



Smernice za označevanje z odvisnostnim sistemom JOS: nestandardna slovenščina

v1.0

Datum zadnje spremembe: 2016-12-21

Specifikacije za nestandardni jezik: Špela Arhar Holdt

Specifikacije SSJ: Peter Holozan, Simon Krek, Matej Pivec, Simon Rigač, Simon Rozman, Aleš Velušček

Specifikacije JOS: Špela Arhar, Tomaž Erjavec, Darja Fišer, Simon Krek, Nina Ledinek, Amanda Saksida, Breda Sivec, Blaž Trebar

Smernice so bile pripravljene v okviru nacionalnega temeljnega projekta "Viri, orodja in metode za raziskovanje nestandardne spletne slovenščine" (J6-6842, 2014-2017), ki ga financira ARRS.

Del smernic, ki se nanaša na standardni jezik, je delno financirala Evropska unija iz Evropskega socialnega sklada ter Ministrstvo za šolstvo in šport v okviru Operativnega programa razvoja človeških virov za obdobje 2007-2013, razvojne prioritete: razvoj človeških virov in vseživljenjskega učenja; prednostne usmeritve: izboljšanje kakovosti in učinkovitosti sistemov izobraževanja in usposabljanja 2007-2013.

Kazalo

0. Uvod	3
1. Sistem označevanja	4
2. Zasnova priročnika	4
3. Program za označevanje	5
4. Pred začetkom označevanja	7
5. Povezave.....	8
5.1 Povezave prvega nivoja.....	8
5.1.1 Povezava DOL	8
5.1.2 Povezava DEL.....	10
5.1.3 Povezavi PRIR in VEZ.....	11
5.1.4 Povezava SKUP	14
5.2 Povezave drugega nivoja	15
5.2.1 Povezava ENA	15
5.2.2 Povezava DVE	16
5.2.3 Povezava TRI.....	17
5.2.4 Povezava ŠTIRI	18
5.2.5 Polstavki.....	18
5.2.6 Prilastkovi odvisniki	19
5.2.7 Predmet ali prislovno določilo?.....	19
5.2.8 Uvajanje odvisnikov s kazalnimi zaimki	20
5.3 Povezave tretjega nivoja	20
5.4 Posebni primeri.....	25
5.4.1 Vezniki	25
5.4.2 Členek ali prislov?	25
5.4.3 Izražanje mnenja ali ocene	25
5.4.4 Sestavljene predložne zveze.....	25
5.4.5 Povedkova določila in povedkovi prilastki.....	26
5.4.6 Modalni glagoli	27
5.4.7 Idiomi.....	29
5.4.8 Polstavčni desni prilastki	29
5.4.9 Pojasnjevalne strukture.....	30
5.4.10 Primerjave	30
5.4.11 Premi govor	31
5.4.12 Vprašalnice	32
5.4.13 Pasivne konstrukcije	33
5.4.14 Členek NE.....	33
5.4.15 Slovnične napake v besedilih.....	34
5.5 Označevanje nestandardne slovenščine.....	35
5.5.1 Spletna mesta, emotikoni, sklici, ključniki.....	35
5.5.2 Tujejezični elementi.....	38
5.5.3 Elipse in fragmenti	42
5.5.4 Nestandardna raba ločil	43
5.5.5 Druge nestandardne jezikovne značilnosti.....	44
I. Priloga A – tabela bližnjic na tipkovnici	45
II. PRILOGA B – informativni seznam glagolov, ob katerih (ustrezne) samostalniške zveze označujemo z odnosom DVA, pridevniške pa z DOL.....	46
III. PRILOGA C – informativni seznam zvez, ki jih povezujemo z odnosom SKUP	47

0. Uvod

Pričujoče smernice dopolnjujejo obstoječe smernice za skladijsko označevanje slovenščine z odvisnostnim sistemom JOS (Holozan 2008).¹ Izhodiščne smernice so bile uporabljene za pripravo učnega korpusa ssj500k, ki je nastal v projektu **Sporazumevanje v slovenskem jeziku**² za učenje sistemov za avtomatsko obdelavo slovenščine. Smernice sledijo sistemu označevanja korpusov, razvitih v okviru projekta **Jezikoslovno označevanje slovenščine**.³

Za skladijski del označevanja je bil v sklopu omenjenih projektov razvit jezikovnospecifičen odvisnostni sistem, ki ga odlikuje visoka stopnja robustnosti in fleksibilnosti. Nabor oznak in sistem označevanja je prilagojen (statističnemu) označevanju obsežnih korpusov, zato lahko prihaja do precejšnjih razlik med slovničnimi opredelitvami v obstoječih jezikovnih virih (SSKJ, Slovenski pravopis 2001) in korpusnimi oznakami. Za skladijsko označevanje je uporabljeno orodje, razvito v okviru projekta JOS, avtor programa je mag. Janez Brank. Sistem je bil apliciran na podkorpusu slovenskega referenčnega korpusa Gigafida, kar pomeni pretežno standardno, v veliki meri tudi jezikovno korigirano (lektorirano) gradivo.

V sklopu projekta **Viri, orodja in metode za raziskovanje nestandardne spletne slovenščine** je bil v letu 2016 pripravljen nov učni korpus, namenjen učenju avtomatske obdelave nestandardne jezikovne rabe, kakršna se pojavlja v različnih besedilnih vrstah računalniško posredovane slovenščine (Čibej et al. 2016).⁴ Vzorec tovrstnega jezika, predhodno normaliziran na leksikalni ravni, je bil označen tudi skladijsko, na podlagi česar so bile smernice za označevanje ustrezno dopolnjene. Označeni dataset je dostopen v okviru repozitorija CLARIN.SI pod imenom »Janes-Syn«

Dopolnila in spremembe, ki se nanašajo na nestandardni jezik, so navedene **v rdečih okvirčkih**. Novo je tudi poglavje 5.5.

¹ Holozan, P., Krek, S., Pivec, M., Rigač, S., Rozman, S. and Velušček, A. (2008). *Specifikacije za učni korpus*. Kamnik: Projekt »Sporazumevanje v slovenskem jeziku« ESS in MŠŠ. Dostopno na: <http://www.slovenscina.eu/Vsebine/SI/Kazalniki/K2.aspx>.

² www.slovenscina.eu

³ <http://nl.ijs.si/jos/>

⁴ Čibej, J., Arhar Holdt, Š., Erjavec, T. and Fišer, D. (2016b). Razvoj učne množice za izboljšano označevanje spletnih besedil. V *Proceedings of the Conference on Language Technologies and Digital Humanities*. Ljubljana. 40-46. Dostopno na: http://www.sdjt.si/wp/wp-content/uploads/2016/09/JTDH-2016_Cibej-et-al_Razvoj-ucne-mnozice.pdf.

1. Sistem označevanja

Sistem za skladiščno označevanje slovenščine, nastal je v okvirih projektov Smart ter JOS, je namenjen ročnemu označevanju skladiščnih odnosov v slovenskih povedih. Zgrajen je bil iz jezika samega, v določeni meri sledeč smernicam uveljavljenega slovenskega jezikoslovja, obenem pa temeljnimi idejami, ki jih zarisujejo že obstoječi sistemi odvisnostnega označevanja.

Snovalci sistema smo imeli možnost zasnovati nabor skladiščnih oznak z upoštevanjem oblikoskladiščnih oznak, ki so pred skladiškim označevanjem besedam že pripisane. Sistem tako prinaša robusten nabor skladiščnih oznak, s katerimi ni težko operirati, bistvo sistema pa leži v kombiniranju teh oznak z že obstoječimi oblikoskladiščnimi.

Imena za povezave med deli stavka oz. povedi so v tej različici navodil za označevanje ostala v svoji delovni obliki, kar morda pomeni manjšo znanstvenost, po drugi strani pa najbrž tudi močnejšo asociativnost. Pri poimenovanju povezav smo se zavestno odločali za imena, ki so drugačna od tistih, ki jih prinaša uveljavljeno slovensko jezikoslovje, saj slednjemu ne sledimo v vseh pogledih in bi prenos terminologije zato vodil v neželeno ter nepotrebno poimenovalno zmedo.

2. Zasnova priročnika

Priročnik je namenjen predstavitvi označevalnega sistema označevalcem, ki ga bodo uporabljali. Na začetku je na kratko predstavljen program za označevanje, sledita razlaga oznak po nivojih ter opis rešitev za določene specifične skladiščne konstrukcije.

Razlaga sistema je seveda podkrepljena s primeri, označenimi v Označevalniku. Na mestu je opozorilo, da se označenost besednih zvez⁵ in okrepljenost posameznih besed oz. puščic v teh primerih pojavljajo zgolj kot pomoč pri ponazoritvi trenutno obravnavanega problema, da torej sami po sebi ne sledijo kakšni globlji označevalni logiki.

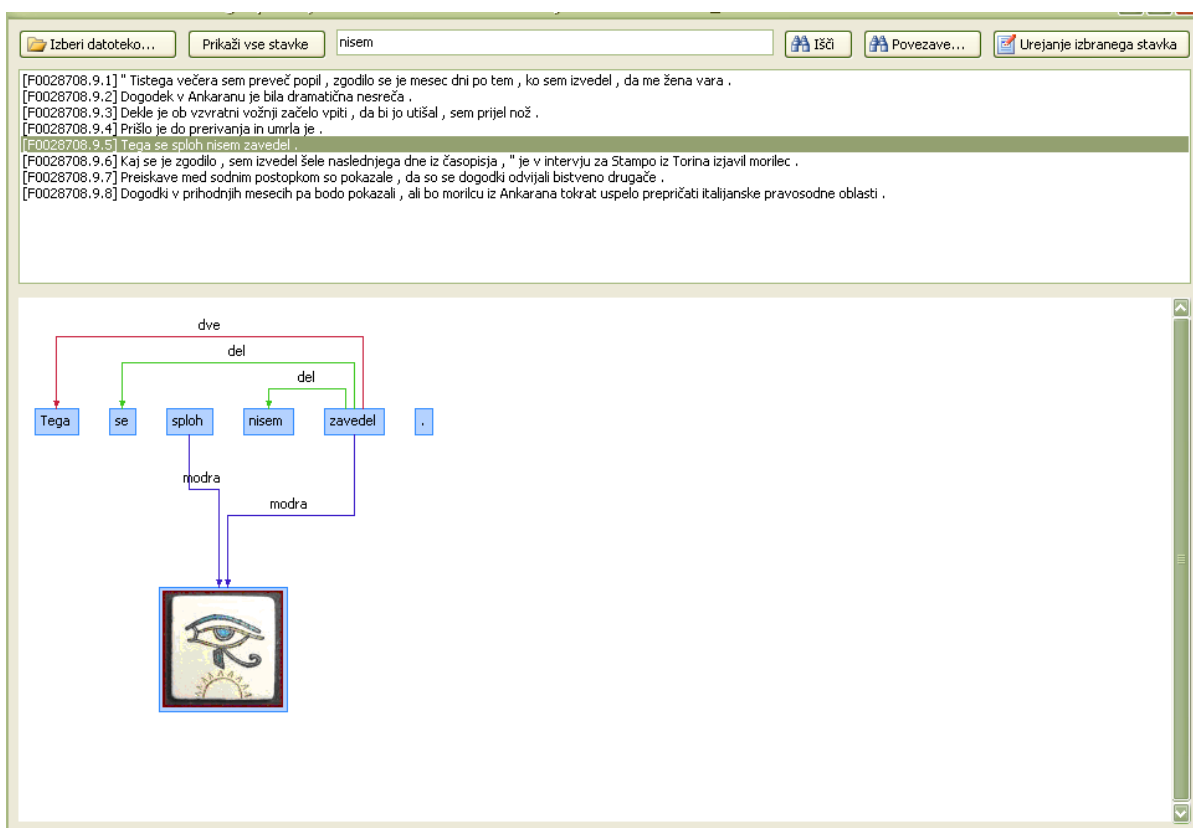
Dodatna literatura za označevalce je Slovenska slovnica (Toporišič 2004, v nadaljevanju SS), ki jo uporabljamo kot osvežitev znanja s področja jezikoslovnih kategorij tradicionalnega tipa. Primeri, ki jih slovnica navaja, so lahko v pomoč pri odločanju, kako obravnavati avtentične primere jezikovne rabe, ki jih označujemo. Potrebno pa je opozorilo, da je označevanje stavkov po sistemu usmerjeno namensko (cilj je izboljšana avtomatska obdelava naravnih jezikov) in kot takšno ugotovitve uveljavljenega jezikoslovja lahko upošteva le do določene mere. Glavno vodilo sistema je identifikacija ter označevanje tistih jezikoslovnih kategorij, ki dajejo za avtomatsko obdelavo relevantne informacije (oznake naj torej ne bi bile presplošne ali prepodrobne za ta namen – ohranjanje ustreznega nivoja specifičnosti oznak pa se odraža v razlikah med tem sistemom ter SS).

⁵Tj. svetlomodri kvadrček okrog besedne zveze, ki prinaša tudi kodo zveze (npr. s1c1).

3. Program za označevanje

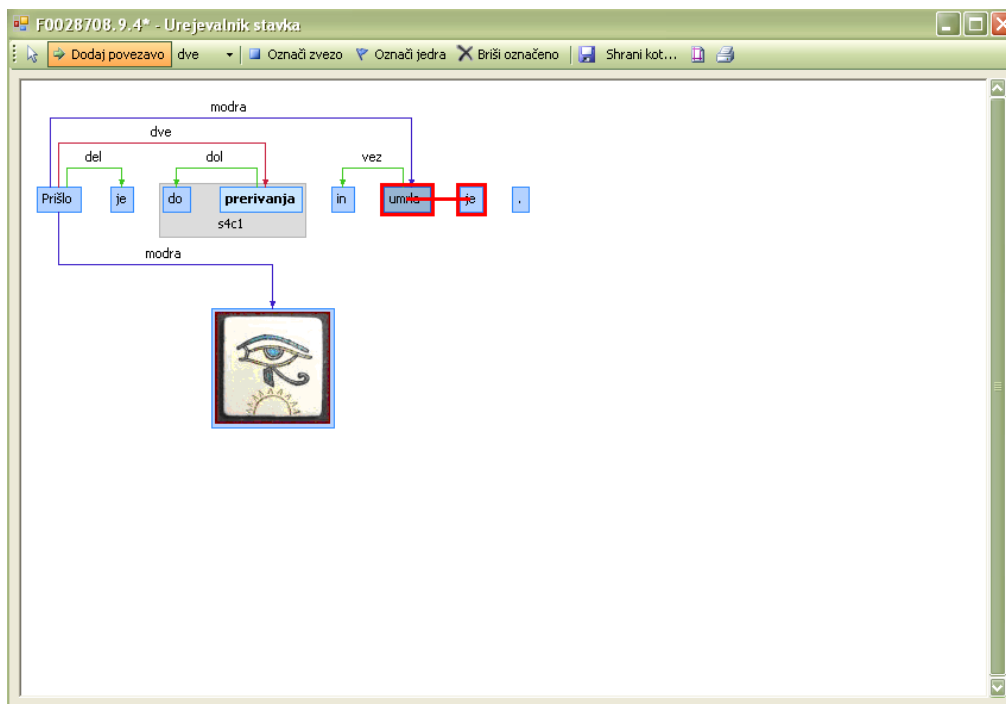
Predvideno je, da se označevalci s programom za označevanje seznanijo v živo, zato na tem mestu samo hiter povzetek funkcij vmesnika.

Prva stran programa je prikazana na spodnji sliki [A]. Z ikono **izberi datoteko** odpremo datoteko s stavki, ki se nato prikažejo v okviru pod ikonami. Po trenutno aktivnih stavkih lahko iščemo z ukazom **išči**, kar rezultira v prikazu samo tistih stavkov, ki ustrezajo iskalnim pogojem. S klikom na **prikaži vse stavke** se vrnemo na izhodiščno stanje prikaza.



[A]

Stavek, ki ga izberemo, se nam v obliki, v kakršni trenutno je (neoznačen ali označen, lahko le delno označen), pokaže v oknu pod naborom stavkov. Z dvojnimi klikom na stavek ali s klikom ikone **urejanje izbranega stavka** se odpre novo okno, v katerem stavek označujemo – glej sliko [B].



[B]

Možnosti označevanja so:

- **dodajanje povezav:** povezave vlečemo z levim klikom miške od izhodišča puščice do cilja puščice (začnemo denimo z jedrnim delom besedne zveze, vlečemo proti nejedrnemu delu). Dokler je povezava aktivna – torej do takrat, ko začnemo početi kaj drugega – lahko spreminjamo ime tipa povezave.
- **poimenovanje povezave:** poteka najhitreje, če uporabljamo bližnjice na tipkovnici (vsakemu tipu povezave ustreza črka ali številka na tipkovnici). Ko izberemo ime povezave, program sam izbere ustrezno barvo puščice. Povezave imamo na voljo tudi v visečem meniju v orodni vrstici na vrhu okna.
- **označevanje zvez ter jeder:** v tej verziji programa je oboje namenjeno samo omogočanju večje preglednosti pri označevanju – v pomoč označevalcem; informacije o označenih zvezah ter jeder se ne zapišejo nikamor, zato oboje lahko uporabljamo, kakor želimo. Zveze označimo tako, da ob držanju tipke CTRL klikamo dele, ki spadajo skupaj, nato pa v orodni vrstici izberemo *Označi zvezo*. Podobno pri označevanju jeder najprej izberemo jedro oz. z uporabo tipke CTRL več jeder hkrati, nato kliknemo *Označi jedra*.

