

(Ne)normativnost računalniško posredovane komunikacije v slovenščini: merilo vejice

Damjan Popič, Darja Fišer

Izvleček

V poglavju predstavimo raziskavo rabe vejice v slovenskih tvitih, in sicer preverimo, v kolikšni meri je stava v tovrstnih vsebinah rabljena v skladu z jezikovnim standardom in na katerih mestih uporabniki s stavo vejice najbolj odstopajo od pravopisa. S tem nadgrajujemo pretekle raziskave, ki so bile usmerjene predvsem v ugotavljanje, v katerih primerih imajo uporabniki največ težav, in poskušamo predstaviti nekoliko bolj celostno sliko rabe vejice v slovenskih uporabniških spletnih vsebinah. Rezultati raziskave kažejo, da je v slovenski računalniško posredovani komunikaciji standardna stava vejice pogostejša kot nestandardna, lahko pa vsaj za določen del gradiva trdimo, da je vejica stavljena nestandardno (tj. izpuščena) namerno, kot del neformalne komunikacije.

Ključne besede: vejica, pravopis, jezikovni standard, računalniško posredovana komunikacija, slovenščina

1 UVOD

Raziskave računalniško posredovane komunikacije, zlasti v razmerju do splošnega jezika,¹ so bile v zadnjem obdobju močno obogatene in razširjene, zlasti s korpusno-jezikoslovno analizo v okviru projekta JANES (prim. npr. Fišer et al. 2017 in Arhar Holdt 2017). V okviru normativnostne obravnave korpusa Janes je bilo še posebej veliko pozornosti posvečene presoji standardnosti računalniško posredovane komunikacije (Arhar Holdt in Dobrovoljc 2016, Fišer et al. 2015, Goli et al. 2015, Osrajnik et al. 2015, Škrjanec et al. 2015), kar je glede na slovensko preskriptivistično tradicijo in direktivno naravo jezikoslovne misli pričakovano, kot je pričakovano tudi to, da je bilo kar nekaj pozornosti posvečene obravnavi vejice (Popič et al. 2016), ki je glede na sodobne empirične raziskave temeljna pravopisna težava v slovenskem jeziku (Kosem et al. 2012, Rozman et al. 2012, Popič 2014, Logar in Popič 2015; Popič in Logar 2015, Popič in Fišer 2015), obenem pa je v slovenski jezikovni tradiciji tudi nosilka prestiža.² To pomeni, da vejici pripisujemo status preizkusnega kamna jezikovne omike (Popič in Logar 2015), kultiviranosti in socialne stratifikacije, in čeprav je prestižnost ena od lastnosti knjižnojezikovne norme oz. standardnega jezika (Nebeská 1999), lahko v slovenski računalniško posredovani komunikaciji vidimo, da se je ta vrednotenjski sistem v veliki meri prenesel tudi na splet (glej Popič in Fišer 2017).³

Iz tega razloga v pričujočem prispevku analiziramo (ne)standardnost stave vejice v slovenskih uporabniških spletnih vsebinah, in sicer želimo številne (tudi lastne; Popič in Fišer 2015, Popič et al. 2016) raziskave nenormativne stave vejice razširiti z analizo normativno ustrezne stave.⁴ Cilj je prikazati nekoliko bolj celostno podobo rabe vejice v tvitih, na način, ki ni obremenjen s predsodki o boljši in slabši slovenščini glede na namen in medij določene jezikovne zvrsti (prim. npr. Jakop 2008). S tem se nanašamo predvsem na prepričanje, da je jezik v računalniško posredovani komunikaciji izrazito nestandarden.

Nedavne raziskave so tovrstne predsodke v precejšnji meri ovrgle. S primerjalno raziskavo korpusov Janes (Erjavec et al. 2018) in Kres (Logar et al. 2012) smo

-
- 1 Izraz *splošni jezik* uporabljamo z namenom, da se izognemo diadi med računalniško posredovano komunikacijo in standardnim oz. knjižnim jezikom. Tovrstna delitev je v prvi vrsti nestrokovna, predvsem pa je problematično to, da je lahko tudi diskriminatorska v razmerju do računalniško posredovane komunikacije (predvsem v smislu, da je ena od temeljnih lastnosti knjižnega jezika oz. njegove norme prestižnost). V besedilu se torej z izrazoma *standardna* in *nestandardna slovenščina* nanašamo izključno na dejansko jezikovno skladnost s pravopisnim standardom.
 - 2 Na to kaže tudi nedavno parlamentarno glasovanje o vejici. 21. junija 2017 je državni zbor namreč glasoval o dopolnitvi prve vrstice Župančičeve pesmi Domovina je ena za inskripcijo na Spomeniku vsem žrtvam vojn (in z vojnamii povezanim žrtvam na območju Republike Slovenije), in sicer je šlo za glasovanje o izbrisu vejice. Glasovanje, ki se je sicer končalo z rezultatom 54 : 6 v podporo izbrisu, je dober indikator odnosa slovenske družbe do jezika, tudi do vejice, predvsem glede vprašanja jezikovne regulacije – za slovensko okolje je namreč značilno prepričanje, da je mogoče o jezikovnih vprašanjih odločati in tudi razsojati s centralne pozicije moči.
 - 3 To dokazujejo tudi nedavne raziskave odnosa do jezika (in same vejice) na Twitterju (Popič in Fišer 2017), ki kažejo, da se vejica tudi v slovenskem računalniško posredovanem diskurzu uporablja kot orodje prestiža in kot sredstvo dokazovanja, da so argumenti nekoga odvisni od jezikovne pravilnosti ubeseditve tega argumenta.
 - 4 Pričujoče poglavje razširja raziskavo o nestandardni rabi vejice v slovenski računalniško posredovani slovenščini (Popič et al. 2016).

tako v preteklih raziskavah pokazali, da raba vejice na družbenih omrežjih ne peša in da so razlike med rabo vejice v tradicionalnih in novomedijskih besedilih manjše, kot pregovorno velja (Popič in Fišer 2015). Pri tem smo se omejili na ožji nabor tipičnih atraktorjev vejic (Verovnik 2003; Žibert 2006) in na samo pogostnost ter razporejenost vejic, zato smo želeli nadgraditi in pripraviti izhodišče za celovito analizo rabe vejice, tako standardne kot tudi nestandardne, predvsem pa razviti čim bolj univerzalen kategorizacijski sistem za označevanje rabe vejice in ga preizkusiti na širšem vzorcu.

