације природног семантичког метајезика, могу да буду од помоћи као основна језгра⁶

⁵ Потенцијална евалуативна вредност придева у оквиру система оцене се одређује на основу формула које су сличне формулама за семантичку оријентацију придева коју је Турнер (2002) развио. О томе опширније погледати Табоада и Грив (2004: 160).

⁶ Енглески термин у овом случају је 'seeds' који описује речи чије су вредности у смислу поларитета, типа оцене или експликације већ одређене, и које би могле бити коришћене у сврху идентификације сличних речи због постојећих колокација.

при проширењу евалуативног речника, тако да је асоцијативним путем могуће пронаћи нове речи које су сличне и које се употребљавају у сличним колокацијским везама. Тако нам, на пример, реч *одвратиан* (*īosīūīak*) која представља негативни евалуатор у оквиру подтипа А може бити употребљен да би се идентификовали неки други евалуатори, као што је то *неморалан* (*īosīūīak*), *нечовечан* (*īosīūīak*), које у различитим језицима, по правилу, не морају да припадају истим подтиповима, нити морају бити исти подтип са основном речју -‘језгром’.

У ранијим радовима је запажено да на квалитет евалуације у тексту утичу модалност, контекстуалне и дискурсивне структуре (видети радове Polanyi & Zaenen 2004; Trnavač & Taboada 2012). Модалност утиче на евалуацију у тексту модификујући је као у примеру (2) у односу на пример (1):

- (1) *To је добар часојис.*
 (2) *To може да буде добар часојис.*

У примеру (2) позитивна оцена из примера (1) је умањена употребом модалне речи.

Дискурсивна структура такође може да модификује евалуативни потенција придева (Trnavač & Taboada 2012, Trnavač et al., у штампи). На пример, концесивни и кондиционални дискурсивни односи могу да утичу на промену поларитета целе реченице (погледати Trnavač et al., у штампи, Trnavač & Taboada 2012.), где постоји конфликт између првог дела реченице који је позитиван и другог дела који је негативан. Као илустрацију, истичемо пример који маркира концесиван дискурсивни однос:

- (3) *Иако се Вања боље осећа, још није у њој њуносџи оздравила.*

Детаљна студија евалуативних маркера може да буде проширена и на друге врсте речи осим придева, као и на дискурсивне односе (Trnavač et al., у штампи). Она доприноси и бољем разумевању процесирања, генерисања и превода афективне терминологије⁷. У том смислу, природни семантички метајезик је посебно интересантан јер може бити

⁷ Утицај афекта/сентимента је детаљније представљен у области језичких технологија, областима афективног природног језичког процесирања (Strappavara 1997; Hobbs & Gordon 2008; Calvo et al. 2015), афективног природног језичкогенерисања (Piwek 2003), као и афективног машинског превођења (Aloy et al. 2014). Ови трендови у лингвистичкој и рачунарској литератури су посебно подстакнути развојем социјалних мрежа на интернету, као и интересовањем маркетиншких кућа да идентификују мишљење (сентимент) потенцијалних купаца о различитој роби.

употребљен у функцији превођења с обзиром да примитиви имају универзални карактер и интеркултуралну вредност.

VII. Закључак

У датом раду су представљене карактеристике система оцене и теорије природног семантичког метајезика које су релевантне за сентимент анализу текста. Комбинација суптилно анализираних експликација у оквиру природног семантичког метајезика и класификације евалуативних придева у оба система показују велики потенцијал за коришћење у сентимент анализи, афективном израчунавању вредности текста и класификацији емотивног дискурса.

ЛИТЕРАТУРА

- Aikhenvald, A. 2004. *Evidentiality*. Oxford: OUP.
- Aloy, M., R. M. Seraj, A. P. García Varela, D. Fass, F. Popowich, A. Sarkar, M. Taboada. 2014. An analysis of affective words in Machine Translation. In: *Proceedings of the Northwest Natural Language Processing Workshop*. Seattle, WA. April 2014. (el. izvor). https://www.sfu.ca/~mtaboada/docs/Aloy_Mayo_et_al_NWNLP_2014.pdf
- Batson, D. C., L. L. Shaw, K. C. Oleson. 1992. Differentiating affect, mood, and emotion: Toward functionally based conceptual distinctions. In: *Emotion. Review of Personality and Social Psychology*. M. S. Clark (ed.). Newbury Park, CA: Sage, 294–326.
- Biber, D., E. Finegan. 1988. Adverbial stance types in English. *Discourse Processes* 11(1), 1–34.
- Bybee, J., S. Fleischmann (eds.). 1995. *Modality in grammar and discourse*. Amsterdam: John Benjamins.
- Calvo, R. A., S. D’Mello, J. Gratch, A. Kappas. 2015. *The Oxford encyclopedia of affective computing*. Oxford: OUP.
- Chafe, W., J. Nichols. 1986. *Evidentiality: The linguistic coding of epistemology*. Norwood, NJ: Ablex.
- Coffin, C., K. O’Halloran. 2006. The role of appraisal and corpora in detecting covert evaluation. *Functions of Language* 13(1), 77–110.
- Eggs, S., D. Slade. 1997. *Analysing Casual Conversation*. London: Cassell.
- Gladkova, A. 2010. *Russkaja kul’turnaja semantika: émocii, cennosti, žiznennye ustanovki* [Russian cultural semantics: Emotions, values, attitudes.] Moscow: Languages of Slavic Cultures.
- Goddard, C. (ed.). 2011. *Cross-linguistic semantics*. Amsterdam: Benjamins.
- Goddard, C., M. Taboada, R. Trnavač. In press. The semantics of evaluational adjectives: Perspectives from Natural Semantic Metalanguage and Appraisal. *Functions of Language*.
- Goddard, C., A. Wierzbicka. 2014. *Words and meanings: Lexical semantics across domains, languages, and cultures*. Oxford: OUP.