Vse oznake lahko brišemo tako, da izberemo, kar želimo odstraniti (tj. kliknemo na povezavo, jedro ipd.), in pritisnemo DELETE ali kliknemo *briši označeno*.

Program omogoča tudi **shranjevanje stavkov** – pred shranjevanjem datoteko ustrezno preimenujemo – in **tiskanje stavkov**.

4. Pred začetkom označevanja

Pri označevanju upoštevamo naslednja načela:

- označuje se SKLADNJA, kar pomeni, da je vodilo pri problematičnih primerih iskanje rešitev v smeri skladdenjsko primerljivih neproblematičnih primerov → če je za označevanje denimo problematičen primer *pri nas v Leningradu*, si je možno pomagati z manj kompliciranim primerom *pri kanti v kotu*;
- drugo vodilo pri problematičnih primerih je označevanje z opiranjem na pomenske informacije – kadar smo v zadregi, ker jezik omogoča več različnih pomenskih interpretacij, označimo po principu »kakor razumeš« → če v primeru označevanja zveze *ameriška državna administracija* denimo ni jasno, ali pridevnik *ameriška* označuje samo samostalnik *administracija* ali celotno zvezo *državna administracija*, se zanesemo na pomensko interpretacijo primera.
- **doslednost** ter **natančnost** pri označevanju sta za uspešno nadaljevanje projekta ključnega pomena. Posebej pri označevanju daljših stavkov se posamezni elementi radi izmuznejo povezavam (npr. priredja nadstavčnega tipa, nepovezani vezniki ali členki ipd.), čemur se je možno ogniti le z natančnim branjem in pregledovanjem že označenih stavkov. Doslednost skušamo vzpostaviti na ravni celotnega procesa označevanja, kar pomeni, da se pri primerih, ki niso jasno opredeljeni v tem priročniku, posvetujemo z drugimi označevalci in rešitve sprejemamo skupaj z njimi,
- izkušnja iz oblikoskladdenjskega dela označevanja besedil nakazuje, da je čas sprejemanja odločitev oz. reševanja problemov bistveno krajši, če ob identifikaciji problema na mejling listi tudi že podamo ideje – ali naštejemo možnosti – rešitve.
- Primeri, ki v smernicah ponazarjajo nestandardni jezik, so bili predhodno **normalizirani** na ravni zapisa skupaj/narazen (*ne navadno – nenavadno*), prilagajanja zapisa besed fonetični podobi (*jest – jaz*) ipd. Če označujemo podatke, ki predhodno niso bili normalizirani, se je potrebno seznaniti z načeli normalizacije pri projektu Janes in specifikke tega postopka pri interpretaciji smernic ustrezno upoštevati. Druga referenčna točka so smernice za pripisovanje lem in MSD-jev nestandardni slovenščini, kjer so bile sprejete malenkost drugačne odločitve, kot pri preteklih projektih za označevanje standardnega jezika.⁶

Posvetovanje z drugimi označevalci je smiselno tudi zato, ker je identifikacija problemov za razvijalce označevalnega sistema izrednega pomena. Nadvse dobrodošle so tudi povratne informacije na temo Označevalnika ter pričujočih navodil za označevanje.

Zgoraj napisano še posebej velja za prvo fazo označevanja, tekom katere se testira tako označevalni sistem kot tudi navodila za označevanje.

⁶ Smernice za normalizacijo in lematizacijo ter MSD označevanje so dostopne na: <http://nl.ijs.si/janes/wp-content/uploads/2014/09/Janes-smernice-v1.0.pdf>.

5. Povezave

Povezave se delijo v tri skupine, vsaki skupini v Označevalniku ustreza ena barva:

- **zelena** je načeloma namenjena povezavam prvega nivoja, tj. znotraj besednih zvez⁷,
- **rdeča** povezuje drugi nivo, tj. stavčne člene,
- **modra** pa se uporablja za povezave tretjega nivoja (to so načeloma tiste, ki ne ustrezajo prvima dvema kategorijama, pogosto so nadstavčnega tipa, ni pa nujno).

Ko v programu za označevanje izberemo ime povezave, s katero želimo besedi označiti, program sam izbere ustrezno barvo povezovalne puščice.

5.1 Povezave prvega nivoja

Za označevanje razmerij znotraj besednih zvez so na voljo povezave **dol**, **del**, **prir**, **skup**, poleg njih pa na tem mestu obravnavamo še povezavo **vez**.

Ime povezave	Bližnjična tipka v Označevalniku
dol	o
del	e
prir	p
skup	s
vez	v

5.1.1 Povezava DOL

Povezava *dol* načeloma povezuje določujoči in določani del podredne besedne zveze – pri čemer glede tega, kaj je določujoče in kaj določano, v veliki meri sledimo načelom SS. Izhodišče puščice je **jedro** besedne zveze, cilj puščice pa beseda, ki jedro določa, torej **določilo**. O jedrih in določilih v besednih zvezah piše SS 2004: 557–574; v nadaljevanju teh navodil se z identifikacijo jedra ter določil ne ukvarjamo podrobno, zato se spleta ta del SS pred označevanjem preglejati.

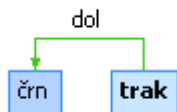
Na tem mestu torej samo shematično: s povezavo *dol* označujemo razmerja znotraj klasičnih:

- **samostalniških besednih zvez** (jedro je samostalnik)

lepa Vida (*kakšna* Vida?), slovenski jezik (*kateri* jezik?), vznožje gore (*katero* vznožje?), Janez Novak (*kateri* Janez?), očetova ura (*čigava* ura?)

Vprašalnice, s katerimi si pomagamo pri določanju jedra, so torej *kakšen/čigav/ kateri* + jedro.

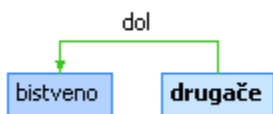
⁷Pojme *besedna zveza* ter *stavčni člani* v teh navodilih uporabljamo splošneje kot je navada v jezikoslovni literaturi, saj kot rečeno ohlapneje dojemamo tudi same jezikoslovne kategorije.



Pri povezavi *dol* je izhodišče puščice **jedro**, cilj puščice pa **določilo** besedne zveze.

- pridevniških oz. prislovnih besednih zvez (jedro je pridevnik oz. prislov)

napol pečen (*kako /koliko* pečen?), meter visok (*koliko* visok), precej izčrpno (*kako /koliko* izčrpno?)

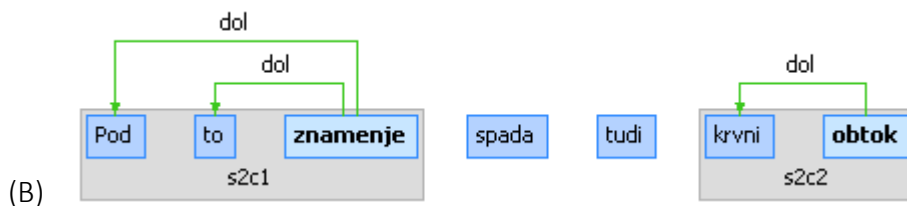
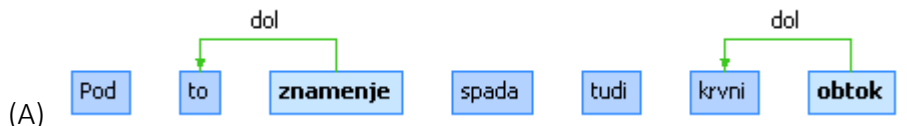


Povezava *dol* se uporablja tudi pri označevanju **povedkovega določila** in povedkovega prilastka (jedro je glagol *biti* oz. polnopomenski glagol), čemur je posvečeno poglavje 0.

Zgoraj navedeni primeri prinašajo enostavne dvobesedne zveze, pri označevanju pa seveda srečujemo kompleksnejše primere, zelo pogosto denimo **besedne zveze s predlogi** (npr. *obsodba na dosmrtno ječo, spodaj ob reki* ipd.)

V takšnih primerih predlog povezujemo z ustreznim samostalnikom (na-ječo, ob-reki), izhodišče puščice je samostalnik, cilj je predlog.

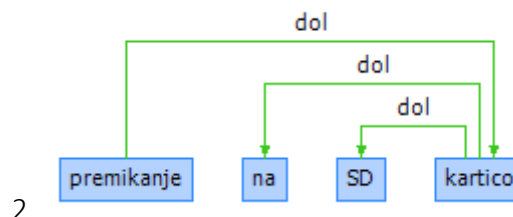
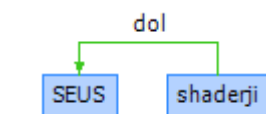
Povezave *dol* v kompleksnih besednih zvezah vedno določamo postopoma, in sicer od spodaj navzgor – najprej povežemo besede, ki so povezane na najosnovnejšem nivoju, nato nadaljujemo s povezavo navzgor, pri čemer se običajno zgodi, da postane beseda, ki je na prvem nivoju jedro, v naslednji fazi določilo. Povezovanje ponavljamo toliko časa, dokler niso vse besede besedne zveze ustrezno povezane, zveza pa ima temu primerno samo eno jedro.



Pogosti v jeziku so primeri, kjer pridevnik ne določa le samostalniškega jedra, ampak **celotno samostalniško besedno zvezo na svoji desni** (tj. tudi pridevnike na svoji desni). V teh primerih postavimo izhodišče puščice v pridevnik, ne samostalnik. Pri odločanju, ali pridevnik določa le samostalnik ali celotno sledečo zvezo, sledimo načelu »označi, kakor razumeš«.



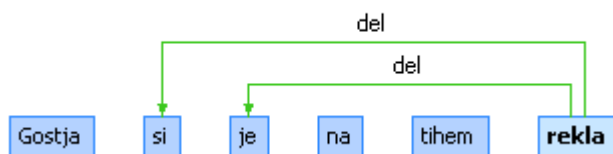
Pri označevanju nestandardnega jezika lahko naletimo na vrste zvez, ki v jezikovnih priročnikih niso natančneje predstavljene oz. so opisane le kot primer neustrezne jezikovne rabe. Tipičen primer so samostalniške **zveze z nesklonljivim prvim delom** (*fitnes center*, *SEUS shader*, *C vitamin*). Označujemo jih kot v primerih spodaj. Za označevanje zvez, ki so v tujem jeziku, gl. poglavje 5.5.2.



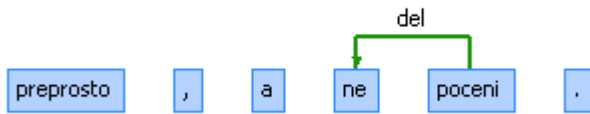
5.1.2 Povezava DEL

Povezava *del* povezuje dele besednih zvez, ki niso povezani na klasičen »odvisnostni« način (na tem mestu torej ne moremo govoriti o prototipnem jedru ter določilu zveze, pač pa zgolj o delih zveze). Tipično uporabljamo povezavo *del* za:

- **dele sestavljenega povedka**: izvor puščice je osebna glagolska oblika oz. deležnik na –I; cilj puščice so morfemi *ne*, *se*, *si*, *bi* oz. oblike glagola *biti* v sestavljenih nesedanjskih časih, tj. *bo*, *je*:



- **določene tipe rabe členka *ne***: kadar nastopa kot negacija sledečega pridevnika, samostalnika ali prislova:



Za nestandardni jezik so značilne elipse, tudi izpuščanje (delov) zloženega osebka. V takih primerih povezujemo po pravilih, ki jih opisuje poglavje 5.5.3.

5.1.3 Povezavi PRIR in VEZ

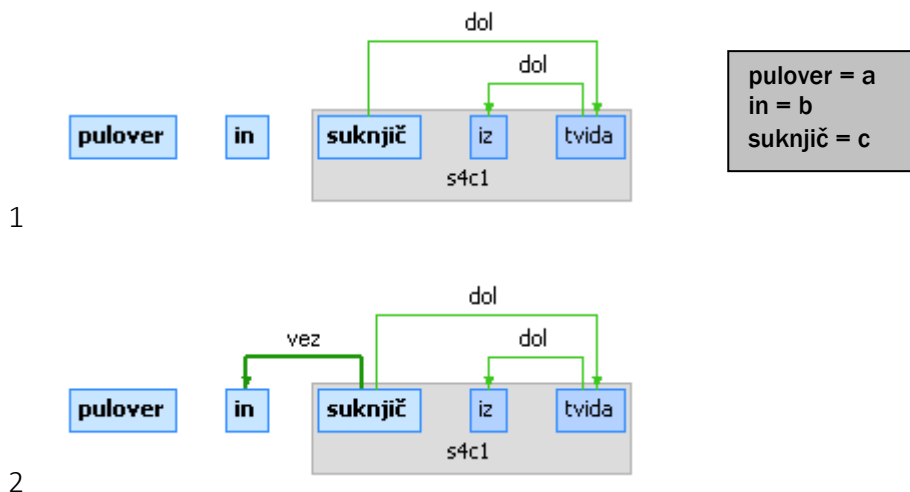
Na prvem nivoju povezav se povezavi *prir* ter *vez* pojavljata v kombinaciji, zato ju obravnavamo na istem mestu. Uporabljamo ju za označevanje prirednih besednih zvez.

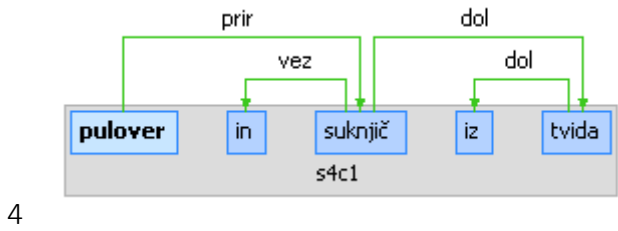
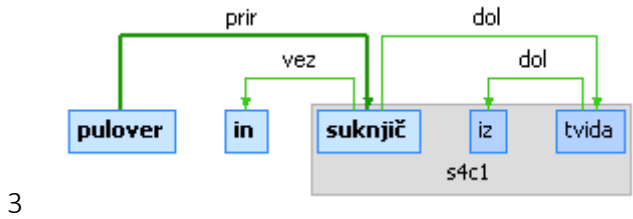
Stavčna priredja označujemo drugače, o tem je govora v poglavju 0.

Pri označevanju priredij na ravni besednih zvez iščemo tri elemente:

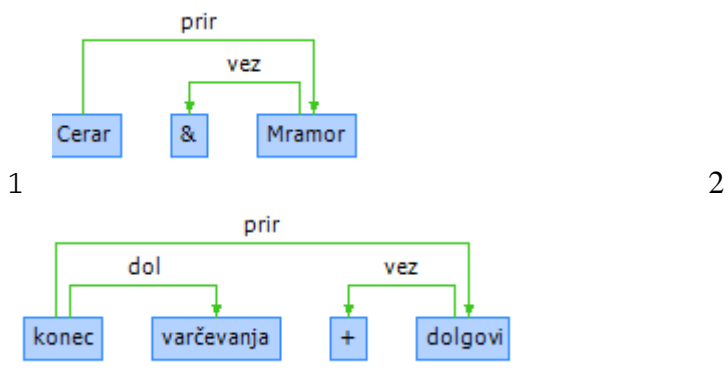
- jedro prvega dela priredja (a)
- priredni veznik ali priredno ločilo, če veznika ni (b)
- jedro drugega dela priredja ©

Elementa b in c povežemo s povezavo *vez*, izhodišče puščice je element c. Nato elementa a in c povežemo s povezavo *prir*, pri čemer je izhodišče puščice element a. Na koncu povezovanja so vse besede povezane, jedro je eno samo (element a):

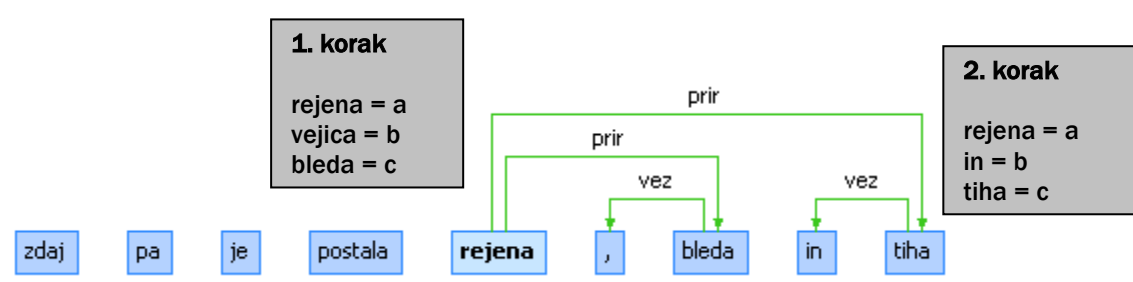


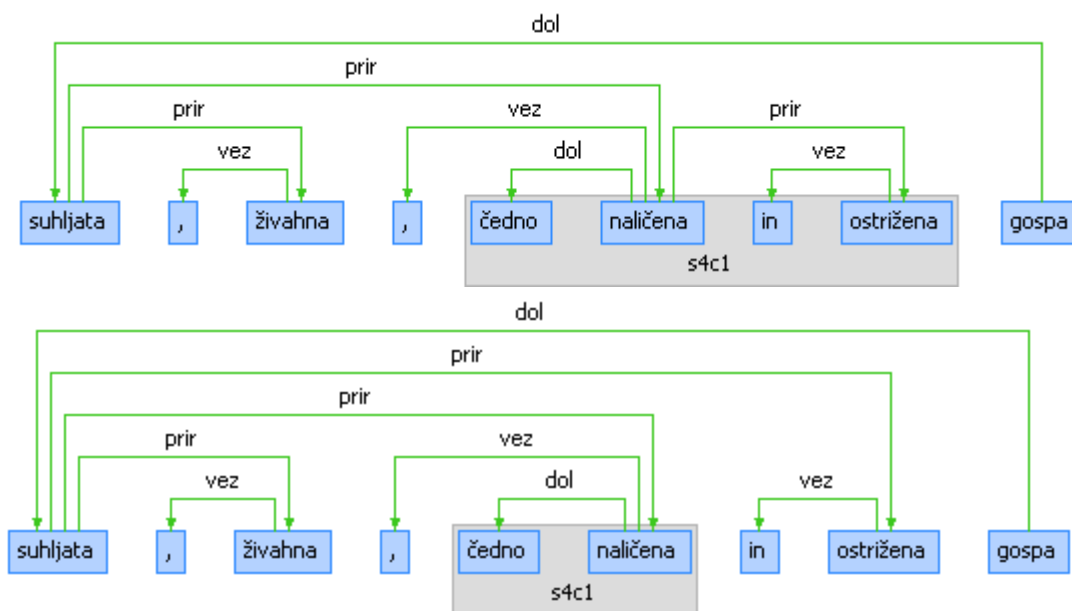


V nestandardni jezikovni se namesto veznikov včasih pojavljajo **simboli** (&, +, /), ki jih povezujemo, kot kažejo spodnji primeri.