V poglavju tako pokažemo, v čem raba vejice v uporabniških spletnih vsebinah odstopa od pravopisnega standarda in v čem mu sledi, obenem pa poskušamo podati preliminarne ugotovitve, v kolikšni meri je nestandardna stava vejice posledica zavestne odločitve in v kolikšni posledica nepoznavanja jezikovnega standarda. Analiza temelji na označevanju napačno in pravilno stavljene vejice na naključnem vzorcu 500 tvitov iz korpusa Janes v0.4 (Popič et al. 2017). Pri tem smo kot merilo upoštevali še stopnjo jezikovne in tehnične standardnosti besedil (Ljubešič et al. 2015) ter v vzorec zajeli po 250 tvitov z naslednjima oznakama: T1L3 (nestandarden jezik, a standardna raba velikih začetnic, presledkov, ločil) in T3L3 (tako tehnično kot tudi jezikovno povsem nestandardni tviti). Ker nas zanima stava vejice v nestandardni slovenščini, tvitov, zapisanih v standardni slovenščini (L1), v analizo nismo zajeli, kar sicer pomeni, da so rezultati nekoliko manj posplošljivi, a so toliko bolj dragoceni zaradi tega, ker lahko na ta način vidimo dejansko vlogo vejice – v besedilih, ki s standardnostjo zapisa vsaj predvidoma niso obremenjena, lahko sami prisotnosti vejice v besedilu pripišemo določeno vrednost.

Na podlagi označenega korpusa tako prikažemo temeljne težave v zvezi s stavo vejice v spletni slovenščini, obenem pa tudi mesta, na katerih tovrstnih težav uporabniki nimajo. Naša osnovna podmena, ki izhaja iz preteklih raziskav, je, da imajo uporabniki težave z vejico predvsem na mestih, ki so skladijsko manj predvidljiva in obenem tudi manj formalizirana (z vidika napovedljivosti stave vejice zgolj na podlagi izraženega veznika). Prav tako je utemeljeno pričakovati, da rezultati ne bodo odgovarjali zgolj na vprašanje, kje imajo jezikovni uporabniki težave, temveč tudi, ali in kje se vejica v računalniško posredovani komunikaciji – vsaj v tisti tehnično manj standardni, s katero se v raziskavi ukvarjamo – zavestno izpušča.

2 VEJICA V (SPLETNI) SLOVENŠČINI

V slovenski jeziko(slo)vni tradiciji je vejica malodane razvpita (glej Popič in Fišer 2017), in sicer velja za nekaj, kar je praktično nemogoče popolnoma osvojiti

oz. obvladati. Tovrstno izročilo je morda nekoliko pretirano, a vendarle nedavne empirične raziskave tovrstnim sodbam v veliki meri pritrjujejo (glej Kosem et al. 2012; Popič 2014; Popič in Logar 2015), saj kažejo, da imajo uporabniki z vejico velike težave. Vejica pa je tudi v jezikoslovnih krogih pogosto povzročala težave, saj gre za »ločilo, o stavi katerega so si bili kritiki najbolj in najpogosteje navzkriž« (Dobrovoljc 2004: 188). Verjetno lahko glavne razloge za to iščemo v sami naravi stave vejice v slovenščini, ki je v slovenski preskripciji izrazito sintaktična, to pa velja za celotni proces normiranja jezika od prvega, Levčevega pravopisa (SP 1899) naprej. Ta je kot prvi slovenski pravopis vejico predpisal predvsem na podlagi skladenjsko-logičnih meril, in sicer je Levec kot izpostavne prepoznal naslednje prvine (Levec 1899: 103–108):

- prosti stavki,
- skrženi stavki,
- priredja,
- podredja,
- skrajšani stavki,
- periode med posameznimi stavki prve in druge polovice (proreka in poreka).

Kot lahko vidimo, je razdelitev izrazito sintaktična in ne omogoča omahovanja na podlagi govora ali pomena. To je sicer značilno za celotno slovensko zgodovino predpisovanja rabe vejice (in drugih elementov jezika nasploh), saj je v zgodovini v resnici prišlo zgolj do enega resnega liberalnega posega v stavo ločil, in sicer s prvim akademjskim pravopisom (SP 1950: 43; poudarila avtorja):

Modulacijo glasu in presledke, s katerimi v govoru vežemo in ločimo pomenske enote, zaznamujemo v pisavi z ločili. Nekatera teh imajo objektivno logično vrednost in so potrebna za lažje in zanesljivo razumevanje pisanega jezika, zato so določena po pravopisnih pravilih. Včasih pa **pisatelj** z ločili izraža osebno ali čustveno razgibanost; **taka ločila so osebna in zanje ni splošno veljavnih pravil.**

To pomeni, da je bila »stava ločil, zlasti vejice, še v SP 1950 utemeljena predvsem s skladenjskim pomensko-logičnim razmerjem« (Dobrovoljc 2004: 189), v SP 1962 pa je bilo v zametkih »upoštevano tudi stavčnofonetično merilo oz. tonski potek kot posledica oblikovanja osnovnih pomenskih enot« (prav tam), zato je v SP 1962 prišlo do izrazitega posega v stavo ločil – od uvajalnega odstavka v rabo ločil je ostal zgolj naslednji zapis:

Modulacijo glasu in presledke, s katerimi v govoru ločimo pomenske enote, zaznamujemo v pisavi z ločili. Potrebna so za lažje in zanesljivo razumevanje pisanega jezika, zato so določena s pravopisnimi pravili (SP 1962: 81).

Z vidika obravnave vejice se aktualni slovenski kodifikacijski priročnik (SP 2001) nekoliko razlikuje, saj kot prvi uvaja skladenjsko in neskladenjsko rabo ločil. Poleg vzpostavljenega razmerja pomen – oblika je z novim slovenskim pravopisom (oz. z načrtom zanj) v slovensko tradicijo prišlo tudi upoštevanje okoliščin oz. sotvarja (Dobrovoljc 2004: 189).

Na področju analiz spletnega jezika se obstoječe študije v veliki meri nanašajo na jezik v novih medijih v razmerju do jezikovnega standarda v tradicionalnih, prim. npr. Jakop (2008). Ugotovitev »večine razprav, ki se skoraj praviloma sklicujejo tudi na ugotovitve tujega jezikoslovja, je, da je jezik, uporabljen v novomedijskih besedilih, slogovno poln posebnosti, kot so kratkost sporočil, okrajšanost besed in besednih zvez ter povedi, izpustnost in neformalnost /.../, /v/endar pa je to ugotovitev mogoče podkrepiti le z jezikovno analizo ustreznih besedil« (Dobrovoljc 2008: 297). Za slovensko tradicijo je v odnosu do slovenščine v novih medijih značilno tudi nerazumevanje, da gre za *nove* medije in da obstoječega normativnega modela, ki temelji na knjižnojezikovni normi, pogosto ni mogoče aplicirati na *spletno slovenščino*, to nerazumevanje pa pogosto spremlja tudi splošno negativno vrednotenje jezika v tovrstnih besedilih, prim. npr. analizo spletnih forumov (Jakop 2008: 326):

Raziskava je pokazala, da je jezikovna podoba slovenskih spletnih forumov z vidika upoštevanja obstoječe pravopisne norme na zelo nizki ravni. Opazna so kršenja pravil o rabi velikih in malih črk, ločil in pisanja besed skupaj ali narazen. Pisanje prevzetih besed na spletnih forumih pa kaže spodbudne težnje jezikovnih uporabnikov po zapisovanju prevzetega besedja v podomačeni obliki.