- Goddard, C., A. Wierzbicka. In press. Explicating the English lexicon of “doing” and “happening”. *Functions of Language*.
- Halliday, M. A. K. 1985. *An introduction to Functional Grammar*. London: Edward Arnold.
- Halliday, M. A. K., C. M. I. Matthiessen. 2014. *An introduction to Functional Grammar*, 4th ed. London: Edward Arnold.
- Harkins, J., A. Wierzbicka (eds.). 2001. *Emotions in cross-linguistic perspective*. Berlin: Mouton.
- Hobbs, J., A. Gordon. 2008. The deep lexical semantics of emotions. In *Proceedings of the International Conference on Language Resources and Evaluation (LREC)*. Marrakech, Morocco.
- Langacker, R. W. 1990. Subjectification. *Cognitive Linguistics* 1(1), 5–38.
- Love, K. 2006. APPRAISAL in online discussions of literary texts. *Text and Talk* 26(2), 217–244.
- Macken-Horarik, M. 2003. APPRAISAL and the special instructiveness of narrative. *Text* 23(2). 285–312.
- Martin, J. R., P. R. White. 2005. *The language of evaluation*. New York: Palgrave.
- Martin, J. R. In press. Attitudinal relations: Continuing the study of lexis. *Estudos e pesquisas em Linguística Sistêmica Funcional*. São Paulo: Mercado de Letras.
- Page, R. E. 2003. An analysis of APPRAISAL in childbirth narratives with special consideration of gender and storytelling style. *Text* 23(2), 211–237.
- Painter, C. 2003. Developing Attitude: An ontogenetic perspective on Appraisal. *Text* 23(2), 183–209.
- Pang, B., L. Lee. 2008. Opinion Mining and Sentiment Analysis. *Foundations and Trends in Information Retrieval* 2(1-2), 1-135.
- Piwek, P. 2003. An Annotated Bibliography of Affective Natural Language Generation. Version 1.3 April, 2003 [Version 1.0 appeared as ITRI Technical Report ITRI-02-02, University of Brighton March 2002].
- Strapparava, C. 1997. Affective NLP. (el. izdanje). <http://people.lett.unitn.it/baroni/tp/materials/AffectiveNLP.pdf>
- Taboada, M., M. Carretero. 2012. Contrastive analyses of evaluation in text: Key issues in the design of an annotation system for Attitude applicable to consumer reviews in English and Spanish. *Linguistics and the Human Sciences* 6(1-3), 275–295.
- Taboada, M., J. Grieve. 2004. Analyzing appraisal automatically. In: *Proceedings of AAAI Spring Symposium on Exploring Attitude and Shanahan, J.G. Affect in Text* (AAAI Technical Report SS-04-07), Y. Qu, J. Wiebe (eds.). Stanford University, CA: AAAI Press, 158–161.
- Taboada, M. (2016) Sentiment analysis: An overview from linguistics. *Annual Review of Linguistics*. ???
- Traugott, E. C. 1995. Subjectification in grammaticalisation. In: *Subjectivity and subjectivisation: Linguistic perspectives*. Stein, D., S. Wright (eds.). Cambridge MA: CUP, 31–54.

- Traugott, E. C. 2010. (Inter)subjectivity and (inter)subjectification: A reassessment. In: *Subjectification, intersubjectification and grammaticalization*. Davids, K., L. Vandelanotte, H. Cuyckens (eds.). Berlin: Mouton, 29–74.
- Trnavac, R. 2006. *Aspect and subjectivity in modal constructions*. Utrecht: LOT.
- Trnavac, R., M. Taboada. 2012. The contribution of nonveridical rhetorical relations to evaluation in discourse. *Language Sciences* 34, 301–318.
- Trnavac, R., D. Das, M. Taboada. 2016. Discourse relations and evaluation. *Corpora*. ???
- Turney, P. 2002. Thumbs up or thumbs down? Semantic orientation applied to unsupervised classification of reviews. In: *Proceedings of the 40th Meeting of the Association for Computational Linguistics*. Stroudsburg, PA: Assoc. Comput. Linguistics, 417–424.
- White, P. R. R. 1998. *Telling media tales: The news story as rhetoric*. Ph.D. dissertation, Sydney: University of Sydney.
- Wierzbicka, A. 1996. *Semantics: Primes and universals*. New York, NY: OUP.
- Wierzbicka, A. 1999. *Emotions across languages and cultures*. Cambridge: CUP.
- Wierzbicka, A. 2009. Language and metalanguage: Key issues in emotion research. *Emotion Review* 1, 3–14.

Радослава Трнавац

СЕНТИМЕНТ-АНАЛИЗ: СИСТЕМА ОЦЕНКИ
И ПРИРОДНО-СЕМАНТИЧЕСКИЙ МЕТАЯЗЫК

Резюме

В работе исследуются два лексических подхода к оценке и оценочным прилагательным в тексте - системно-функциональная система оценки (Martin & White 2005) и природно-семантический метаязык (Goddard et al., в печати). Целью работы является описание классификации оценочных прилагательных в двух системах на примере сербского языка, а также оценка потенциала этих систем для комбинированного употребления в сентимент-анализе, задачей которого является автоматическая экстракция субъективного содержания в тексте (Pang & Lee 2008).