Tudi če je elementov priredja več, je jedro prvega dela priredja vedno tisto, ki bo na koncu povezovanja jedro celotne zveze. Ta del je torej stalnica v označevalnem trikotniku. Spodnji primer prikazuje zvezo, kjer so v prirednem odnosu trije pridevniki. Povežemo jih na način, kot je opisan zgoraj:





Pri kompleksnejših prirednih strukturah se priredja lahko gnezdijo eno znotraj drugega in kombinirajo s podrednimi odnosi. Takšne primere označujemo glede na to, **kako stavek razumemo**, po potrebi po posvetu z ostalimi označevalci. V primeru *suhljata, živahna, čedno naličena in ostrižena gospa* lahko pridevnik *čeden* razumem kot določujoči del celotne priredne zveze (*bila je čedno naličena IN čedno ostrižena*) ali pa morda kot določujoči samo za prvi del (*bila je čedno naličena in bila je tudi ostrižena*). Spodaj sta primera označevanja tako prve kot druge interpretacije.

Podoben primer je npr. zveza *delavci in delavke brez zaposlitve* – so brez zaposlitve le delavke ali tako delavke kot delavci. Ob upoštevanju širšega besedilnega konteksta tovrstnih problemov sicer ne bi imeli (*gospa* v gornjem primeru je čedno naličena, pa tudi ostrižena je čedno, brez zaposlitve pa so tako delavci kot delavke); na ravni povedi brez konteksta pa je dvomnosti smiselno predvideti.

5.1.3.1. Besede v vezniški vlogi

Povezava *vez* se uporablja na zgoraj opisani način na besednozvezni ravni, v nadaljevanju bo opisana raba te povezave tudi pri označevanju odvisnih stavkov vseh tipov (0), polstavčnih desnih prilastkov (0), primerjav (0) itd. Na tem mestu je potrebno poudariti, da s povezavo *vez* ne povezujemo le besed, ki so besednovrstno že označene za veznike, ampak vse besede, ki nastopajo v vezniški vlogi na mestu, kjer bi to pričakovali (torej tudi prislove, zaimke itd., npr. *zanima me, kaj reče župnik*).

Veznike povezujemo le znotraj meja povedi. Če se veznik navezuje na predhodno poved, se izognemo nepovezanosti tako, da ga povežemo na metaelement⁸ s povezavo tretjega nivoja (0).

⁸Kot bo vidno iz nadaljevanja, je metaelement oz. korenski element točka zunaj povedi (v Označevalniku jo predstavlja sličica v okvirčku), ki služi povezovanju elementov povedi, ki bi sicer ostali nepovezani.

5.1.3.2 Povezovanje vejice

Kadar v stavku primernege veznika ni, stavčno razmerje pa nakazuje ločilo (običajno je to vejica), povežemo s povezavo *vez* vejico. V vezalnoprrednih razmerjih vejica npr. nadomešča veznik *in*. Pri nekaterih skladenjskih konstrukcijah dosledno povežemo vejico, npr. pri polstavčnih strukturah (0).

Ne smemo mešati uporabe besede *pa* v vezniški vlogi – *gledal pa poslušal* – s členkom *pa*, ki se prav tako lahko pojavlja v prirednih zvezah – *Kajenje ni bilo drogiranje, epizoda v ovalni dvorani pa ne seks*. S povezavo *vez* povežemo besedo *pa* le takrat, ko je popolnoma zamenljiva z veznikom *in*, sicer besedo *pa* večinoma povežemo s povezavo *modra* barve na metaelement – s povezavo *vez* pa v tem primeru označimo vejico.

Za primere nestandardne jezikovne rabe, kjer tovrstna **vejica manjka**, glej poglavje 5.5.4.

5.1.4 Povezava SKUP

S povezavo *skup* povežemo (večinoma – ne pa nujno – nepolnopomenske) besede, ki imajo zelo močno tendenco po sopojavljanju, tvorijo torej neke vrste večbesedne enote, ne gre pa za klasične podredne ali priredne besedne zveze. S povezavo *skup* označujemo:

- besede, ki imajo variantni zapis skupaj ali narazen (npr. *kakor koli* – *kakorkoli*, *čim večji* – *čimvečji itd.*),
- določene večbesedne veznike (*ne le da*),
- druge večbesedne zveze tega tipa.

S povezavo *skup* povežemo le sosednje (skupaj stoječe) besede. Seznam zvez, ki naj bi jih označevali s to povezavo, je **pripravljen vnaprej** – temelji na pogostnosti pojavljanja zvez tega tipa v korpusu FidaPLUS. S tem naborom si lahko tudi pomagamo pri odločitvi, ali je določena zveza, čeprav je na seznamu ni, ustrezna kandidatka za označevanje s povezavo *skup* (o tem, ali je besedno zvezo smiselno označevati z oznako *skup* in jo dodati na že pripravljen seznam, se posvetujemo z drugimi označevalci oz. s pripravljalci označevalnega sistema).

Izhodišče puščice pri povezavi *skup* je prvi del zveze oz. tisto, kar je označeno kot jedro na seznamu skupov. Cilj puščice je sosednja beseda. Če je skup sestavljen iz niza besed, vlečemo puščice zaporedno, kot prikazuje spodnji primer:



5.2 Povezave drugega nivoja

Na ravni stavčnih členov pripisujemo povezave **ena**, **dve**, **tri** in **štiri**. Načeloma te povezave sovpadajo s stavčnočlenskimi odnosi, kakršne poznamo iz Slovenske slovnice oz. opisov slovenskega jezikoslovja, vendar ne nujno povsem, zato so poimenovane na jezikoslovno nevtralen način. Sorodnost je naslednja:

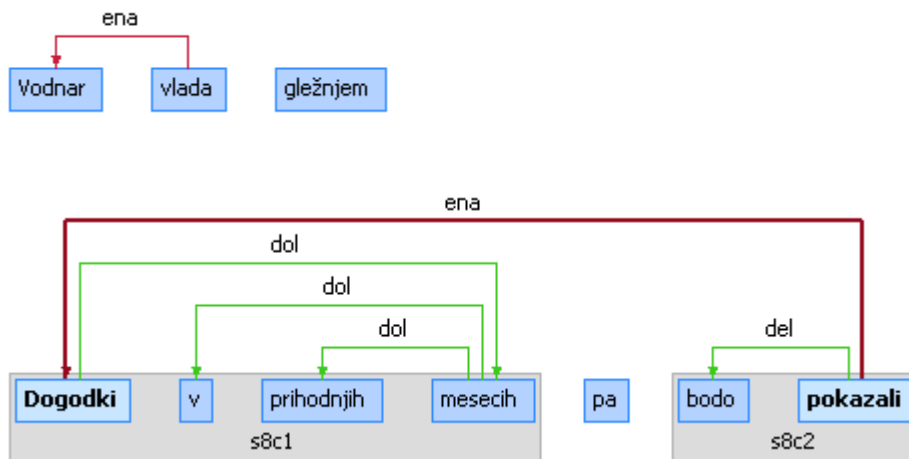
ena	osebek stavka ← povedek stavka
dve	predmet stavka ← povedek stavka
tri	prislovno določilo lastnosti ← povedek stavka
štiri	ostala prislovna določila ← povedek stavka

Označevalnik povezavam te skupine pripisuje rdeče povezovalne puščice. Bližnjice na tipkovnici so ustrezni numerični znaki (1, 2, 3, 4).

Povezave te kategorije pripisujemo v večini primerov znotraj stavčnih meja, v določenih primerih pa tudi preko le-teh (kadar je v vlogi stavčnega člana celoten stavek, kakor npr. pri osebkovih, predmetnih, prilastkovih odvisnikih – ti primeri so razloženi sproti).

5.2.1 Povezava ENA

Na **znotrajstavčni ravni** povezava *ena* načeloma ustreza razmerju med povedkom ter nestavčnim osebkom stavka: pri določanju si pomagamo z vprašalnico *kdo ali kaj + povedek*. Pri označevanju je izhodišče razmerja jedro povedka, cilj razmerja pa jedro osebka:



Na **medstavčni ravni** s to povezavo povezujemo osebkove odvisnike z ustreznim nadrednim stavkom: izhodišče puščice je jedro povedka nadrednega stavka, cilj puščice je jedro povedka odvisnika (o osebkovih odvisnikih piše Slovnica 2004 na strani 638):

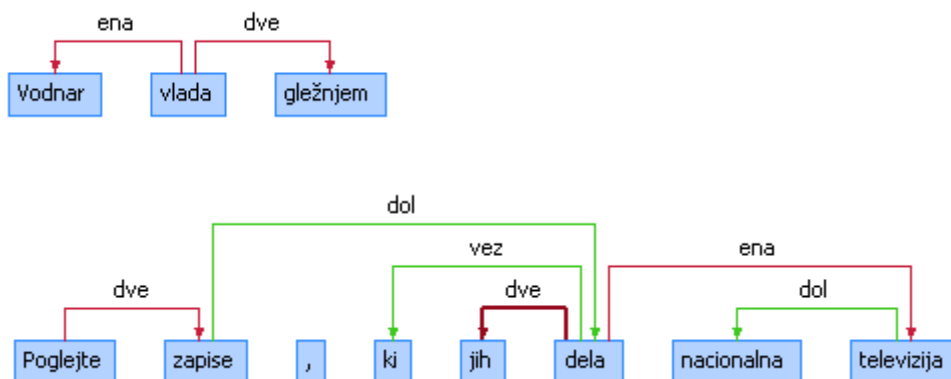


Logičnega osebka ne označujemo.

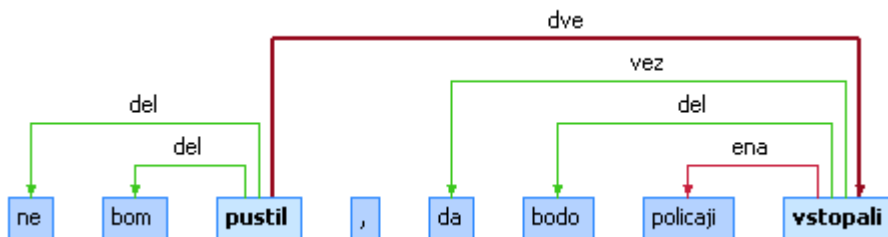
V primerih tipa *Lužinovi se je zdelo* bi bil samostalniik *Lužinovi* po Slovenski slovnici potencialno označen kot osebek, ker je nosilec dogajanja. Po sistemu označevanja pa odnosov tega tipa ne označujemo s povezavo ENA, ampak s povezavo DVE (sledeč vprašalnici *komu – čemu se je zdelo?*). Kot osebki so lahko označeni samo zanikani osebki v roditeljski in partitivni roditeljici.

5.2.2 Povezava DVE

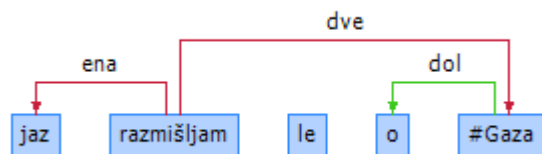
Na znotrajstavčni ravni povezava *dve* načeloma ustreza razmerju med povedkom ter nestavčnim predmetom stavka: pri določanju si pomagamo z vprašalnico za neimenovalniške sklone (*koga ali česa, komu ali čemu, koga ali kaj ...*) + povedek. Pri označevanju je izhodišče razmerja jedro povedka, cilj razmerja pa jedro predmeta:



Na medstavčni ravni s to povezavo povežemo tudi predmetne odvisnike z ustreznim glavnim stavkom: izhodišče puščice je jedro povedka glavnega stavka, cilj puščice je jedro povedka odvisnika (o predmetnih odvisnikih piše Slovnica 2004 na straneh 639–640):



Če se v nestandardnem jeziku pojavljajo stavčni členi na atipične načine, npr. v nepričakovanih sklonih, povežemo po analogiji s standardnem jezikom oz. kakor zahteva pomen stavka. Spodnji primer prikazuje povezavo elementa *#Gaza*, ki je v stavku v imenovalniku, vendar ga povežemo kot predmet s povezavo *dva*. (Več o povezovanju elementov, kot so ključniki, emotikoni, spletni naslovi ipd., v poglavju 5.5.1.)

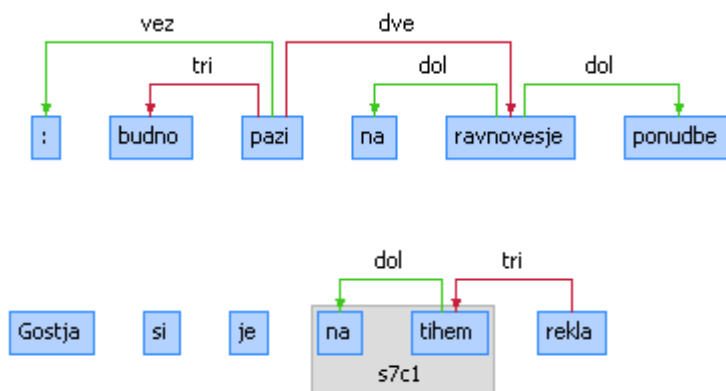


5.2.3 Povezava TRI

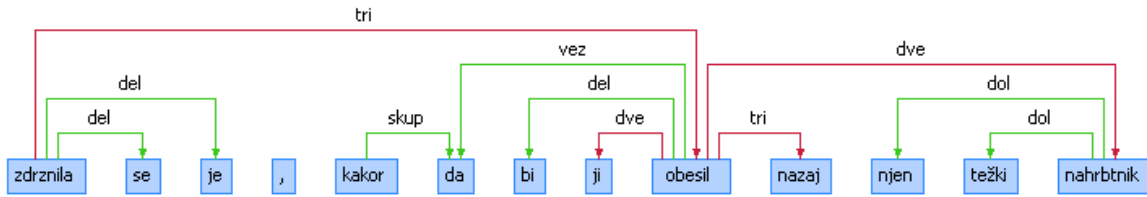
Na **znotrajstavčni ravni** povezava *tri* načeloma ustreza razmerju med povedkom ter nestavčnim prislovnim določilom lastnosti: pri določanju si pomagamo z naslednjimi vprašalnici (sledimo klasifikaciji Slovenske slovnice 2004, za primere glej 408 ter 622–625):

prislovna določila lastnosti identificirajo vprašalnice: *kako* (na kakšen način), *koliko*, *kolikokrat*, *kolikanj*, *kolikič*, *katerikrat*, *čemu podobno*, *kako* (s čim), *do katere mere* (kako zelo), *s kakšnim rezultatom* oz. *uspehom* /.../

Pri označevanju je izhodišče razmerja jedro povedka, cilj razmerja pa jedro prislovnega določila:

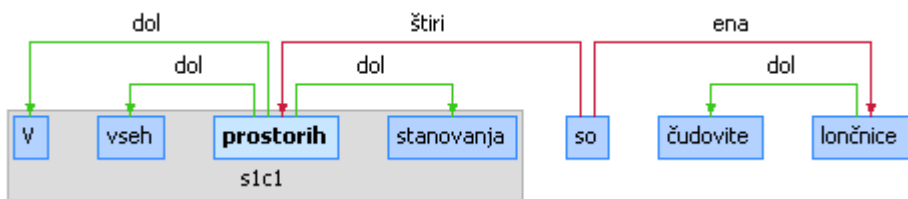
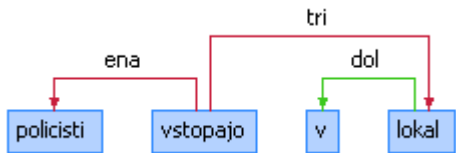


Na **medstavčnem nivoju** s to povezavo povežemo tudi lastnostne odvisnike oz. polstavke z ustreznim glavnim stavkom: izhodišče puščice je jedro povedka glavnega stavka, cilj puščice je jedro povedka odvisnika (o lastnostnih odvisnikih piše Slovnica 2004 na straneh 641–643):