Pri teh težnjah bi težko domnevali, da gre za prizadevanje piscev za »dobro slovenščino«, temveč za drugačne komunikacijske okoliščine, s katerimi jezikovni standard, temelječ na knjižnojezikovni normi, nima prav dosti skupnega oz. za njimi v marsičem zaostaja. Razlogi za tovrstne odmike od standardne slovenščine so odvisni tudi (ali predvsem) od prenosnika, ne le od jezikovnih ali metajezikovnih dejavnikov, ki se po navadi znajdejo na tnalu kot počelo nestandardnosti slovenščine v novih medijih, prim. npr. Jakop (2008: 326): »Med jezikovnimi dejavniki so v ospredju nepoznavanje oziroma slabo razumevanje nekaterih poglavij iz slovenskega pravopisa, zlasti pravil o rabi ločil in pisanju besed skupaj ali narazen.«

Pomislek, da lahko v novomedijskih besedilih najdemo pravopisne napake, zlahka zvrnemo tudi na kateri koli segment ubesedovanja v slovenščini, težave z obvladovanjem jezikovnega standarda pa so akutne tako pri poklicnih piscah (glej Popič 2014) kot tudi pri šolarjih (Kosem idr. 2012; Popič in Logar 2015), vendarle pa bi le težka pričakovali, da se bodo uporabniki novih medijev »kultivirali«, predvsem v smislu, kot ga navaja Jakop (2008: 327):

Morda smo preveč optimistični, vendar verjamemo, da z izobraževanjem mladih glede sporazumevanja na spletu dolgoročno prispevamo k višji kultiviranosti tvorcev besedil v spletnem diskurzu in s tem tudi k bolj kultiviranemu dialogu nasploh. Vse skupaj v duhu pregovora: Kar se Janezek nauči, to Janez zna.

Malo verjetno je, da se bodo uporabniki v spletnem okolju obnašali podobno kot v drugih, formaliziranih okoljih, saj so ta okolja pač neformal(izira)na in jim je zato neformalno izražanje inherentno. Ravno to je eden od ciljev pričujočega prispevka – preveriti, ali kljub težnji uporabnikov, da ločila uporabljajo, vendarle obstaja tudi močna tendenca po manj standardiziranem izražanju. Prav zaradi tega ima jezikoslovna stroka toliko težav že pri poimenovanju jezika v novih medijih, saj ga ni mogoče opisati z obstoječo socialno-funkcijsko stratifikacijo. Računalniško posredovana komunikacija lahko namreč zaseda katero koli mesto v tej stratifikaciji, od izrazito privzdignjene govornice do povsem neartikuliranega, profanega izražanja.

3 OPIS STAVE VEJICE V SLOVENSKI RAČUNALNIŠKO POSREDOVANI KOMUNIKACIJI

V pričujočem poglavju predstavimo procesa označevanja standardne in nestandardne stave vejic v našem korpusu. Delo je potekalo v orodju WebAnno (Eckart de Castilho et al. 2014), za označevanje pa smo uporabili tipologijo, predstavljeno v Tabeli 1. V nadaljevanju najprej predstavimo zasnovo in izpopolnjevanje tipologije za označevanje rabe vejic, zatem pa opišemo še proces označevanja in delotok.

3.1 Tipologija

Za opis rabe vejice smo razvili tipologijo, pri tem pa smo se oprli na aktualni jezikovni predpis (Slovenski pravopis 2001, Pravila), z nekaj (minimalnimi) dopolnitvami in prilagoditvami.⁵ Pri zamejevanju kategorij in definiranju tega, kaj naj zajemajo, smo se poskušali v čim večji meri držati jezikovnega predpisa, za podporo sorodnim raziskavam in vpogled v sistem označevanja pri naših raziskavah smo smernice za označevanje objavili na spletu.⁶ Pri sestavi kategorizacije smo se držali načela, da naj bo ta čim bolj univerzalna (da torej lahko s čim bolj podobnimi kategorijami opiše čim več primerov tako standardne kot tudi nestandardne rabe vejice, tj. odvečne vejice, manjkajoče vejice itd.). To ni bilo povsem

5 Osnovne značilnosti tipologije, vezane na nestandardno stavo vejice, so podane v Popič in Fišer (2015) in Popič et al. (2016).

6 Smernice za označevanje so na voljo na spletni strani <http://nl.ijs.si/janes/wp-content/uploads/2014/09/vejice-smernice.pdf>.

izvedljivo, zato tipologija vsebuje tudi kategorije, vezane predvsem na specifične primere (tako denimo kategorija Stavčni člen opisuje predvsem nestandardno »stavčnočlensko« rabo vejice, ki je vezana zgolj na odvečno vejico), povečini pa so kategorije vendarle univerzalne in namenjene opisu katerega koli segmenta jezika – tipologija torej ni namenjena zgolj opisu (ne)standardne stave vejice v spletni rabi, temveč je bil namen sestavljavcev prav ta, da lahko zadosti opisu katerega koli žanra ali besedilnega tipa.

Poleg jezikovnosistemskega opisa stave vejice smo se pri pripravi tipologije oprli tudi na že opravljene empirične, izgradivne preglede stave vejice v različnih okoljih, in sicer v spletnem (Popič in Fišer 2015), šolskem (Logar in Popič 2015, Kossem et al. 2012) ter v okolju poklicnih piscev (Popič 2014). Na podlagi pridobljenih spoznanj smo lahko v določeni meri predvideli problematična mesta pri stavi vejice in pripravili začetno različico tipologije. Z njo smo označili testni nabor tvitov in na podlagi sprotnih opažanj tipologijo dopolnjevali. Glavne razlike med začetno in končno tipologijo so se nanašale predvsem na nestandardno odvečno vejico, pri kateri tipična razdelitev glede na jezikovnosistemski opis odpove, saj temelji na skladenjski razčlenitvi – odvečnost vejice pa temelji ravno na tem, da uporabnik skladenjske razčlenitve ne (pre)pozna. Glede na to, da slovenski jezikovni predpis daje specifične napotke o tem, kje vejica mora stati, manj pozornosti pa namenja temu, kje vejice ne bi smelo biti, je za predvidevanje odvečnih vejic v okviru tipologije za označevanje potrebna predhodna empirična raziskava.⁷

Kot prikazuje Tabela 1, smo pri vseh segmentih skladenjske rabe vejice uvedli še podkategoriji levo- in desnosmerne vejice (Korošec 2003), saj v tem prepoznavamo različne (potencialne) vzgibe in razloge za nestandardno stavo vejice, v vsakem primeru pa smo želeli vsako posamezno stavo vejice opisati kar najnatančneje.

Kot lahko vidimo, kategorije v veliki meri temeljijo na razdelitvi, podani v jezikovnem predpisu, dodani sta kategoriji Stavčni člen in Besedna zveza, pri neskladenjski rabi vejice pa smo želeli čim podrobneje razčleniti, za katero vrsto nestandardne stave gre, zato so vključene še štiri podkategorije. Za neskladenjsko rabo vejice je rezerviran tudi parameter »napačen znak«, saj v tovrstnih primerih kategoriji manjkajoče in odvečne vejice nista (nujno) relevantni (npr. nestandardna raba pike namesto vejice). V določeni meri smo jezikovnosistemske kategorije tudi združevali (glej predvsem kategorijo 1.6), in sicer na mestih, kjer drobljenje informacij po naših pričakovanjih ne bi bistveno izboljšalo informacij o nestandardni rabi vejice, bi pa močno povečalo število kategorij. Pri združevanju kategorij smo bili pozorni predvsem na to, da smo združevali »skladdenjsko sorodne« kategorije, tako da so informacije, ki jih z analizo stave vejice dobimo, čim bolj posplošljive na vse elemente znotraj kategorije.

⁷ Verjetno je utemeljeno pričakovati, da bo slovenska preskripcija v prihodnje zajemala tudi napotke o najpogostejših napakah, osnovane na empiričnem pregledu gradiva.

Tabela 1: Pregled skladijskih in neskladijskih oznak po kategorijah.

1 Skladijska	1.2.5 Načinovni
1.1 Priredje	1.2.5.1 Desnosmerna
1.1.1 Vežalno	1.2.5.2 Levosmerna
1.1.1.1 Desnosmerna	1.2.6 Vzročni
1.1.1.2 Levosmerna	1.2.6.1 Desnosmerna
1.1.2 Stopnjevalno	1.2.6.2 Levosmerna
1.1.2.1 Desnosmerna	1.2.7 Namerni
1.1.2.2 Levosmerna	1.2.7.1 Desnosmerna
1.1.3 Ločno	1.2.7.2 Levosmerna
1.1.3.1 Desnosmerna	1.2.8 Pogojni
1.1.3.2 Levosmerna	1.2.8.1 Desnosmerna
1.1.4 Protivno	1.2.8.2 Levosmerna
1.1.4.1 Desnosmerna	1.2.9 Dopustni
1.1.4.2 Levosmerna	1.2.9.1 Desnosmerna
1.1.5 Vzročno	1.2.9.2 Levosmerna
1.1.5.1 Desnosmerna	1.2.10 Prilastkov
1.1.5.2 Levosmerna	1.2.10.1 Desnosmerna
1.1.6 Posledično	1.2.10.2 Levosmerna
1.1.6.1 Desnosmerna	1.3 Polstavek
1.1.6.2 Levosmerna	1.3.1 Desnosmerna
1.1.7 Pojasnjevalno	1.3.2 Levosmerna
1.1.7.1 Desnosmerna	1.4 Stavčni člen
1.1.7.2 Levosmerna	1.5 Besedna zveza
1.2 Odvisnik	1.6 Pri-, pa- in dostavek ter izpostavek⁸
1.2.1 Osebkov	1.6.1 Desnosmerna
1.2.1.1 Desnosmerna	1.6.2 Levosmerna
1.2.1.2 Levosmerna	1.7 Soredje
1.2.2 Predmetni	2 Neskladijska
1.2.2.1 Desnosmerna	2.1 X → vejica
1.2.2.2 Levosmerna	2.2 Vejica → X
1.2.3 Krajevni	2.3 Tipkarska napaka
1.2.3.1 Desnosmerna	2.4 Drugo
1.2.3.2 Levosmerna	
1.2.4 Časovni	
1.2.4.1 Desnosmerna	
1.2.4.2 Levosmerna	

8 Kategorija je bila zastavljena kot zbir formalno sorodnih skladijskih podkategorij, ki bi posamezno močno povečale število skupnih kategorij, a brez posebnega doprinosa k razumevanju (ne)standardne stave vejice. Nekoliko problematično je združevanje različnih kategorij, ki denimo ne morejo biti desno- ali levosmerne, vendarle pa je sama *smernost* v tovrstnih primerih izjemno koristna informacija, prepričani smo, da v večji meri kot podatek o tipu strukture.

Izdelana klasifikacija se je izkazala za uporabno, tako da je za označevanje standardne stave vejice nismo spreminjali, seveda pa to pomeni, da so bile določene kategorije, uvedene predvsem za opisovanje napačne stave vejice, tokrat izpuščene. Tudi pri označevanju nestandardne stave se je klasifikacija izkazala za uporabno, kar je bilo tudi pričakovano, saj temelji na jezikovnem predpisu (ravno tako kot tudi standardna stava). V tem oziru lahko rečemo, da je klasifikacija univerzalna, da pa terja razširitev, predvsem s podrobnejšimi jezikovnosistemskimi informacijami (denimo besednovrstnimi), če bi želeli stavo vejic (zlasti nestandardno) opisati natančneje.

3.2 Označevanje

Med procesom označevanja je bilo dvojno označenih 500 tvitov, in sicer smo najprej označili nestandardno stavo vejic, rezultate (glej Popič in Fišer 2015 ter Popič et al. 2016) pa smo želeli obogatiti dopolniti še s podatki o tem, katera mesta pri stavi vejice so neproblematična, zato smo dvojno označili še vse ustrezno stavljene vejice. To smo izvedli tako, da smo dvojno označili v istem naboru 500 tvitov (z vsebovanimi oznakami po prvem sklopu označevanja).

Pri označevanju je bilo mogoče za posamezni primer izbrati le eno kategorijo, večkategorialnih oznak – npr. ko je utemeljitev za stavo vejice lahko tako levokot desnosmerna – nismo dopuščali. Če je bilo mogočih več interpretacij, smo izbrali kategorijo, ki je bila v danem kontekstu bolj povedna ali pa verjetnejši razlog za (ne)stavo vejice. Po koncu označevanja (ne)stave vejice je kuratorka pregledala vseh 500 dvojno označenih tvitov in pri neskladjih med označevalcema sprejela dokončno odločitev. Po podatkih orodja WebAnno je Cohenov koeficient ujemanja med označevalcema kappa znašal 0,57, kar pomeni srednje dobro ujemanje. Zelo dobrega ujemanja zaradi težavnosti problema, sprotnega vzpostavljanja tipologije in smernic ter neizkušenosti označevalcev ni bilo mogoče pričakovati. Neskladja med označevalcema so bila najpogostejša predvsem pri naslednjih segmentih:

- ločevanje posameznih medmetov znotraj pastavka;
- prepoznavanje vezalnosti, protivnosti in posledičnosti veznikov *pa* ter *in*;
- obravnava smeškov in drugih metajezikovnih struktur, tipičnih za uporabniške vsebine;
- prepoznavanje vrst odvisnikov, zlasti v primerih, pri katerih so vezniške strukture zapisane nestandardno.

Pri prvih treh problematičnih mestih smo se odločili za princip minimalne intervencije, zato se je kuratorica izogibala posegov v izvorno stavo vejic, razen če je poseg nedvoumno predvidevalo določeno pravopisno pravilo ali kontekst, medtem ko je bila pri zadnjem problematičnem torišču potrebna vsebinska odločitev. Za čim ustrežnejšo obravnavo težavnih segmentov smo sprejeli natančne smernice, s čimer bo označevanje v prihodnje enostavnejše in bolj usklajeno, s tem pa smo poskrbeli tudi za sledljivost odločitev, ki smo jih sprejemali med procesom označevanja. Skupno je bilo pri 500 tvitih označenih 405 mest z nestandardno rabo vejice, kar pomeni v povprečju 0,8 oznake na posamezni tvit.