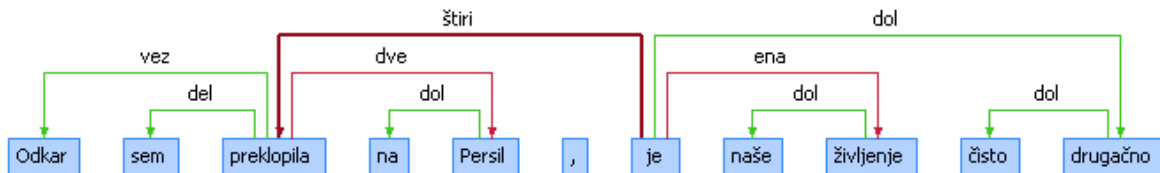


5.2.4 Povezava ŠTIRI

Na znotrajstavčnem nivoju povezava *štiri* načeloma ustreza razmerju med povedkom ter ostalimi nestavčnimi tipi prislovnih določil (vsemi, ki niso zajeti s povezavo *tri*): pri določanju si pomagamo z vprašalnico za ta prislovna določila + *povedek*. Pri označevanju je izhodišče razmerja jedro povedka, cilj razmerja pa jedro prislovnega določila:

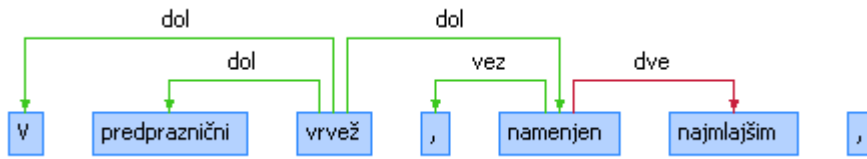


Na medstavčni ravni s to povezavo povezujemo tudi odvisnike ustrežajočih tipov z glavnim stavkom: izhodišče puščice je jedro povedka glavnega stavka, cilj puščice je jedro povedka odvisnika (o odvisnikih piše Slovnica 2004 na straneh 640–645):



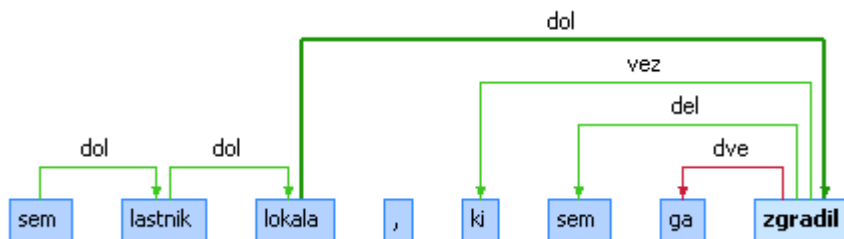
5.2.5 Polstavki

Povezave drugega nivoja določamo tudi znotraj **polstavkov**. Izhodišče puščice je v tem primeru polstavčno jedro (o polstavkih in tipičnih jedrih teh konstrukcij v Slovnici 2004: 632), cilj puščice je enako kot pri zgoraj opisanih primerih jedro ustreznega stavčnega člena. Za označevanje polstavka glej tudi poglavje 5.4.8 Polstavčni desni prilastki.



5.2.6 Prilastkovi odvisniki

Prilastkove odvisnike povezujemo s povezavo **dol**: izhodišče puščice je beseda, na katero se prilastkov odvisnik nanaša, cilj puščice je jedro povedka odvisnika (o prilastkovem odvisniku piše Slovničica 2004 na straneh 645–646):



5.2.7 Predmet ali prislovno določilo?

Posebej v primeru predložnih zvez je problematično ločevanje med stavčnim predmetom ter prislovnim določilom – npr. *spremenila se je v gospo – v koga/kaj se je spremenila?* ali *kako se je spremenila?* Slovničica 2004 (615-616) dilemo rešuje na takle način:

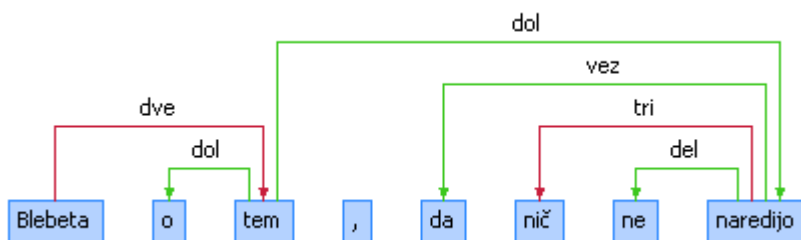
V zvezi s predlogi imamo ob takih vprašalnicah [tj. *kdo/kaj, koga/česa ...*] opraviti s predmeti le, če se po istem ni mogoče vprašati tudi s *kje, kdaj, kako, zakaj* ipd., tj. z vprašalnicami za prislovna določila (prim. *Ne hodi za mano: za mano je predmet, če hoditi za pomeni 'dvoriti'; pa prislovno določilo, če glagol pomeni premikanje*).

Kadar smo torej v dilemi, je varnejša odločitev za prislovno določilo, tako vsaj pravi Slovničica. Zelo ohlapna pomoč je lahko občutek, da so orodniške ter mestniške zveze pogosteje prislovna določila, roditeljske ter tožilniške bolj predmeti – vendar ti podatki še niso preverjeni na večji količini realnega teksta, zato slepo opiranje nanje ni priporočljivo.

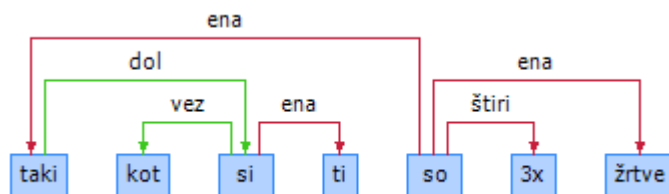
Ker nam na tem mestu ločevanje med predmetom ter prislovnim določilom v grobem pomeni tudi ločevanje med obveznimi ter neobveznimi stavčnimi sestavinami, se v prvi vrsti vprašamo, **kaj problematični del stavka pomeni za glagol**. Če ga dojemamo kot del glagolske fraze (*zaljubiti se v nekoga, spremeniti se v nekoga*), potem označujemo s povezavo DVE. Če ga ne dojemamo kot del glagolske fraze (*kapljati v zvezah kapljati s sveče / kapljati s strehe*), ga označimo s TRI oz. ŠTIRI. »Del fraze« pri tem razumemo skladenjsko in ne pomensko (tudi primera *kapljati na zabavo* ne bi označili z DVE).

5.2.8 Uvajanje odvisnikov s kazalnimi zaimki

Kadar pri označevanju naletimo na uvajanje odvisnika s kazalnim zaimkom (stavki tipa *moti me to, da ne poveš do konca*), jih označimo, kot prikazuje spodnji primer (uporabimo povezavo **dol**, podobno kot pri prilastkovih odvisnikih 0):



Za nestandardni jezik je značilno, da je tovrstne kazalnosti več, zato spodaj navajamo še primer z vrinjenim odvisnikom. Glede nestandardne rabe ločil gl. poglavje 5.5.4.



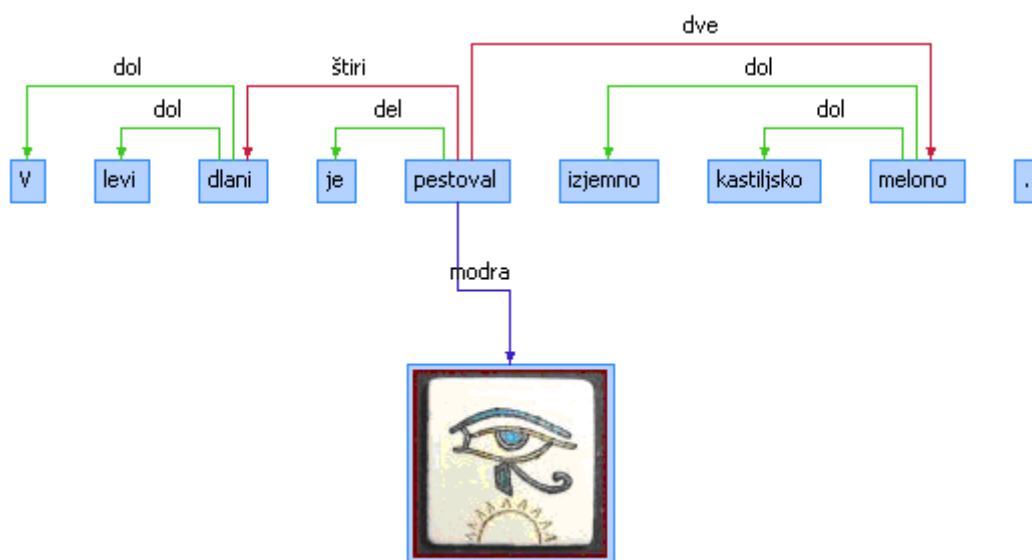
5.3 Povezave tretjega nivoja

Povezave tretjega nivoja so bolj ali manj tiste, ki jih ne pokrijemo s prvim ali drugim nivojem: ta tip povezav je torej namenjen povezovanju elementov, ki bi sicer ostali nepovezani. Povezavo prikazuje modra puščica. Kljub temu da na tej ravni povezujemo različne tipe razmerij, je povezava vedno poimenovana **modra**.

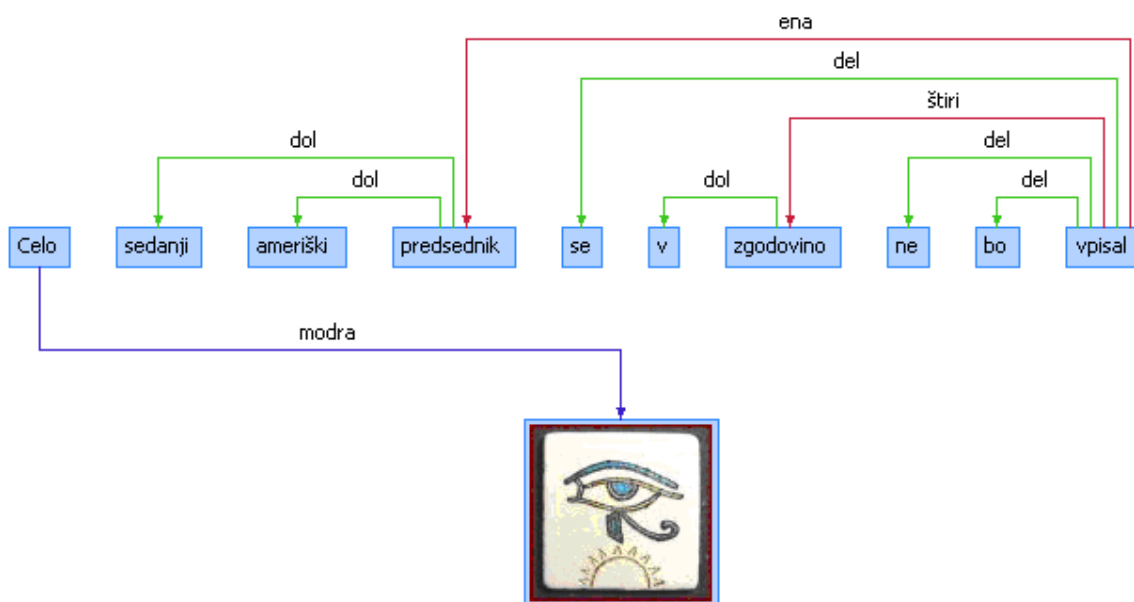
S povezavo *modra* povezujemo:

- jedro povedka glavnega stavka povedi na metaelement [A],
- členke na metaelement [B],
- veznike, ki se navezujejo na prejšnjo poved, na metaelement [C],
- jedra pastavkov ali vrinjenih stavkov na metaelement [Č],
- eliptične dele stavkov, ki jih ne moremo povezati nikamor drugam in bi sicer ostali nepovezani; predvsem v nestandardnem jeziku je tovrstnih primerov veliko [D],
- daljše segmente v tujem jeziku [E],
- skladenjsko proste elemente, značilne za računalniško posredovano komunikacijo, kot so spletni naslovi, emotikoni in emoji, ključniki in sklici na uporabniška imena [F].

Kadar povežemo nepovezane dele stavka z metaelementom, je izvor puščice vedno stavčni del, metaelement je cilj:⁹

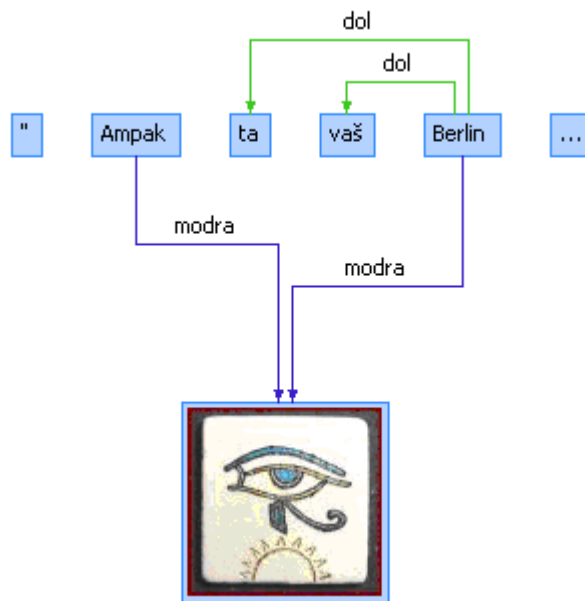


[A]

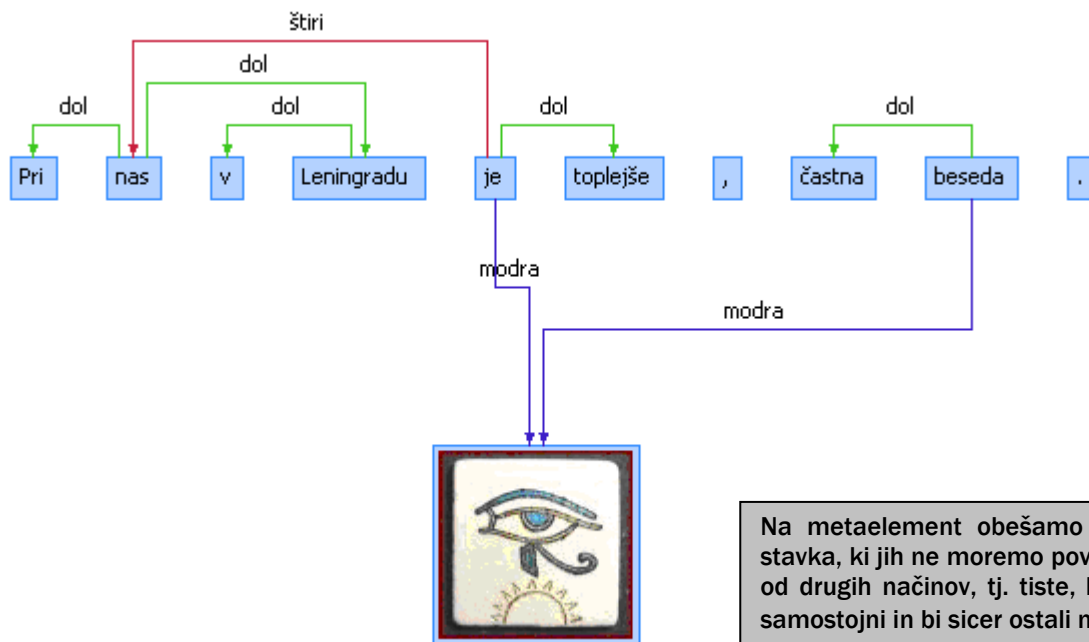


[B]

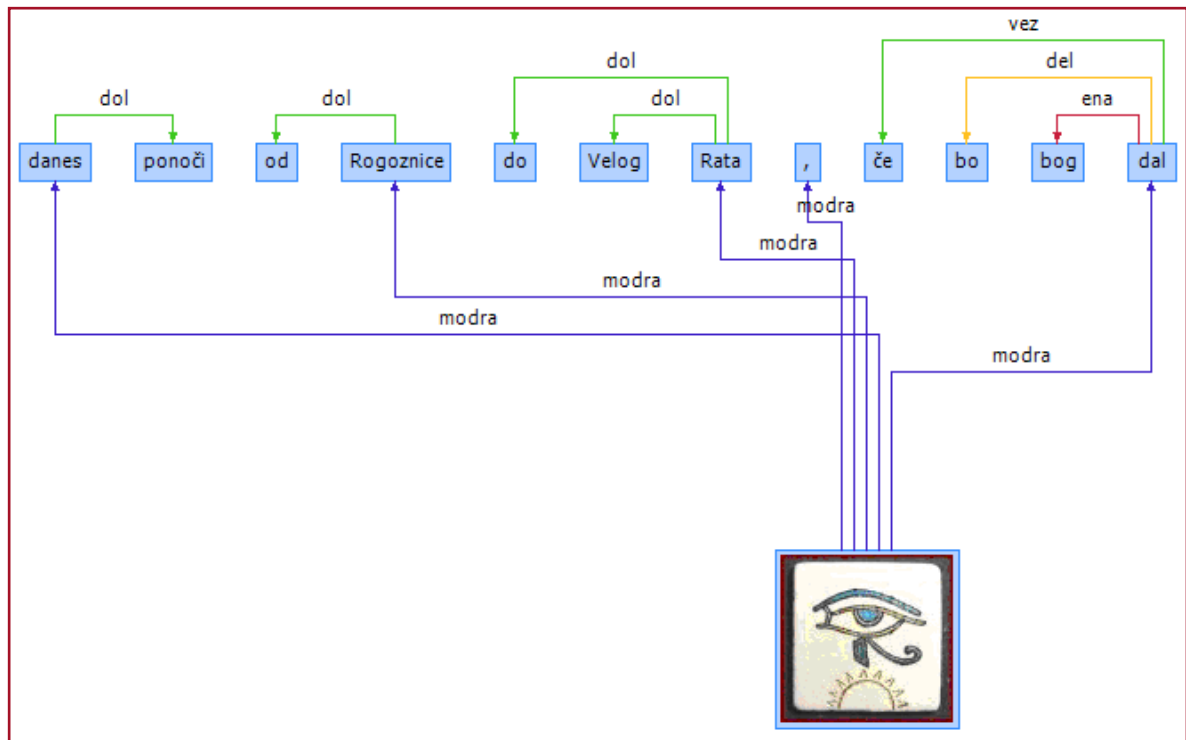
⁹ V kasnejših označevalnih projektih se je to načelo spremenilo tako, da je izhodišče puščice metaelement.