Da je bilo v tvitih več ustreznih vejic kot neustreznih, kaže že podatek o številu oznak za standardno rabo vejic, in sicer je bilo označenih 628 mest, tj. 1,26 oznake (in s tem ustrezno stavljene vejice) na tvit. Tudi pri označevanju standardne stave vejice smo naleteli na nekaj težav, a bistveno manj kot pri označevanju nestandardne stave; težave pa so bile vezane predvsem na prepoznavanje odvisnikov v primerih, ko je bil zapis (tudi samih veznikov) izrazito nestandarden ali pa je bilo treba določene dele besedila preprosto inferirati. Po zaključenem postopku označevanja je kurator⁹ pri neskladjih med označevalcema sprejel končno odločitev, pri čemer je bila večina neskladij vezana na identifikacijo posamezne jezikovne enote, v določenih primerih pa je zaradi zgoščenosti oznak znotraj označevalskega vmesnika prišlo tudi do pomot.

4 ANALIZA STAVE VEJICE V OZNAČENEM KORPUSU TVITOV

V nadaljevanju predstavimo izsledke analize nestandardne rabe vejic v tvitih. Najprej predstavljamo kumulativne rezultate po posameznih kategorijah, ki jih podajamo v Tabeli 2, v nadaljevanju pa se osredotočimo na posamezne (pod)kategorije, zlasti tiste, ki so glede na rezultate empirične raziskave v uporabniških spletnih vsebinah še posebno problematične.

4.1 Nestandardna stava

V Tabeli 2 so podane frekvence oznak za nestandardno stavo vejic po posameznih kategorijah.

⁹ Z vidika metodologije je nujno izpostaviti, da označevalska ekipa v obeh postopkih ni bila (povsem) enaka, prav tako se je zamenjal kurator, oboje iz drugotnih, pragmatičnih razlogov.

Tabela 2: Pregled skladijskih in neskladijskih oznak po kategorijah. Pri nadkategorijah so podani deleži oznak znotraj skladijskega in neskladijskega segmenta.

Kategorija	Odveč	Manjka	Nap. znak
1 SKLADENJSKA	19 (4,7) %	382 (95,3) %	0
1.1 Priredja	8 (2) %	35 (8,7) %	0
1.1.1 Vezalno	6	8	0
1.1.2 Stopnjevalno	0	2	0
1.1.3 Ločno	2	0	0
1.1.4 Protivno	0	13	0
1.1.5 Vzročno	0	1	0
1.1.6 Posledično	0	10	0
1.1.7 Pojasnjevalno	0	1	0
1.2 Odvisniki	2 (0,5) %	230 (57,4) %	0
1.2.1 Osebkov	0	26	0
1.2.2 Predmetni	1	89	0
1.2.3 Krajevni	0	1	0
1.2.4 Časovni	0	20	0
1.2.5 Načinovni	1	11	0
1.2.6 Vzročni	0	16	0
1.2.7 Namerni	0	0	0
1.2.8 Pogojni	0	25	0
1.2.9 Dopustni	0	1	0
1.2.10 Prilastkov	0	41	0
1.3 Polstavek	0	2 (0,5) %	0
1.4 Stavčni člen	3 (0,7) %	0	0
1.5 Besedna zveza	6 (1,5) %	0	0
1.6 Pa-, pri- do- in izpostavek	0	115 (28,7) %	0
2 NESKLADENJSKA	0	3 (75) %	1 (25) %

Kot lahko razberemo iz Tabele 2, so oznake med kategorijami razporejene zelo heterogeno, zelo neenakomerno pa so posejane tudi oznake glede na vrsto skladijskega razmerja. Ta heterogenost je značilna za obe ravni tehnične standardnosti tvitov,¹⁰ ki sta zajeti v analiziranih podatkih, je pa treba omeniti, da je v tvitih, ki so tehnično nestandardni, vejica stavljena bistveno manj standardno (v tvitih z oznako T1L3 lahko denimo najdemo 118 manjkajočih vejic, v tvitih T3L3 pa 266 manjkajočih vejic). To lahko pripišemo tudi motiviranosti uporabnikov za stavljenje ločil, saj je pri odvečnih vejicah razmerje precej bolj

¹⁰ V korpusu Janes imajo vsa besedila pripisano stopnjo tehnične (zapis presledkov, ločil, velikih in malih začetnic) in jezikovne standardnosti (zapis besed, raba nestandardnega besedišča). Uporabljena je trostopenjska lestvica, pri čemer 1 označuje zelo standardna, 3 pa zelo nestandardna besedila (glej Ljubešič et al. 2018).

enakovredno – celo obratno (12 odvečnih vejic v tehnično standardnih tvitih, 7 pa v tehnično nestandardnih tvitih).

Kot lahko vidimo, je bilo v 500 tvitih zanemarljivo malo (4) primerov napačno stavljene neskladenjske vejice, prav tako pa se je pri izbranem vzorcu pokazalo, da je v uporabniških spletnih vsebinah skorajda zanemarljiv pojav odvečna vejica (19 primerov), saj je velika večina (382 primerov) oznak vezana na manjkajočo vejico. Pri tem je treba poudariti, da zaradi majhnega vzorca te ugotovitve ni mogoče sploševati. Velika večina popravkov se nanaša na skladenjsko rabo vejice (401 od skupno 405 primerov oz. 99 %). Glede na rezultate empirične analize lahko vidimo, da sta še posebno akutni kategoriji manjkajočih vejic pri odvisnikih (230 primerov oz. 57 % vseh skladenjskih oznak) in pastavčnih strukturah (115 primerov oz. 28,7 % vseh skladenjskih oznak).

4.1.1 Manjkajoča vejica

Kot lahko vidimo v Tabeli 2, je stava vejice z vidika manjkajoče vejice problematična predvsem pri odvisniških in pristavčnih strukturah, zato se v nadaljevanju nekoliko podrobneje posvetimo prav tema kategorijama. V Tabeli 3 je podan prikaz najpogostejših odvisniških vrst z označeno manjkajočo vejico.

Tabela 3: Pregled manjkajočih vejic pri odvisnikih po frekvenci in deležu znotraj skladenjske kategorije.

Odvisnik	Število oznak
Predmetni (D)	80 (19,8 %)
Prilastkov (D)	29 (7,1 %)
Osebkov (D)	24 (5,9 %)
Vzročni (D)	16 (4,0 %)
Pogojni (D)	13 (3,2 %)
Časovni (D)	13 (3,2 %)
Pogojni (L)	12 (2,9 %)
Prilastkov (L)	12 (2,9 %)
Načinovni (D)	10 (2,5 %)
Predmetni (L)	9 (2,2 %)
Časovni (L)	7 (1,7 %)
Osebkov (L)	2 (0,5 %)
Načinovni (L)	1 (0,2 %)
Krajevni (D)	1 (0,2 %)
Dopustni (D)	1 (0,2 %)

Kot lahko vidimo, je od skupno 230 manjkajočih vejic pri odvisnikih edina zares problematična kategorija desnosmerne vejice pri predmetnih odvisnikih, pri katerih najdemo več kot tretjino primerov z manjkajočo vejico. Tudi na splošno se je izkazalo, da je desnosmerna vejica precej bolj problematična kot levosmerna, saj zaseda šest najpogostejših mest z manjkajočo vejico (175 od skupno 230 manjkajočih vejic). Edini odvisniški vrsti, pri katerih smo zaznali nekaj več manjkajočih levosmernih vejic, sta kategoriji prilastkovih in vzročnih odvisnikov. Za slednje lahko sklepamo, da pogosto začenjajo povedi, pri prilastkovih odvisnikih pa uporabniki pogosto najverjetneje pozabijo skleniti oziralni odvisnik.

Tabela 4: Pregled oznak pri pa-, pri-, do- in izpostavnih strukturah po frekvenci in deležu skladijskih oznak.

	Odvečna	Manjkajoča
Levosmerna	/	79 (19,7 %)
Desnosmerna	/	36 (9,0 %)

Pri pristavkih, dostavkih, izpostavkih in pastavkih je problematična predvsem vejica na začetku, relativno pogosta pa je tudi nestandardna stava brez vejice na koncu stavka. V številnih primerih gre za nestandardno stavo vejice pri denimo medmetih (npr. hahaha), ki jih – tako domnevamo – uporabniki spletnih omrežij ne dojemajo kot pastavčne tvorbe, temveč kot metajezikovno pojavnost oz. verbalizirane emotikone. Zelo pogosta je tudi pojavnost nestandardne rabe vejice pri členkovnih pastavkih na začetku in koncu stavka, zlasti v navezavi s sklicem na uporabniško ime. Tovrstne rabe v korpusu nismo označevali kot odmik od standarda, predvsem zaradi pogostosti in zaradi tehničnih zahtev Twitterja, saj lahko stavljenje vejice v sklice povzroči težave s sklicem. V vsakem primeru pa način uporabe sklicev potrjuje domnevo, da uporabniki izražajo metajezikovne prvine z uporabo jezikovnih sredstev.

4.1.2 Odvečna vejica

Vseh primerov odvečne vejice od skupno 401 skladijske oznake v zbirki podatkov je bilo 19, kar pomeni manj kot 5 odstotkov vseh oznak, njihova porazdelitev po kategorijah pa je naslednja:

Tabela 5: Pregled odvečnih vejic glede na kategorijo po frekvenci in deležu skladijskih oznak.

Oznaka	Število
Vezalno priredje	6
Besedna zveza	6
Stavčni člen	3
Ločno priredje	2
Predmetni odvisnik	1
Načinovni odvisnik	1

Kot lahko vidimo, je bila vejica največkrat odvečna v besednih zvezah (denimo med sestavljenimi vezniki), med strukturami v vezalnem ali ločnem priredju, za stavčnim členom in med enakovrednima odvisnikoma. Sodeč po skupnem številu označenih nestandardnih mest, bi pričakovali večje število odvečnih vejic, predvsem stavčnočlenskih, zlasti glede na raziskave na standardnojezikovnem gradivu (prim. Popič 2014). Preden lahko zaključimo, da je neskladenjska raba vejice pretežno manjkajoča, ker je to ena od strategij krajšanja sporočil v tovrstnem načinu komuniciranja in ker je vejico psihološko in tehnično lažje izpustiti kot pa napisati po nepotrebnem, bi bilo nujno treba opraviti analizo na večjem vzorcu besedil tega žanra.

4.1.3. Neskladenjska vejica

Kot smo pokazali v predhodnih poglavjih, so bili v celotnem naboru podatkov označeni zgolj štirje primeri nestandardne neskladenjske vejice, ki so se pojavili v treh tvitih. Te izpisujemo v celoti v Tabeli 6:

Tabela 6: Pregled primerov z označeno nestandardno rabljeno neskladenjsko vejico.

Setamo po Lj in pride Zoki mim, se ustavi pa da Emanuelu petko. Tamalmu nc jasn. Recem(,) to je Zoki kralj(,) in se tamal zadere: Zoki kralj! :)

@user Ja bi mogla tud jst naletet na dobrega :) ampak ima res ogromno folka sfaljenega (ne ostri do 2.8). Je, kar je.

@user »Maybe yes, maybe no, maybe go home(,)« so radi rekli Šerpe na odpravi na Annapurno, kadar smo jih vprašali, če bo vreme OK

Kot lahko vidimo, so od tega tri vejice v resnici skladijske, in sicer gre za tri primere soredja (1. in 3. tvit; v oklepaju so podane manjkajoče vejice na mestih, ki so bila označena tudi med označevanjem), ki ga namenoma nismo vključili v tipologijo, saj v tem mediju nismo pričakovali primerov premega govora. Pri drugem primeru (krepko je tiskana pojavnica, ki je bila označena med označevanjem) pa je dejansko neuporabljena neskladijska vejica, vendar lahko v tem primeru oznako problematiziramo, saj gre – kolikor lahko razberemo iz sobesedila – za navedbo odprtosti zaslonke foto-grafske leče, ki se tipično podaja s piko. Podobno kot pri nekaterih drugih kategorijah smo glede na dosedanje raziskave, izvedene na standardnem gradivu (glej npr. Popič 2014), pričakovali bistveno več primerov nestandardne rabe neskladijske vejice.

4.2 Standardna stave vejice

Kot prikazuje Tabela 7, lahko 628 oznak za standardno stavljeno vejico v korpusu tvtov razporedimo v 19 kategorij, in medtem ko so oznake povečini homogeno razporejene po korpusu, lahko vidimo, da štiri kategorije izrazito izstopajo, in sicer gre za (levosmerno) pristavčno vejico ter (desnosmerno) vejico pri vezalnem priredju, predmetnem odvisniku in protivnem priredju.

Tabela 7: Pregled oznak standardne stave vejice v korpusu Janes-Vejica.

Kategorija	Št. oznak	
Pa-, pri- do- in izpostavek (L)	127	20,22%
Vežalno priredje (D)	99	15,76%
Predmetni odvisnik (D)	88	14,01%
Protivno priredje (D)	70	11,15%
Prilastkov odvisnik (D)	34	5,41%
Pa-, pri- do- in izpostavek (D)	31	4,94%
Pojasnjevalno priredje (D)	28	4,46%
Vzročni odvisnik (D)	24	3,82%
Pogojni odvisnik (D)	23	3,66%
Načinovni odvisnik (D)	18	2,87%
Časovni odvisnik (D)	17	2,71%
Pogojni odvisnik (L)	14	2,23%
Posledično priredje (D)	14	2,23%
Namerni odvisnik (D)	10	1,59%
Časovni odvisnik (L)	8	1,27%
Osebkov odvisnik (D)	8	1,27%
Prilastkov odvisnik (L)	5	0,80%
Stopnjevalno priredje (D)	5	0,80%
Predmetni odvisnik (L)	5	0,80%

Kot lahko vidimo, je kategorija z najvišjo frekvenco pri standardni stavi vejice obenem tudi ena najbolj problematičnih pri nestandardni stavi, s primerljivo frekvenco (127 : 115), kar pomeni, da je kljub zelo frekventni standardni stavi vejice pri pri-, pol-, do- in izpostavnih strukturah v polovici primerov pri tovrstnih strukturah stava vejice še vedno nestandardna. Če združimo levo- in desnosmerno standardno stavo vejic pri tovrstnih strukturah (158 oz. več kot četrtnina vseh standardno stavljenih vejic), vidimo, da je to relativno akutna zapisovalna težava, saj je 42 % vseh vejic stavljenih nepravilno (v vseh primerih gre za manjkajočo vejico), vendarle pa drži, da je zgolj 30 primerov nestandardne vejice.