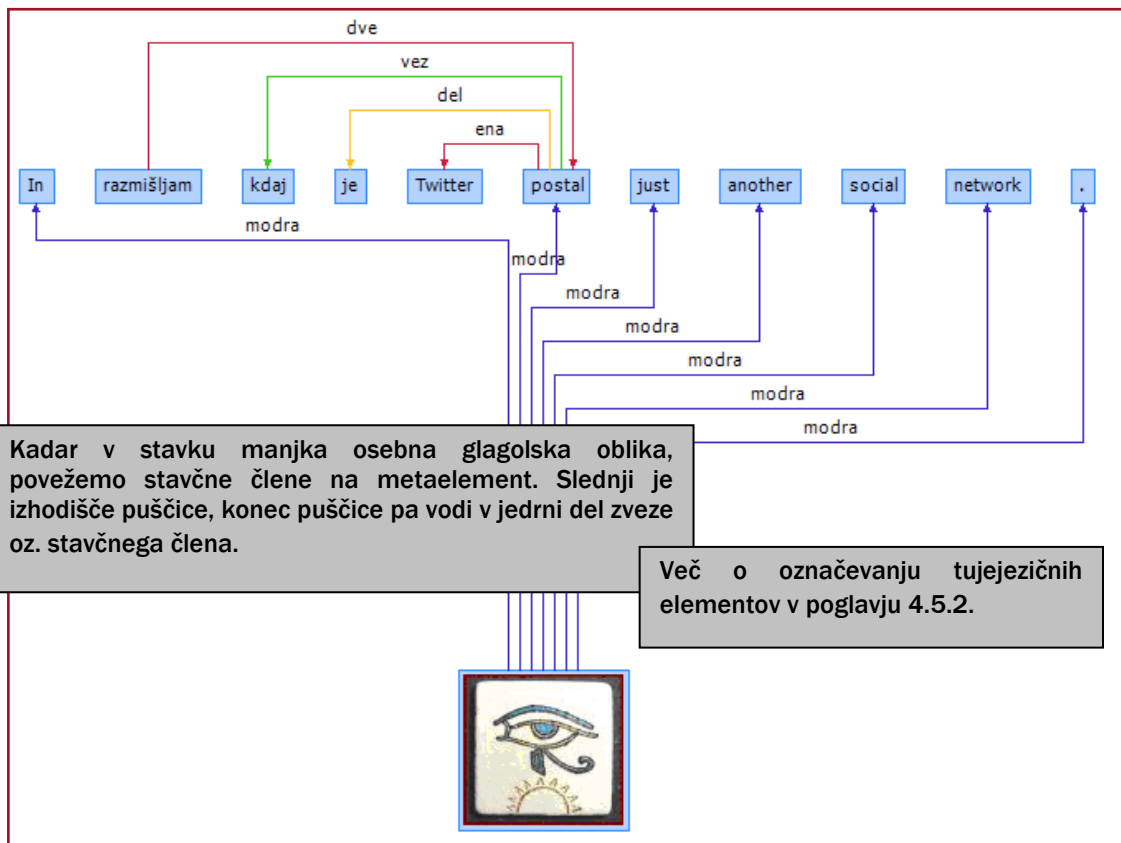
[C]



[Č]



[D]

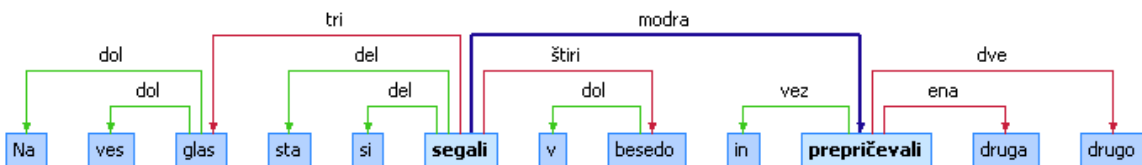


Kadar v stavku manjka osebna glagolska oblika, povežemo stavčne člene na metaelement. Slednji je izhodišče puščice, konec puščice pa vodi v jedrni del zveze oz. stavčnega člena.

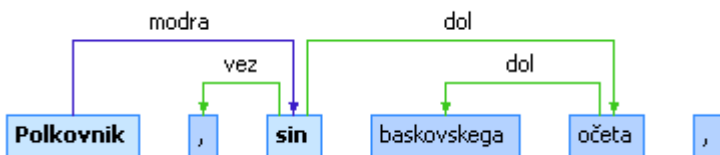
Več o označevanju tujejezičnih elementov v poglavju 4.5.2.

[E]

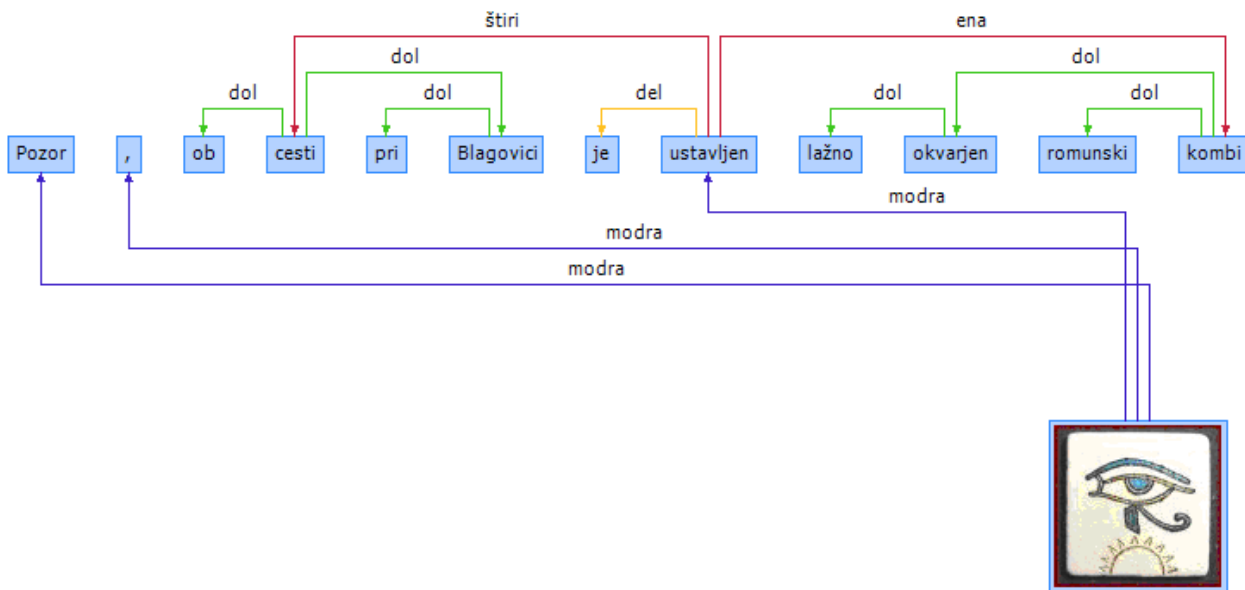
Prav tako s povezavo *modra* povežemo med sabo povedkovna jedra prirednih stavkov, in sicer po sistemu, ki velja za označevanja priredij znotraj stavka (0) – prvi element priredja je izhodišče puščice, naslednji je cilj. Ta tip povezovanja uporabljamo za vsa medstavčna priredja, ne le vezalnega.



Nazadnje s to povezavo povežemo pristavke, kjer povežemo jedro pristavka z jedrom predhodne zveze. Kar loči te zveze od skladenjsko podobnih, je dejstvo, da sta imata s povezavo *modra* povezana dela stavka enako referenco (*polkovnik, sin baskovskega očeta* pomeni *polkovnik = sin baskovskega očeta*), ne gre za pojasnjevanje lastnosti (*trgovine, odprte tudi ob nedeljah*):



V kasnejših kampanjah za označevanje standardne slovenščine je bilo odločeno, da se povedkovna jedra prirednih stavkov ter izpostavki vseh vrst povezujejo **neposredno na metaelement**. V projektu Janes sledimo tem novostim (vendar smernic ne popravljamo v delu, ki se tiče standardnega jezika). V spodnjem primeru gl. povezovanje besede *pozor*.



5.4 Posebni primeri

5.4.1 Vezniki

Večdelne veznike povezujemo s povezavo:

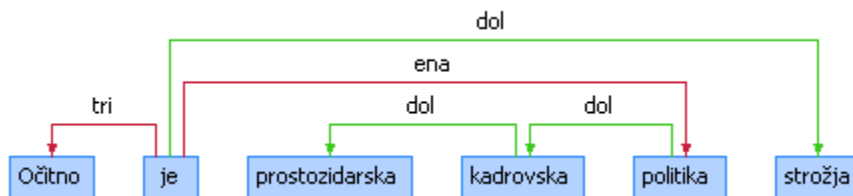
- **skup**, kadar gre za zvezo z vnaprej pripravljenega seznama skupov ali zvezo, sorodno tem, ki se na seznamu pojavljajo,
- **dol**, kadar gre za zvezo, kjer lahko določimo element, ki je v zvezi najbolj ključen oz. »vezniški« (npr. *da/ko* v kombinacijah *kakor da*, *vedno ko* itd.),
- **večdelnih veznikov, katerih deli ne stojijo drug ob drugem, ne povezujemo** (npr. parov *niti – niti*).

5.4.2 Členek ali prislov?

Kaj je členek in kaj prislov, je vprašanje, na katerega ni enostavnega odgovora. Načeloma na skladenjskem nivoju za členek smatramo tiste dele stavka, po katerih se ne moremo vprašati in jih posledično ne moremo povezati z nobeno od povezav prvega ali drugega nivoja.

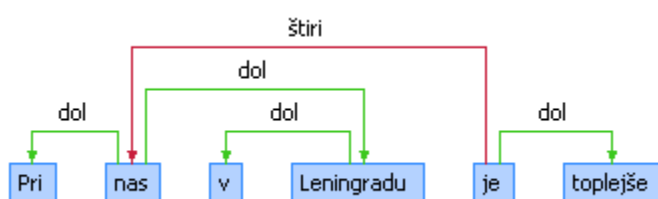
5.4.3 Izražanje mnenja ali ocene

Izražanje mnenja in ocene z zvezami, kot so npr. *očitno je*, *zanimivo je* itd., označujemo s povezavo TRI:



5.4.4 Sestavljene predložne zveze

Sestavljene predložne zveze tipa *pri nas v Leningradu* obravnavamo kot eno zvezo, kljub temu da je možno z vprašalnici zvezo razdeliti na dva dela (*kje? pri nas + kje? v Leningradu*). Jedro določamo po analogiji s podobnimi, a enostavnejšimi primeri (*pri kanti v kotu*):



5.4.5 Povedkova določila in povedkovi prilastki

Povedkova določila in povedkove prilastke označujemo s povezavo **dol**. Povlečemo jo med povedkovim jedrom (tj. kopulo, glagolom *biti*, oz. polnopomenskim glagolom) ter povedkovim določilom in povedkovim prilastkom. Izhodišče puščice je v povedkovem jedru, njen cilj pa v jedru povedkovega določila oz. povedkovega prilastka. Glede ločevanja med povedkovimi določili ter prilastki sledimo smernicam Slovnice 2004 (588):

[Povedkove prilastke] lahko razstavimo na dve enoti, ki sta sporočilno [...] samozadostni: *hoditi bos – hoditi, biti bos; vrniti se (kot zmagovalec) – vrniti se, biti zmagovalec*. [Primeri s povedkovimi določili] so sporočilno samozadostni le, če so drug ob drugem; *biti star* se ne da razstaviti po zgledu *hoditi bos*, ker bi glagol pri tem dobil drugačen pomen, kot ga ima v dani zvezi.

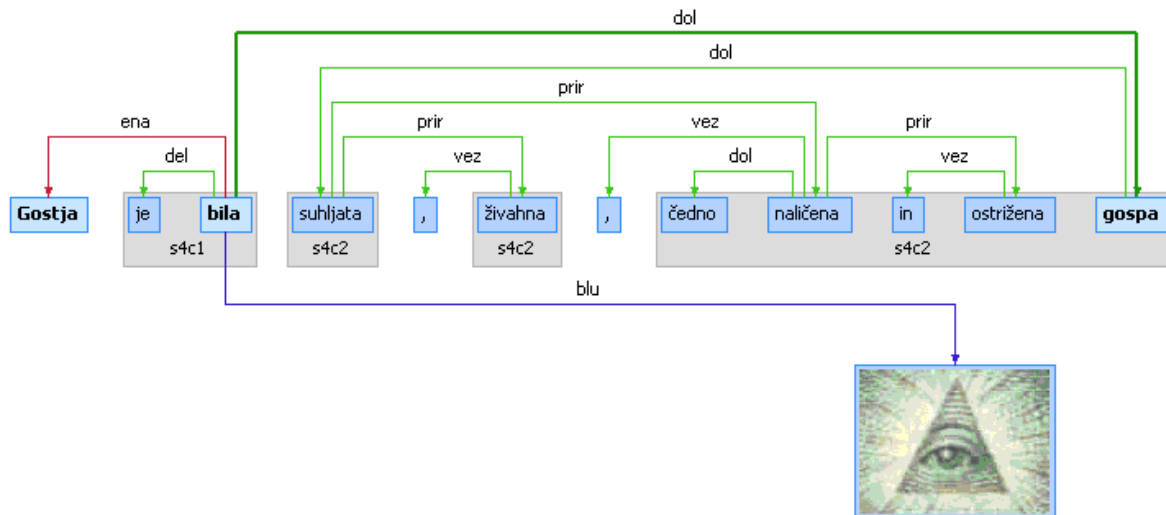
Glede na smernice sistema je označevanje nekoliko poenostavljeno v smislu, da lahko povedkovo določilo nastopa le ob (**nepolnopomenskem**) glagolu *biti* (*bil je ekonomist, nadarjena je bila, njegov načrt je bil, da bi napadli sovražnika, bili so vesele narave, popravljati napake je vaša naloga*), vsi drugi glagoli so obravnavani kot polnopomenski. Če nismo prepričani, ali naj strukturi ob glagolu *biti* pripišemo povedkovodoločilni status oz. ali naj jo označimo kot osebkovo ali prislovnodoločilno, ji določimo status povedkovega določila in jo povežemo z odnosom *dol*. Če torej oklevamo, ali je glagol *biti* polnopomenski ali ne oz. ali je struktura ob njem povedkovo določilo oz. osebek ali prislovno določilo, ima prednost povedkovodoločilna interpretacija (povezava *dol*).

Strukture ob »pomensko oslavljenih glagolih«¹⁰ (zgolj informativni in ne dokončen seznam glagolov oz. ustreznih struktur je naveden v prilogi B), ki jih Slovenska slovnica (str. 590–591) označuje kot povedkova določila, v sistemu označimo s povezavo DVA, če so samostalniške zveze (*ostal je učitelj, imam te za prijetnega človeka, imenovali so ga za predsednika ZDA, velja za poštenjaka* ipd.), in s povezavo dol, če gre za pridevniške zveze (*postaja bled, dela se neumnega, čutil se je odgovornega, zdi se zaspan* ipd.).

Povedkov prilastek označujemo s povezavo *dol*. Pojavlja se ob pomensko popolnih glagolih (v sistemu vsi glagoli razen *biti*), ob glagolu *biti* pa samo, če gre za polnopomenski glagol (*Oče je doma vesel* 'oče se nahaja doma in je vesel'), ob kopuli *biti* pa le, če je v stavku tudi povedkovo določilo (*Peter je bil igralec prvi od nas* 'Peter je postal igralec prvi, drugi smo bili igralci šele kasneje'). Strukturi lahko pripišemo status povedkovega prilastka le, če je od samostalniške zveze, ki jo preko glagola oz. povedka posredno določa, prostorsko ločena in če jo lahko izpustimo, ne da bi pri tem stavek postal neslovničen (spremeni se le njegov pomen). Povedkov prilastek torej povedku oz. samostalniški zvezi pripisuje samo dodatno lastnost (*Z dopusta se je vrnil dobre volje, Zamišljen je pokimal, Umril je mlad* ipd.)