Vežalno priredje zaradi svoje narave v normativnem smislu ni problematično, veliko težav pa povzročajo predmetni odvisniki. V označenem korpusu lahko tako najdemo 88 primerov standardne stave (desnosmerne) vejice pri predmetnih odvisnikih, medtem ko je primerov manjkajoče (desnosmerne) vejice celo več (90), kar nakazuje na to, da gre pri tovrstnih odvisnikih za pereč problem. To lahko v večji ali manjši meri pripišemo tudi dejstvu, da gre za pogost pojav v jeziku, obenem pa ugotovitve pritrjujejo dosedanjim raziskavam, ki kažejo, da uporabniki vejico pred predmetnim odvisnikom izpuščajo, zlasti v primerih, ko je predmetni veznik uveden z manj pogostim (ali formaliziranim) veznikom. Na to nakazujejo tudi frekvence posameznih pojavnic, pred katerimi v korpusu Janes-Vejica najdemo največ oznak za standardno stavo vejic in ki so podane v Tabeli 8 (zajete so pojavnice z vsaj 10 oznakami v korpusu).

Tabela 8: Frekvenca pojavnic za standardno stavljenimi vejicami v korpusu Janes-Vejica.

Pojavnica	Št. oznak
da	105
pa	31
če	27
ko	20
ne	16
ki	16
ampak	16
ker	14
je	12

Kot lahko vidimo, gre pri vseh pojavnicah za izrazito predvidljive slovnične besede, pri katerih raba vejice načeloma ne omahuje in ne peša. Seveda pa to ne pomeni, da je raba tudi vedno pravilna, zlasti v luči tega, da zaradi kompleksne sintakse vejica pred njimi pogosto mora biti izpuščena. Da pa bi lahko ugotovili, katere pojavnice so najbolj problematične, v Tabeli 9 podajamo pojavnice, pred

katerimi vejica najpogosteje umanjka (ponovno so zajete zgolj pojavnice, ki se v korpusu pojavijo vsaj 10-krat; tiste, ki jih najdemo med najpogostejšimi tudi pri standardni stavi vejice, so pisane krepko).

Tabela 9: Pojavnice z največ oznakami za manjkajočo vejico v korpusu Janes-Vejica.

Pojavnica	Št. oznak
da	71
k	26
pa	17
in	12
če	10
ker	10

Kot prikazuje Tabela 9, so najpogostejše oznake v veliki meri prekrivne, kar nakazuje na to, da je opuščanje vejic vsaj pred določenimi pojavniciami (denimo bazičnimi vezniki, kot so *da*, *ki*, *ko*, *če* in *ker*) zavestna odločitev, ne napaka. Vendar pa je pri tem treba omeniti, da lahko to sklepamo predvsem na podlagi raziskav jezikovne rabe poklicnih oz. izobraženih piscev, ne pa tudi šolarjev, kot prikazuje Tabela 10.

Tabela 10: Pojavnice z največ oznakami za manjkajočo vejico v korpusih Lektor in Šolar.

Lektor	Frekvenca	Šolar	Frekvenca
in	269	da	1524
kot	73	je	668
je	69	saj	585
ali	49	kako	394
ne	46	se	370
»	36	kar	325
da	28	kaj	312
v	26	kot	292
so	23	in	291
pa	22	ki	280
naj	22	ko	266
se	20	ker	222
«	19	naj	221
s	18	v	220
ter	16	vendar	192

Kot lahko vidimo, se težave z vejico pri šolarjih in poklicnih piscih in prevajalcih precej razlikujejo. Medtem ko rezultati iz Lektorja kažejo, da uporabniki vejico pozabijo/izpustijo v bolj zapletenih stavčnih strukturah, rezultati iz Šolarja kažejo, da šolarji vejico pogosto izpuščajo tudi pred povsem predvidljivimi in bazičnimi strukturami, kar je zaradi šolske narave seveda razumljivo, obenem pa lahko vidimo, da se najpogostejše pojavnice za izpuščenimi vejicami v bistveno večji meri ujemajo s podatki iz Šolarja kot iz Lektorja.

4 SKLEP

V prispevku smo predstavili tipologijo za označevanje rabe vejice v slovenščini in empirično raziskavo o rabi vejice, temelječo na korpusu Janes-Vejica. Temeljne rezultate raziskave lahko v grobem povzamemo z naslednjimi točkami:

- 1) Tudi v jezikovno nestandardni računalniško posredovani komunikaciji je standardna vejica pogostejša od nestandardne.
- 2) Točki 1 navkljub lahko vidimo, da uporabniki vejico pogosto tudi izpuščajo oz. ne stavijo pravilno, kot to velja za večino pisanja v slovenščini, ne glede na medij. Vsaj za določen del gradiva lahko utemeljeno trdimo, da gre za nameren izpust vejice, to pa je v spletnem okolju vsaj za del pisne produkcije tudi pričakovano in samoumevno.
- 3) Ko gre za nestandardno vejico, je bistvena manjkajoča vejica, odvečna vejica je v našem vzorcu malodane neproblematična.
- 4) Z vidika normativnosti imajo uporabniki težave predvsem s prepoznavanjem skladenjskih enot, kar bi moralo biti vodilo pri prihodnjem delu slovenske preskripcije. Sistem, opisan v pričujočem prispevku, se ponuja kot ustrezen za empirično raziskovanje jezikovne rabe (za širok razpon besedilnih zvrsti).
- 5) Tipologija, razvita za namene pričujoče raziskave, se je izkazala za ustrezno oz. vsaj zadostno za tovrstno raven jezikovnega opisa. Vsakršna globalna analiza bi terjala nadaljnje drobljenje kategorij.

Kot kažejo rezultati študije, je standardna vejica v jezikovno nestandardnih tvitih pogostejša od nestandardne. Pri tem seveda govorimo zgolj o frekvenci, kar pomeni, da je nujno izvesti še statistično-kvalitativno raziskavo, s katero bomo lahko natančno prikazali trende rabe, nerabe in neznanja uporabe vejice. Rezultati pričujoče raziskave prav tako potrjujejo preteklim analizam, da je nestandardna stava vejice v slovenščini precejšen problem, ne glede na medij ali besedilno zvrst, kar pomeni, da je tako z raziskovalnega kot tudi preskriptivističnega (najbrž pa tudi pedagoškega) vidika ta tematika potrebna čim obširnejše obravnave, po

možnosti temelječe na realnem in sodobnem gradivu. Pričujoča razprava predstavlja eno od možnosti, kako tovrstno analizo zastaviti.