Tipično povedkovo določilo sestoji iz treh elementov, osebk (A), povedka (B) ter povedkovega določila (C). npr. *moja mama je čarovnica // gostja je bila suhljata gospa // sem pilot* – uporaba vprašalnic v teh primerih je nekako *kdo ali kaj + je + kdo ali kaj*. V primerih, da iz stavka ni povsem jasno, kateri del stavka je osebek, kateri pa določilo (*simbol kastracije je mati*), se posvetujemo z ostalimi označevalci. Sicer je kombinacija oznak tipično takšna, kot v spodnjem primeru: deležnik je jedro, ki tvori z morebitnim pomožnikom razmerje *del*, z jedrom osebk razmerje *ena* ter z jedrom povedkovega določila razmerje *dol*.

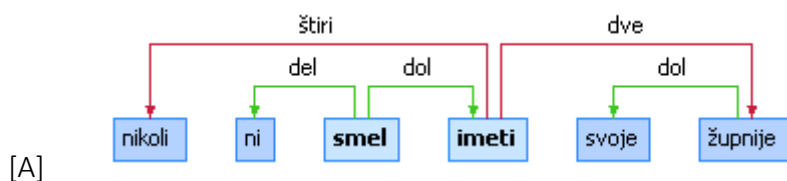
¹⁰ V sistemu JOS so obravnavani kot polnopomenski glagoli.



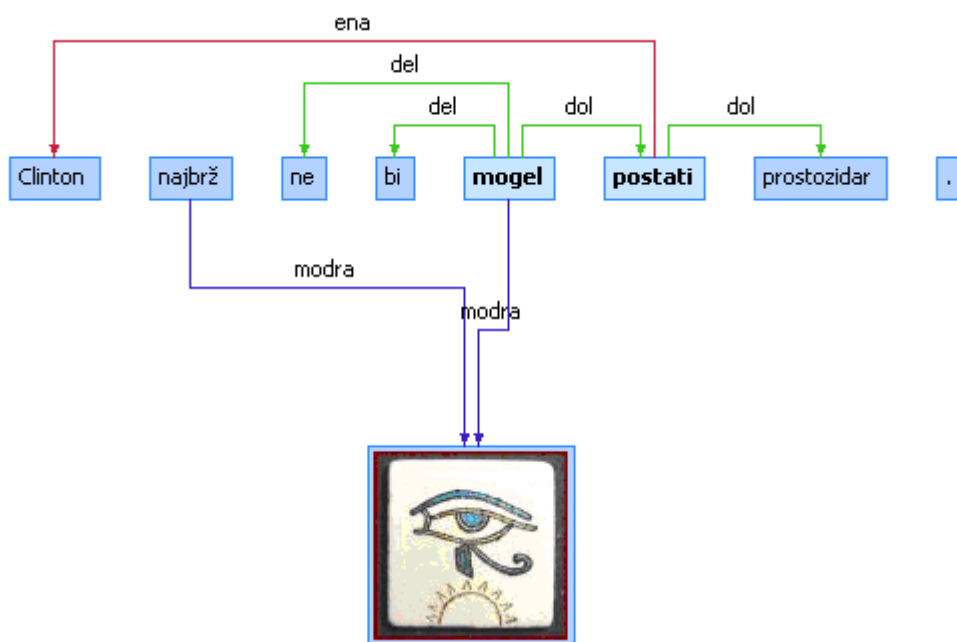
5.4.6 Modalni glagoli¹¹

5.4.6.1 Morati, smeti, hoteti itd.

Konstrukcije z modalnimi glagoli označimo tako, da za izhodišče izberemo modalni glagol, nedoločnik oz. namenilnik pa označimo kot določilo modalnega glagola. Kadar konstrukciji sledijo drugi stavčni členi, jih običajno povezujemo iz nedoločnika – slednji je torej izhodišče za povezave ENA, DVE, TRI, ŠTIRI [A], pa tudi *dol* v primeru povedkovega določila oz. povedkovega prilastka [B].



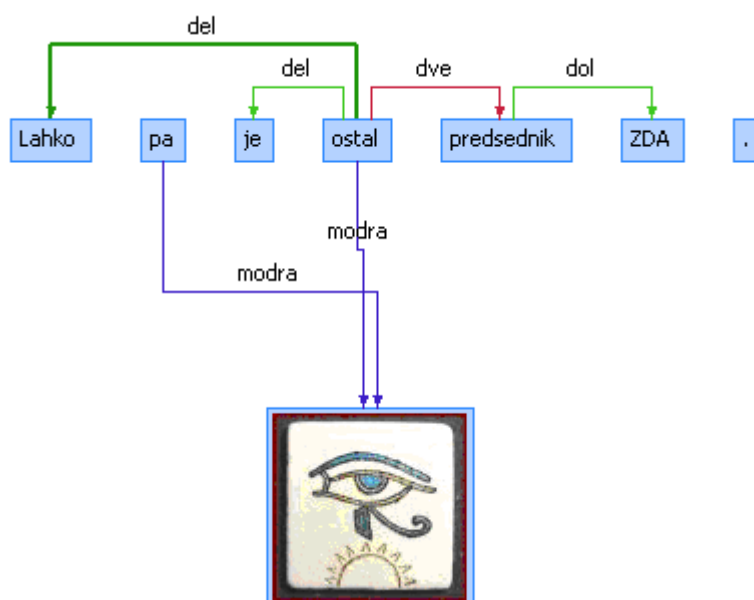
¹¹Na enak način obravnavamo vse glagole, ki nastopajo v zvezah z nedoločniki in namenilniki.



[B]

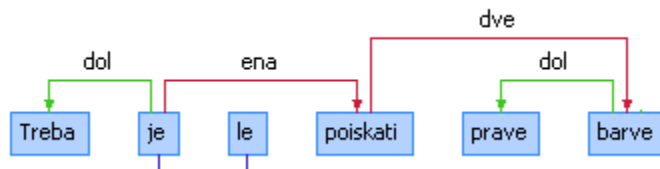
5.4.6.2 Lahko je

Nekoliko poseben primer je izražanje modalnosti z *lahko* (*lahko je ostal predsednik*), ker je edina kombinacija, kjer ob glagolu nastopa osebna glagolska oblika. Zvezo označujemo tako, da za izhodišče zveze izberemo osebno glagolsko obliko, *lahko* pa nanjo povežemo s povezavo *del*:



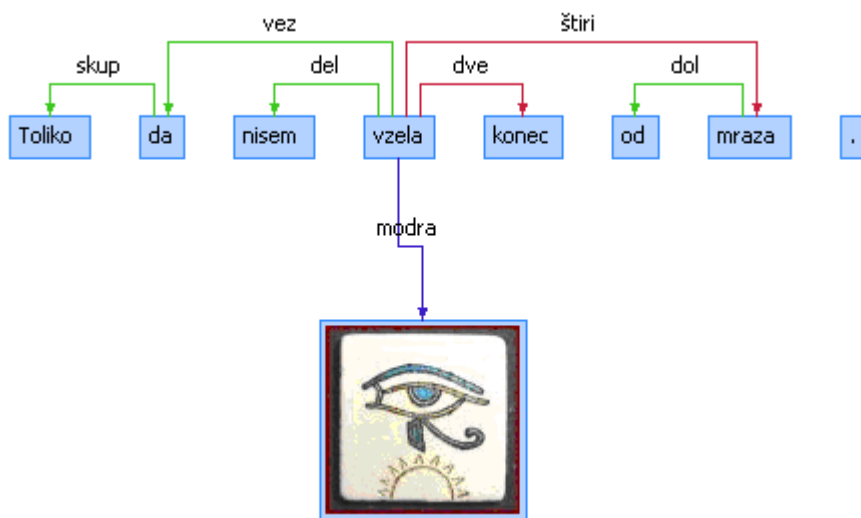
5.4.6.3 Treba je

Na tem mestu še primer označevanja zvez tipa *treba je + nedoločnik*, ki jih označujemo podobno kot 0:



5.4.7 Idiomi

Obravnava idiomov se ne razlikuje od obravnave običajnih stavčnih konstrukcij, kar pomeni, da pri označevanju pomensko komponento glagolske kolokacije namenoma prezremo: zveze *vzeti konec* ali *vzeti usodo v svoje roke* npr. označimo na enak način kot *vzeti kladivo v svoje roke*. Pri določanju oznak torej sledimo skladenjskim smernicam, pomagamo pa si s **zamenjavo** tako, da stalni del fraze nadomestimo s skladenjsko ustrežajočo nestalno varianto.

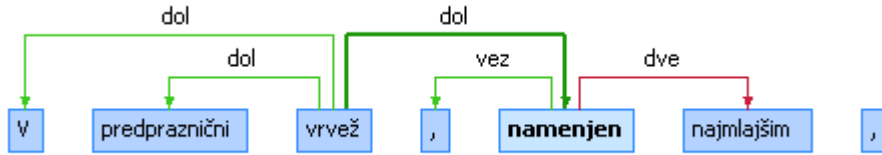


Posebna pozornost je potrebna pri variantah, kjer je pomembna **skladijska komponenta** fraze – pri zvezah *gre za seks in droge* in *gre za plot*, kjer pri prvi varianti kot izhodišče razumemo zvezo glagola in predloga *iti za*, v drugi pa glagol brez predloga (obenem pa je drugačen tudi pomen), varianti označujemo različno (*iti za koga/kaj?* -> **DVE** /*iti kam?* -> **ŠTIRI**). O tej temi glej še 0.

5.4.8 Polstavčni desni prilastki

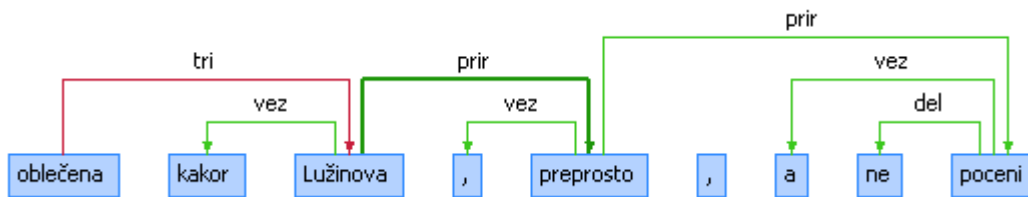
Polstavčne desne prilastke – pridevniške, deležniške, nedoločniške – tipa *gospa, oblečena kot Lužinova* (= *kakšna/katera gospa?*) označujemo podobno kot ostale tipe samostalniških besednih

zvez: med jedrom polstavka ter besedo, ki jo polstavek določa, označimo povezavo *dol*. V označevanje vključimo tudi vejico, ki jo povežemo z jedrom polstavka z razmerjem *vez*. Podobno kot pri priredjih (0) torej tudi tukaj iščemo trielementno razmerje:



5.4.9 Pojasnjevalne strukture

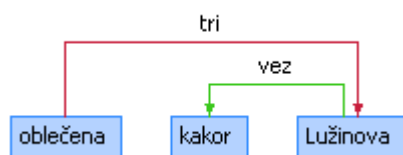
Pojasnjevalne strukture tipa *oblečena kakor Lužinova, preprosto, a ne poceni* interpretiramo kot pojasnjevalno priredje (*oblečena kakor Lužinova, IN SICER preprosto, a ne poceni*). Priredje označimo s povezavo *prir*, ki jo tukaj povežemo med prirednima deloma na podoben način, kot bi to storili pri klasičnem vezalnem priredju – v spodnjem primeru torej sopostavimo kot priredna dela *Lužinova* ter *preprosto*, ker na vprašanje *kako oblečena?* odgovarjamo *kakor Lužinova = preprosto, a ne poceni*.



Gornji primer prinaša tudi primer označevanja **protivnega priredja** – tudi slednje poteka analogno na označevanje vezalnega priredja. V zvezi *preprosto, a ne poceni* je torej priredni odnos med semantičnima nosilcema priredja, tj. *preprosto* in *poceni*, s tem da v označevalni trikotnik seveda pritegnemo tudi veznik, izhodiščni element razmerja pa je kot običajno prvi element priredja.

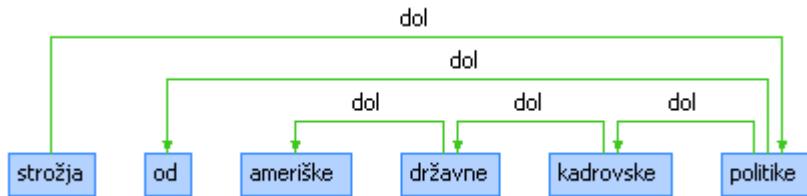
5.4.10 Primerjave

Primerjave tipa *oblečena kakor Lužinova* označujemo tako, kot je prikazano na spodnji sliki – primerjalni veznik označimo s povezavo *vez*, povezava med prvim ter drugim delom v primerjavi je tipično *TRI* (ker odgovarjamo na vprašanje *kako + prvi del primerjave*):



Na enak način označujemo tudi primere tipa *prijetneje kot v sprejemnici*, kjer bi se sicer lahko vprašali *prijetneje kot kje?* in posledično pripisali druge povezave med semantičnima nosilcema primerjave. Dogovorjeno je, da se v teh primerih vedno označuje s povezavo *TRI*.

Kadar označujemo primerjavo tipa *strožja od ameriške politike*, jo označujemo s povezavami *dol*, kakor je prikazano spodaj:

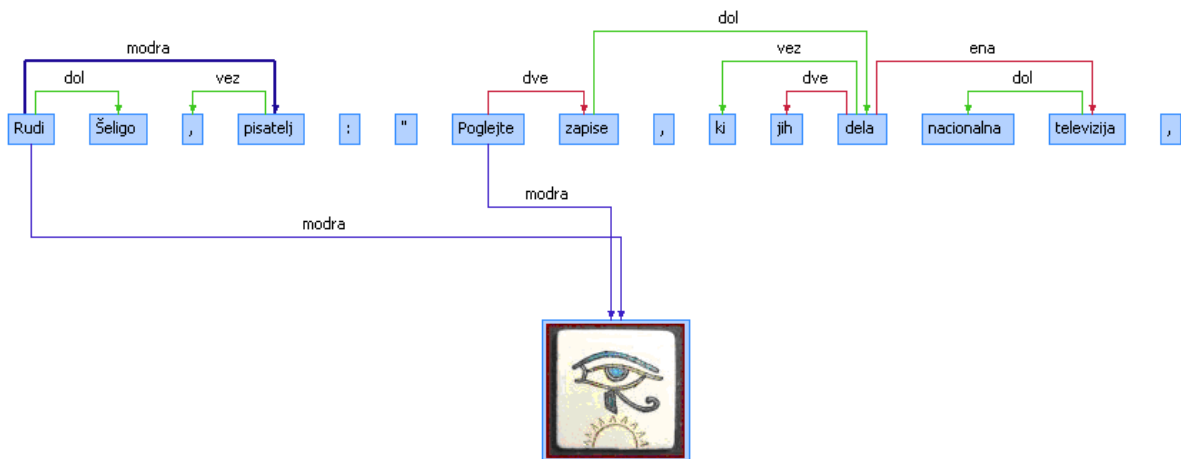


5.4.11 Premi govor

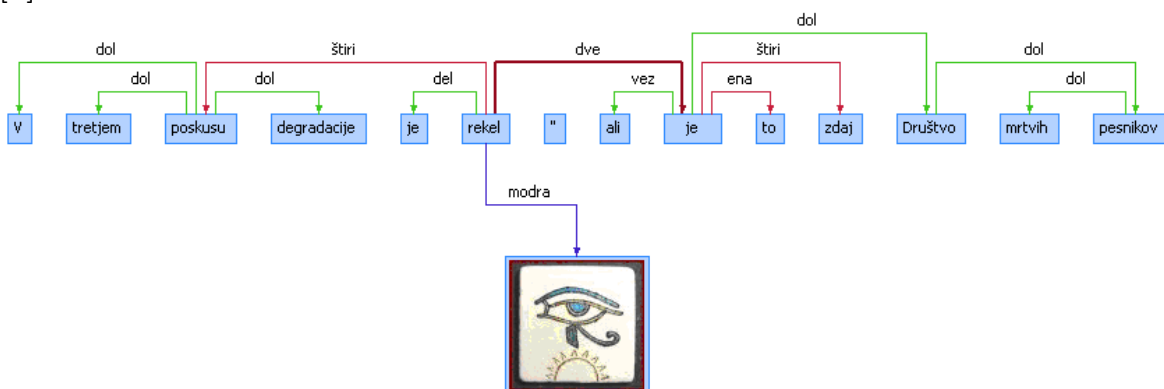
Za označevanje premege govora veljajo naslednja načela:

- narekovajev ne povezujemo v strukturo, enako velja za dvopičje,
- narekovaji in dvopičja sugerirajo stavčno mejo, ki jo v določenih primerih upoštevamo [A], v drugih ne [B], odvisno od tega, kakšno vlogo igra v stavku navedek (ali ga interpretiramo kot samostojno enoto ali npr. predmetni oz. osebkov odvisnik – *koga ali kaj je rekel?*)
- pika je vsekakor meja, prek katere ne označujemo, kar pomeni, da je pri daljših navedkih prvi del navedka npr. lahko predmetni odvisnik uvajalnemu stavku, preostale povedi navedka pa obravnavamo ločeno – kot da niso navedek.