Zahvala

Avtorja se za pomoč pri anotaciji, kuriranju in pripravi tipologije zahvaljujeva Katji Zupan, Poloni Logar in Teji Kavčič.

Literatura

- Arhar Holdt, Špela in Kaja Dobrovoljc, 2016: Vrednost korpusa Janes za slovensko normativistiko. *Slovenščina 2.0*, 4 (2): 1–37.
- Dobrovoljc, Helena, 2004: *Pravopisje na Slovenskem*. Ljubljana: Založba ZRC.
- Dobrovoljc, Helena, 2008: Jezik v e-poštnih sporočilih in vprašanja sodobne normativistike. Košuta, Miran (ur.): *Slovenščina med kulturami. Zbornik Slavističnega društva Slovenije 19*. Celovec: Slavistično društvo Slovenije. 295–314.
- Eckart de Castilho, Richard, Chris Biemann, Iryna Gurevych in Sied Muhie Yimam. 2014. WebAnno: a flexible, web-based annotation tool for CLARIN. *Proceedings of the CLARIN Annual Conference (CAC) 2014*. Soesterberg, Netherlands. https://www.clarin.eu/sites/default/files/cac2014_submission_6_0.pdf
- Erjavec, Tomaž, Ljubešič, Nikola in Darja Fišer, 2018: Korpus slovenskih spletnih uporabniških vsebin Janes. Fišer, Darja (ur.): *Viri, orodja in metode za analizo spletne slovenščine*. Ljubljana: Znanstvena založba Filozofske fakultete. 16–43.
- Fišer, Darja, Tomaž Erjavec, Ljubešič, Nikola, 2017: The compilation, processing and analysis of the Janes corpus of Slovene user-generated content. Ciara, R. Wigham in Gudrun Ledegen (ur.): *Corpus de communication médiée par les réseaux: construction, structuration, analyse*. Collection Humanités Numériques. Pariz: L'Harmattan (v tisku).
- Goli, Teja, Osrajnik, Eneja in Darja Fišer, 2016: Analiza krajšanja slovenskih sporočil na družbenem omrežju Twitter. *Proceedings of the Language Technologies and Digital Humanities Conference*. Ljubljana, Slovenia. 77–82.
- Jakop, Nataša, 2008: Pravopis in spletni forumi – kva dogaja? Košuta, Miran (ur.): *Slovenščina med kulturami. Zbornik Slavističnega društva Slovenije 19*. Celovec: Slavistično društvo Slovenije. 315–327.
- Korošec, Tomo, 2003: K pravilom za skladijsko vejico v Slovenskem pravopisu 2001. *Slavistična revija 5/2*: 247–266.
- Kosem, Iztok, Stritar, Mojca, Može, Sara, Zwitter Vitez, Ana, Arhar Holdt, Špela, Rozman, Tadeja, 2012: *Analiza jezikovnih težav učencev: korpusni pristop*. Ljubljana: Trojina, zavod za uporabno slovenistiko.
- Levec, Fran, 1899: *Slovenski pravopis*. Dunaj: Cesarska kraljeva zaloga šolskih knjig.

- Ljubešić, Nikola, Darja Fišer, Tomaž Erjavec, Čibej, Jaka, Marko, Dafne, Pollak, Senja in Škrjanec, Iza, 2015: Predicting the level of text standardness in user-generated content. *Proceedings of Recent Advances in Natural Language Processing*, 7.–9. september 2015. Hisar, Bolgarija. 371–378.
- Ljubešić, Nikola, Tomaž Erjavec in Darja Fišer, 2018: Orodja za procesiranje nestandardne slovenščine. Fišer, Darja (ur.): *Viri, orodja in metode za analizo spletne slovenščine*. Ljubljana: Znanstvena založba Filozofske fakultete v Ljubljani. 74–99.
- Logar Berginc, Nataša, Miha Grčar, Marko Brakus, Tomaž Erjavec, Špela Arhar Holdt in Simon Krek, 2012: *Korpusi slovenskega jezika Gigafida, KRES, cc-Gigafida in ccKRES: gradnja, vsebina, uporaba*. Ljubljana: Trojina, zavod za uporabno slovenistiko; Fakulteta za družbene vede.
- Logar, Nataša in Damjan Popič, 2015: Vejica: rezultati anketne raziskave med dijaki in študenti. *Jezikoslovni zapiski* 2/2. 45–59.
- Michelizza, Mija, 2014: Slovenščina v elektronskih medijih. *Razpotja* 15. <http://www.razpotja.si/slovenscina-v-elektronskih-medijih/>
- Nebeská, Iva, 1999: *Jazyk. Norma. Spisovnost*. Praga: Karolinum.
- Osrajnik, Eneja, Darja Fišer in Damjan Popič, 2015: Primerjava rabe ekspresivnih ločil v tvitih slovenskih uporabnikov in uporabnic. Fišer, Darja (ur.): *Zbornik konference Slovenščina na spletu in v novih medijih*. Ljubljana: Znanstvena založba Filozofske fakultete, 50–74.
- Popič, Damjan, 2014: *Korpusnojezikoslovna analiza vplivov na slovenska prevodna besedila*. Doktorska disertacija. Ljubljana: Filozofska fakulteta.
- Popič, Damjan in Darja Fišer, 2015: Vejica je mrtva, živela vejica. Žele, Andreja (ur.): *OBDOBJA 34: Slovnica in slovar – aktualni jezikovni opis*. Ljubljana: Znanstvena založba Filozofske fakultete. 609–618.
- Popič, Damjan in Nataša Logar, 2015: Med dvema ognjema: kje stoji vejica v slovenskih gimnazijah? *OBDOBJA 34: Slovnica in slovar – aktualni jezikovni opis*. Ljubljana: Znanstvena založba Filozofske fakultete. 619–627.
- Popič, Damjan, Darja Fišer, Katja Zupan in Polona Logar, 2016: Raba vejice v uporabniških spletnih vsebinah. Erjavec, Tomaž in Darja Fišer (ur.): *Zbornik konference Jezikovne tehnologije in digitalna humanistika*, 29. september–1. oktober 2016, Filozofska fakulteta, Univerza v Ljubljani, Ljubljana, Slovenija. 106–110.
- Popič, Damjan in Darja Fišer, 2017: Fear and Loathing on Twitter: Attitudes towards Language. *Proceedings of the 5th Conference on CMC and Social Media Corpora for the Humanities (cmccorpora 17)*, oktober 2017, Eurac Research, Italija. 61–64.
- Rozman, Tadeja, Irena Krapš Vodopivec, Mojca Stritar in Iztok Kosem, 2012: *Empirični pogled na pouk slovenskega jezika*. Ljubljana: Trojina, zavod za uporabno slovenistiko.
- Slovenski pravopis, 2001. Pravila. Ur. Toporišič, J. et al. Ljubljana: Založba ZRC, ZRC SAZU.
- Verovnik, Tina, 2003: Vejica premalo, vejica preveč (2). *Pravna praksa* 22/21. 51.
- Žibert, Živa, 2006: Slovenska vejica: balast ali skladišna nujnost slovenskega knjižnega jezika? Diplomsko delo. Ljubljana: Fakulteta za družbene vede.