[A]



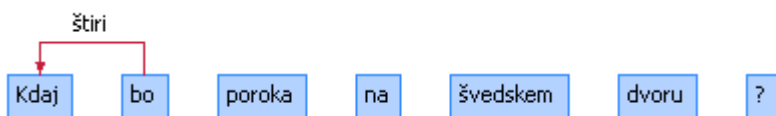
[B]



5.4.12 Vprašalnice

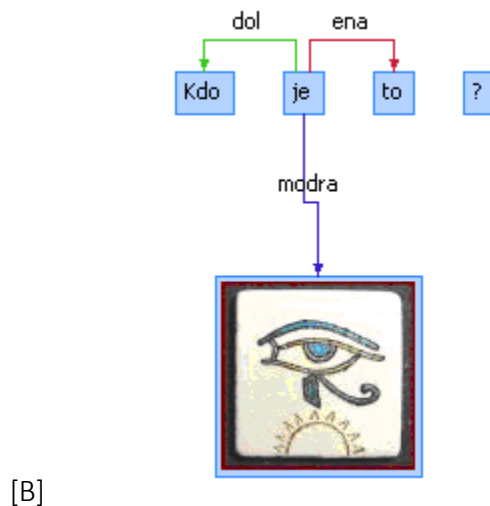
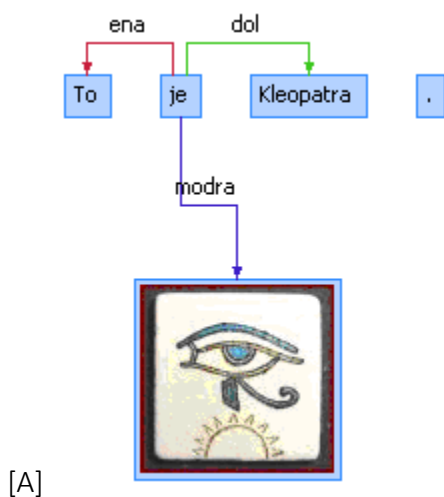
Prislove (*kako, kje, zakaj ...*) oz. zaimke (*kdo, kaj ...*), ki lahko nastopajo v vlogi vprašalnice ali v vezniški vlogi, označujemo glede njihovo **trenutno vlogo** v stavku. Vezniško vlogo prepoznamo po tem, da običajno uvaja podredni stavek (primer iz slovarja: *začutil je, kako mu nekaj leze po roki*). V tem primeru besedo *kako* povežemo s povezavo **vez**.

V primerih, ko te besede nastopajo kot vprašalnice (*kdaj bo poroka na švedskem dvoru?*), jih povezujemo s povezavami **drugega nivoja**, in sicer s tisto, ki pritiče stavčnemu členu, po katerem se z vprašalnico vprašamo:



Kadar imamo z določanjem tipa povezav težave, pretvorimo stavek v trdilnega oz. **nadomestimo vprašalnico z ustrezajočo nevprašalno besedo**, označimo povezave, nato te povezave prevedemo nazaj na problematični vprašalni stavek – v gornjem primeru smo pri označevanju dela stavka *kdaj bo poroka?* izhajali iz povezav, kakršne bi pripisali v primeru *poroka bo v nedeljo* – tukaj je s povezavo ENA torej povezana beseda *poroka*, zveza v *nedeljo* pa kot prislovno določilo časa s povezavo ŠTIRI.

Še primer s povedkovim določilom: problematični primer vprašalnega stavka *kdo je to?* pretvorimo v neproblematičnega *to je X*, označimo [A], prenesemo povezave nazaj [B]:

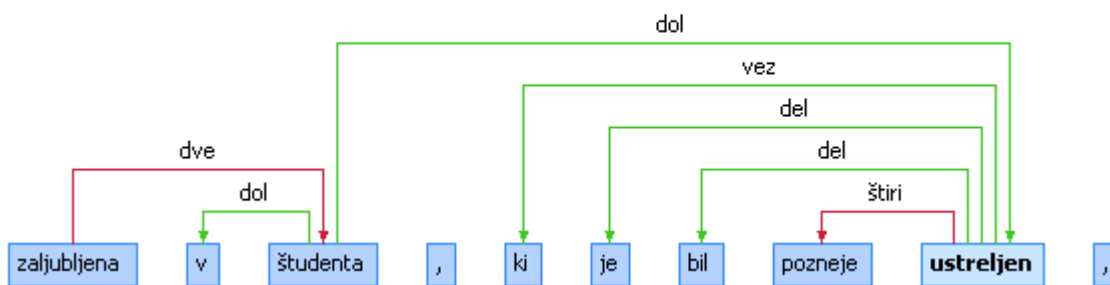


5.4.13 Pasivne konstrukcije

Kadar pri označevanju naletimo na pasivno konstrukcijo (npr. *študent, ki je bil pozneje ustreljen*), izhodišče povezav določimo analogno na tvorno obliko – v gornjem primeru je izhodišče povezav *del, vez, ŠTIRI* torej beseda *ustreljen* (glej primer spodaj).

Večji problem je **ločevanje dejanskih pasivnih oblik od deležnikov stanja**. Pomagamo si z vprašalnici: pasiv odgovarja na vprašanje *kaj so mu storili / kaj se mu je zgodilo*, medtem ko deležniki sporočajo *kakšno je bilo njegovo stanje*. Trik, s katerim si lahko pomagamo, je ta, da **časovne opredelitve** v stavku vedno nastopajo s pasivi – v primerih tipa *včeraj je bil zgrajen, v dveh letih bo zgrajen* gre torej za pasivno konstrukcijo in ne za deležnik stanja.

Če se po vsem tem (in posvetu z ostalimi označevalci) še vedno ne moremo odločiti, je varnejša izbira deležnik stanja (ker je pogostejši).

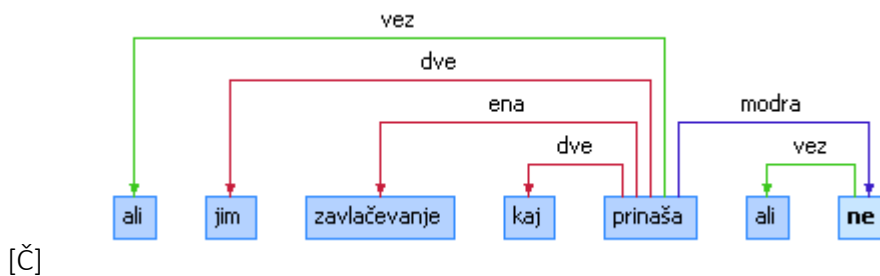
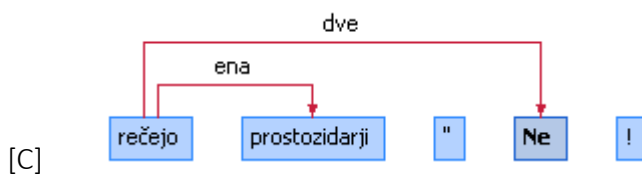
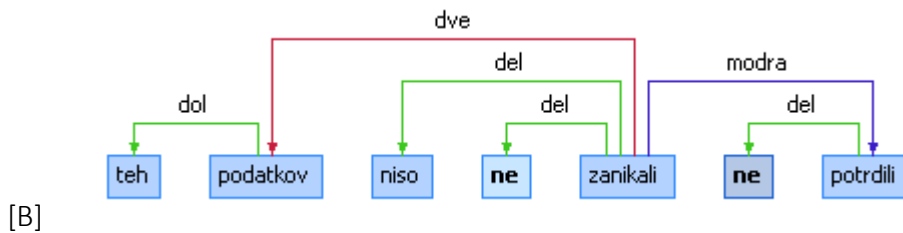
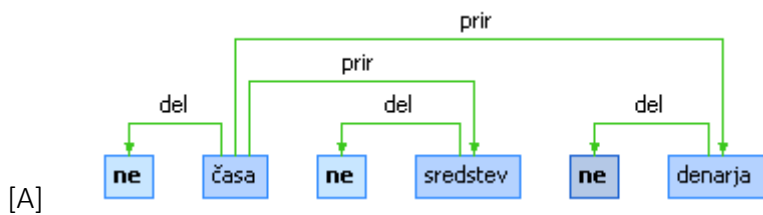


5.4.14 Členek NE¹²

Členek *ne* je problematičen, ker se v stavkih pojavlja na tako različne načine, da ga je težko konsistentno označevati. Smernice za označevanje so trenutno takšne:

- Kadar *ne zanika* sledečo besedo, med to besedo ter členkom pripišemo povezavo **del**. Pri zvezah tipa *ne – ne* (*ne lep ne grd*) med polnopomenskima enotama pripišemo povezavo *prir*. [A] Če sta zanikana glagola, med polnopomenskima enotama pripišemo povezavo *modra*, ker gre za nadstavčno razmerje. [B]
- Kadar stoji beseda *ne* v vlogi prepoznavnega stavčnega člena (*rekel je NE drogam in alkoholu – koga/kaj je rekel?*), ga temu primerno označimo s povezavo drugega nivoja. [C]
- Kadar nastopa *ne* v kombinaciji z elipso (*verjeli ali ne = verjeli ali ne verjeli*), ga obravnavamo, kakor da nadomešča izpuščeni del stavka – v primeru izpusta glagola ga torej povežemo s prvim glagolom s povezavo *modra*. [Č]

¹² Na enak način obravnavamo tudi ustrezne zveze z *ni*.



5.4.15 Slovnčne napake v besedilih

Z deli besedila, ki se ne skladajo z normo slovenskega jezika, se obremenjujemo čim manj. Označevanje nadaljujemo tako, kot da napake ne bi bilo (če gre le za odvečne vejice, jih ignoriramo ipd.), v zapletenejših primerih se posvetujemo z ostalimi označevalci.

Če označujemo nestandardni jezik, upoštevamo smernice, ki sledijo v naslednjem poglavju.

5.5 Označevanje nestandardne slovenščine

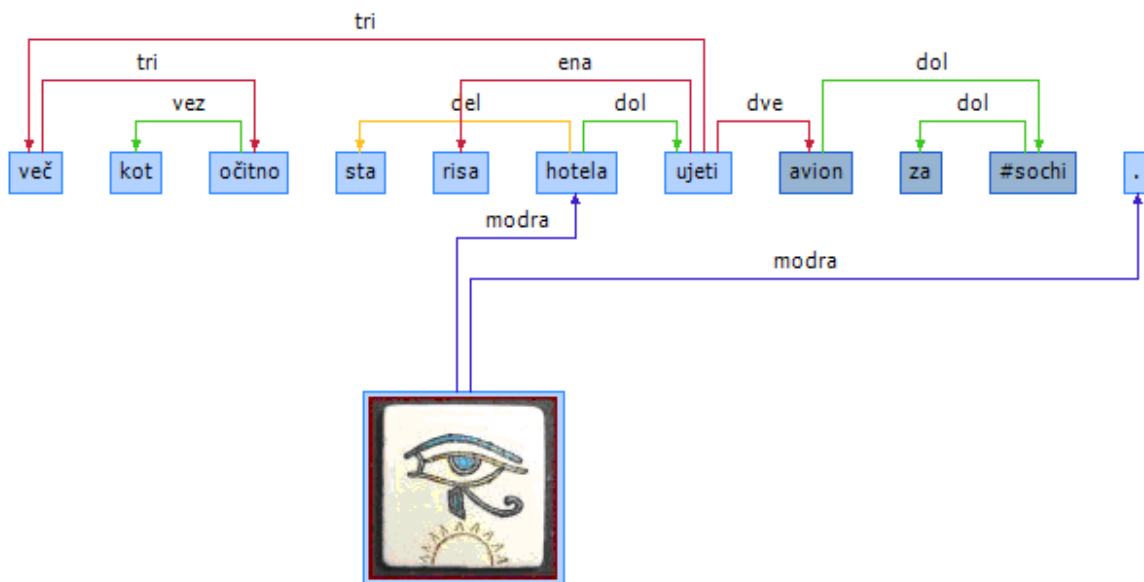
5.5.1 Spletna mesta, emotikoni, sklici, ključniki

V računalniško posredovani komunikaciji se pojavljajo specifični elementi, kot so navedbe spletnih mest, uporabniških imen, raba emotikonov in emojijev, ključniki (hashtags) in podobno. V skladenjskem smislu so ti elementi lahko dveh vrst:

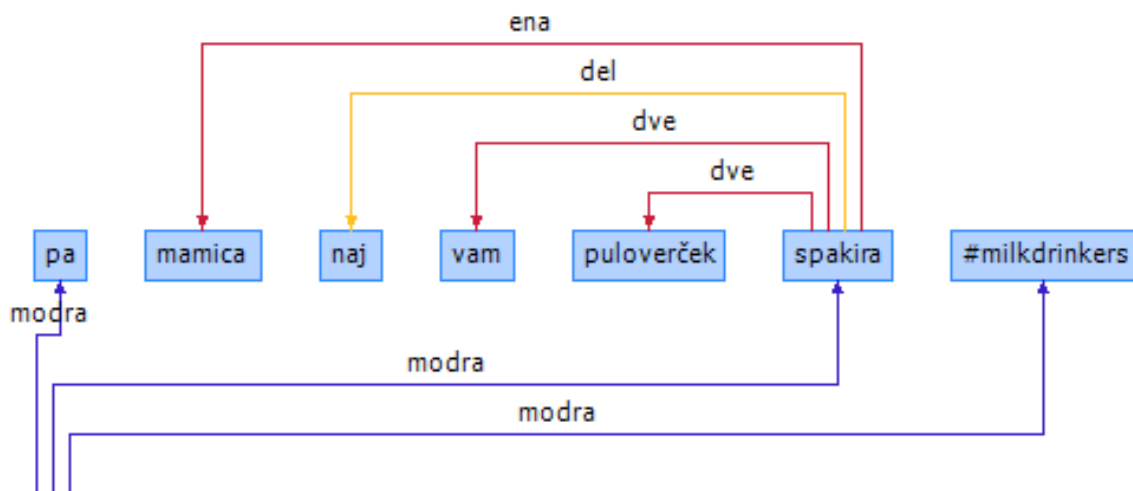
[A] bodisi so skladenjsko vpeti, kar pomeni, da tvorijo neizpušljivi del besedne zveze oz. igrajo vlogo stavčnega člena.

[B] Lahko pa so skladenjsko prosti, kar (npr. pri tvitih) običajno pomeni, da stojijo na začetku ali koncu sporočila in ne vstopajo v skladenjska razmerja s preostalimi deli stavka.

[A]

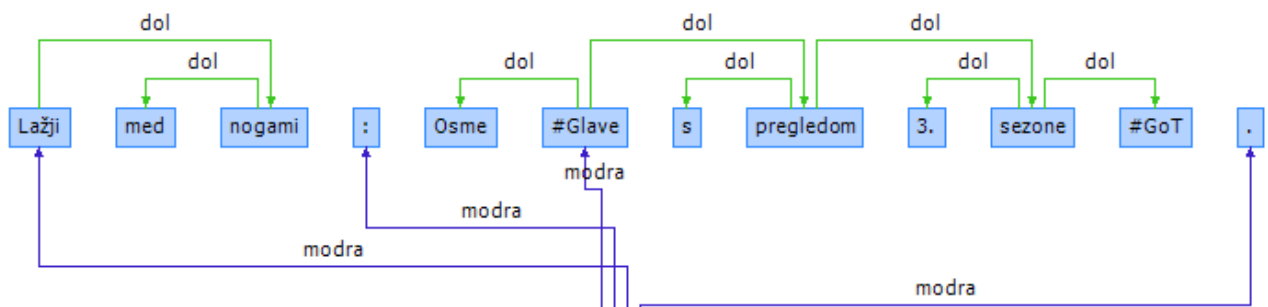


[B]



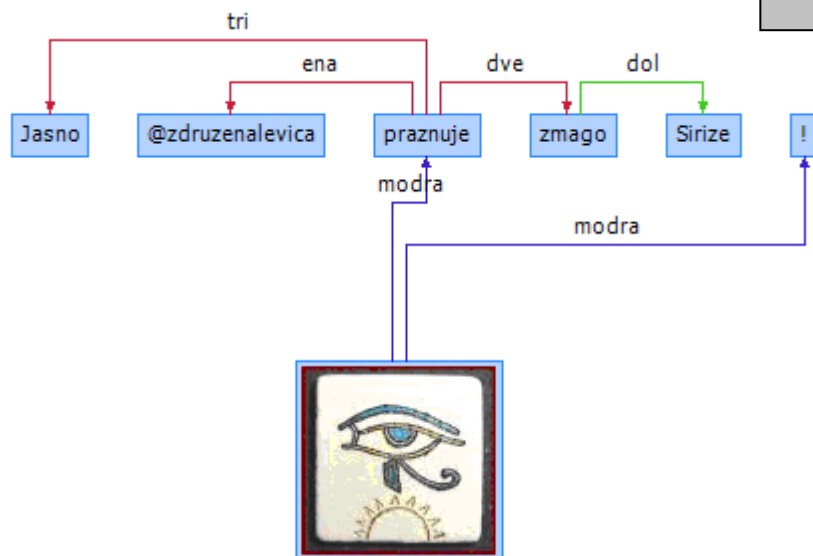
V drevesnico povezujemo samo tiste elemente, ki so v skladenjskem smislu neizpušljivi. Ostale primere povežemo na metaelement. Kaj razumemo kot izpušljivo in kaj ne, je v določeni meri odvisno od besedilnega žanra, ki ga označujemo. Pri označevanju tвитov so npr. po večini izpušljivi vsi emotikoni in emoji ter velika večina naslovov spletnih strani. Uporabniška imena, ki se nahajajo na začetku ali koncu tvita bi sicer lahko razumeli kot del stavka (nagovor), vendar je njihov glavni namen usmerjanje komunikacije (najava replik), zato jih ne povezujemo. Če pa so uporabniška imena vključena sredi stavka, v pomenskem smislu posebej izpostavljena in igrajo vlogo osebka ali predmeta, jih povežemo. Podobno velja za ključnike. Sledi nekaj primerov:

5.5.1.1 POVEZANO V DREVESNICO

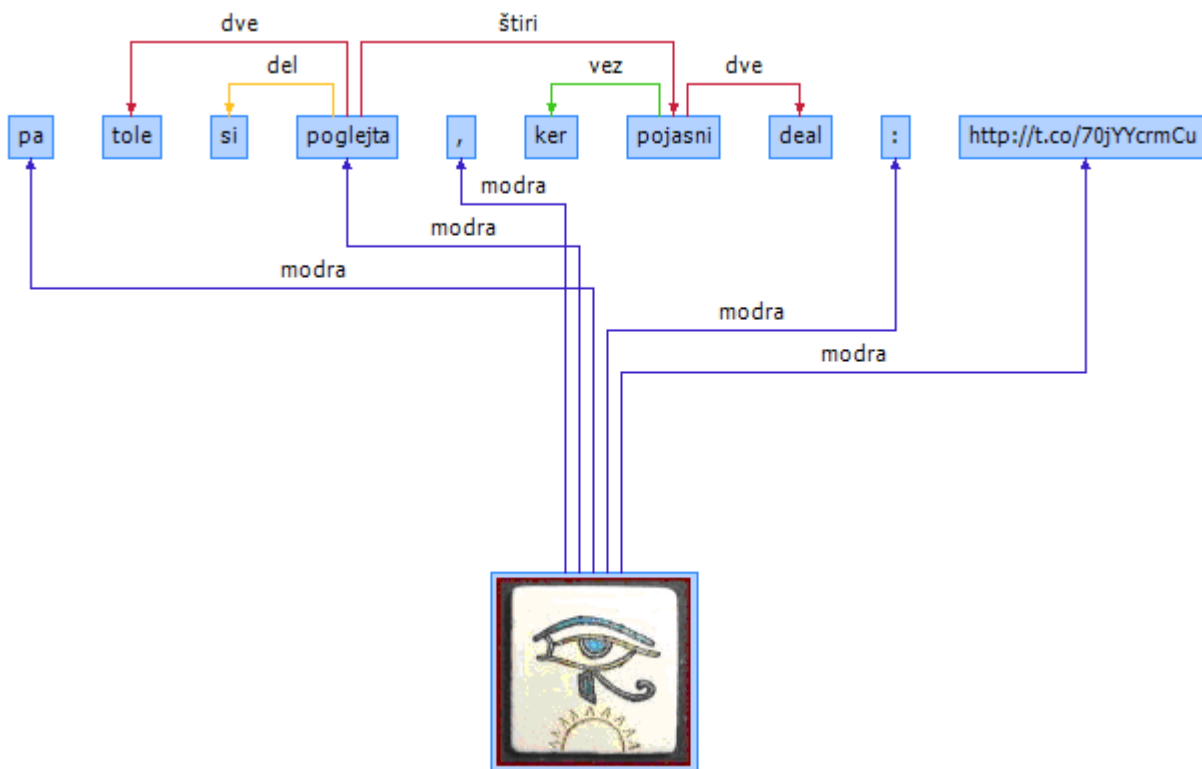
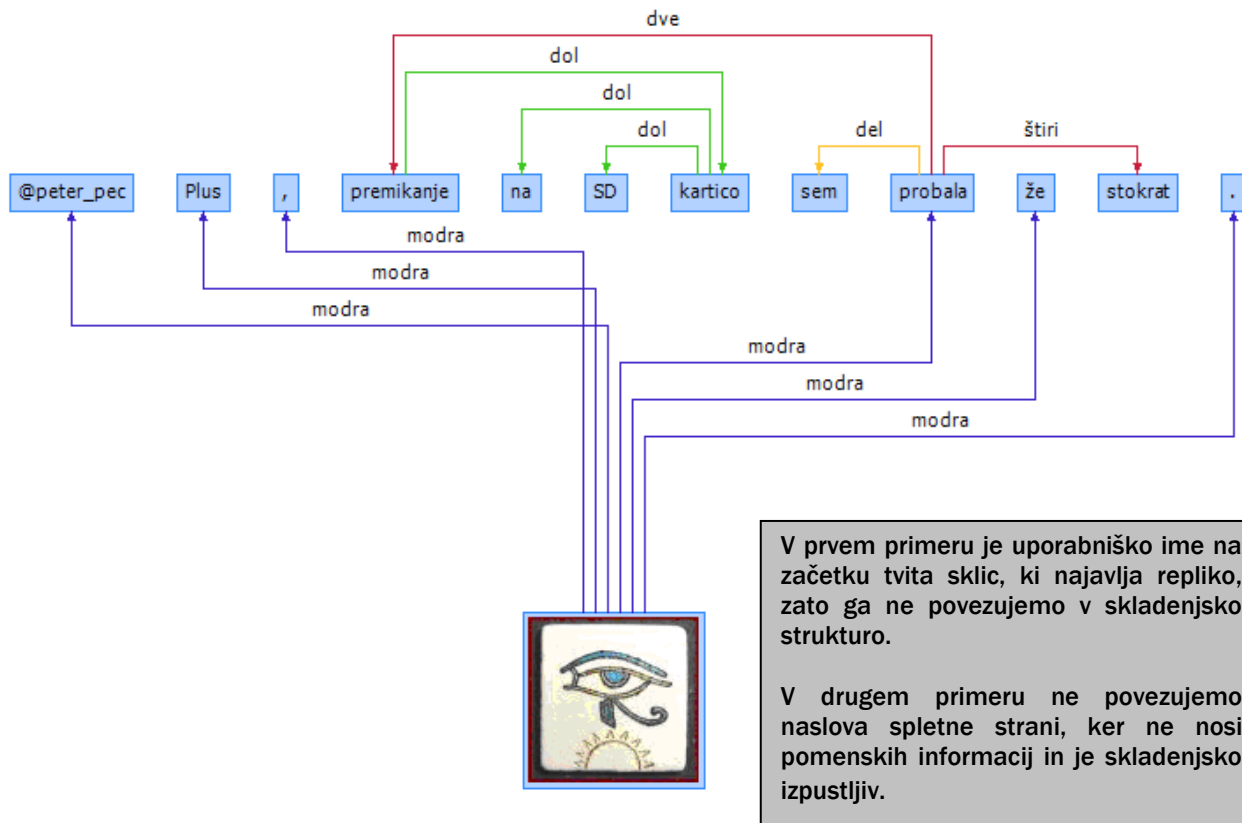


V prvem primeru ključniki nastopajo kot del besedne zveze (*Osme #Glave; 3. sezona #GoT*).

V drugem primeru uporabniško ime nastopa v vlogi osebka. Nahaja se na sredini stavka, nosi pomen, ki ni samo sklic na repliko.



5.5.1.2 POVEZANO NA METAELEMENT



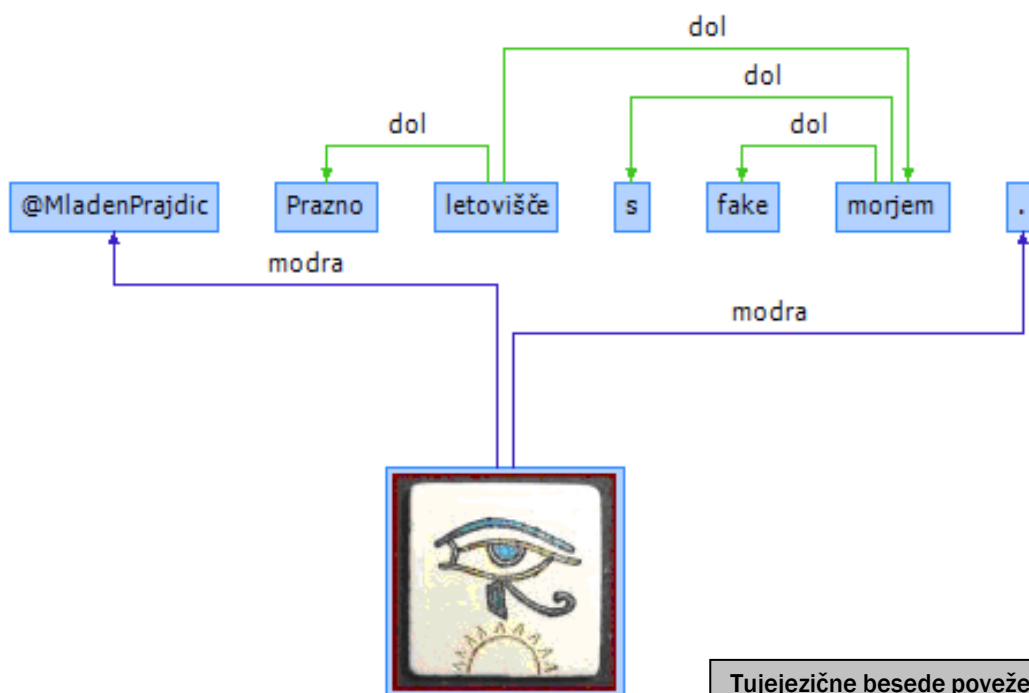
5.5.2 Tujejezični elementi

V nestandardnem jeziku se pojavlja višje število tujejezičnih jezikovnih fragmentov, kot to velja za standardni jezik. Kako ločevati »tujejezično« od »slovenskega« je izziv že za predhodne ravni označevanja (lematizacija, pripis MSD-oznak). Odločitev, kako jih bomo povezovali na skladenski ravni, je povezana z odločitvami na prejšnjih nivojih, ki se lahko od projekta do projekta razlikujejo.

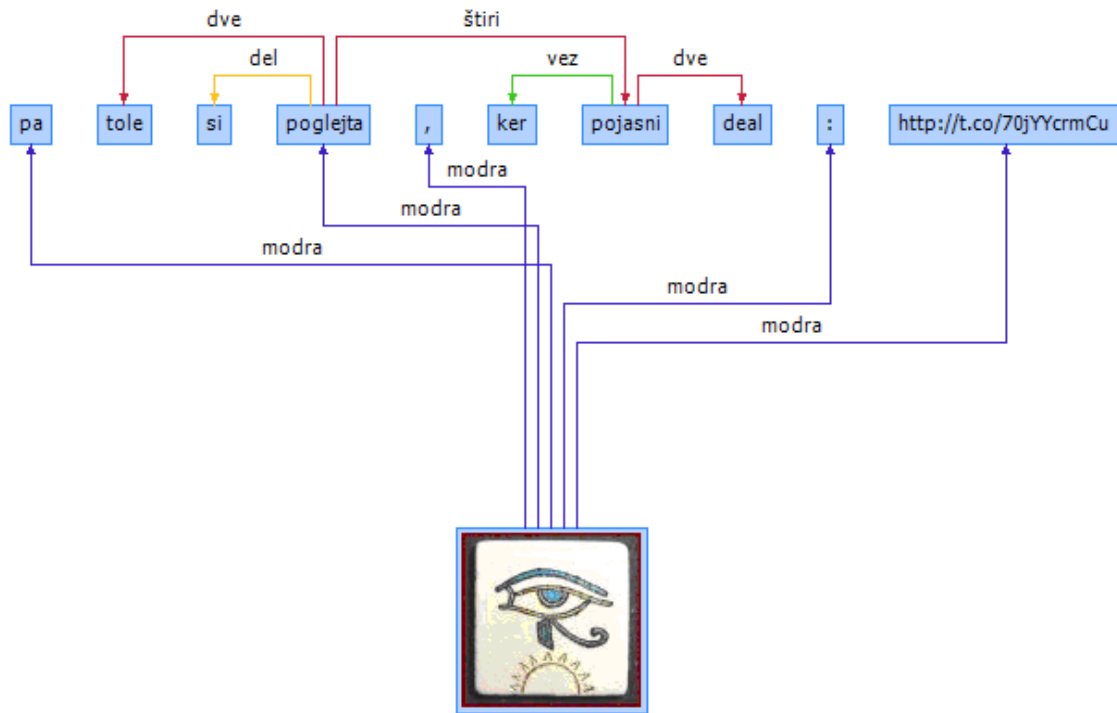
Vodila projekta Janes se ravnaajo glede na dolžino tujejezičnega niza:

- enobesedne pojavitve se povezujejo v drevesnico analogno s slovenskimi besedami,
- enako velja za dvodelne strukture z jasnim odvisnostnim razmerjem (pretežno zveze, ki so kombinacija dveh samostalnikov ali pridevnika in samostalnika),
- daljši segmenti v tujem jeziku se ne povezujejo v drevesnico, ampak se vse besede vežejo na metaelement.

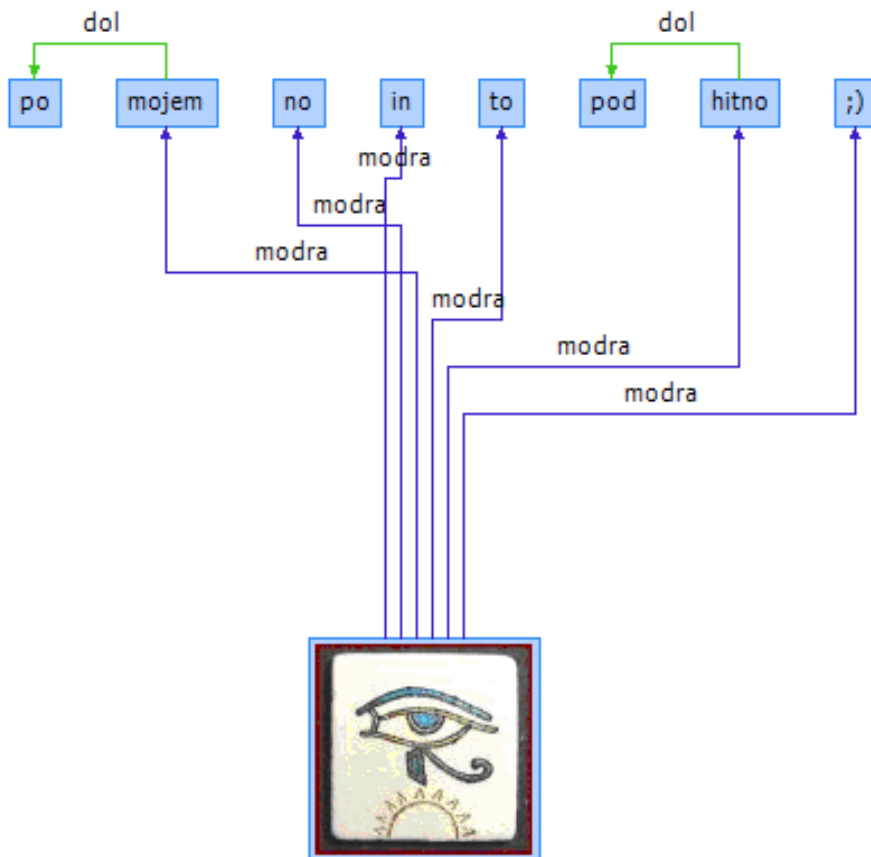
5.5.2.1 POVEZOVANJE POSAMEZNIH BESED

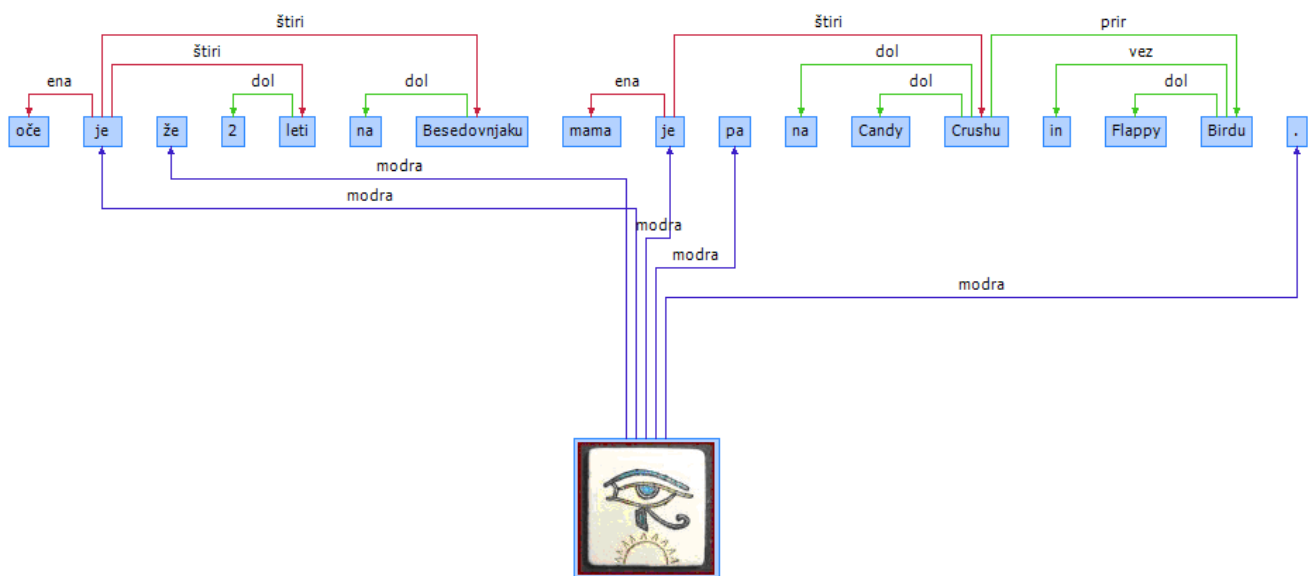
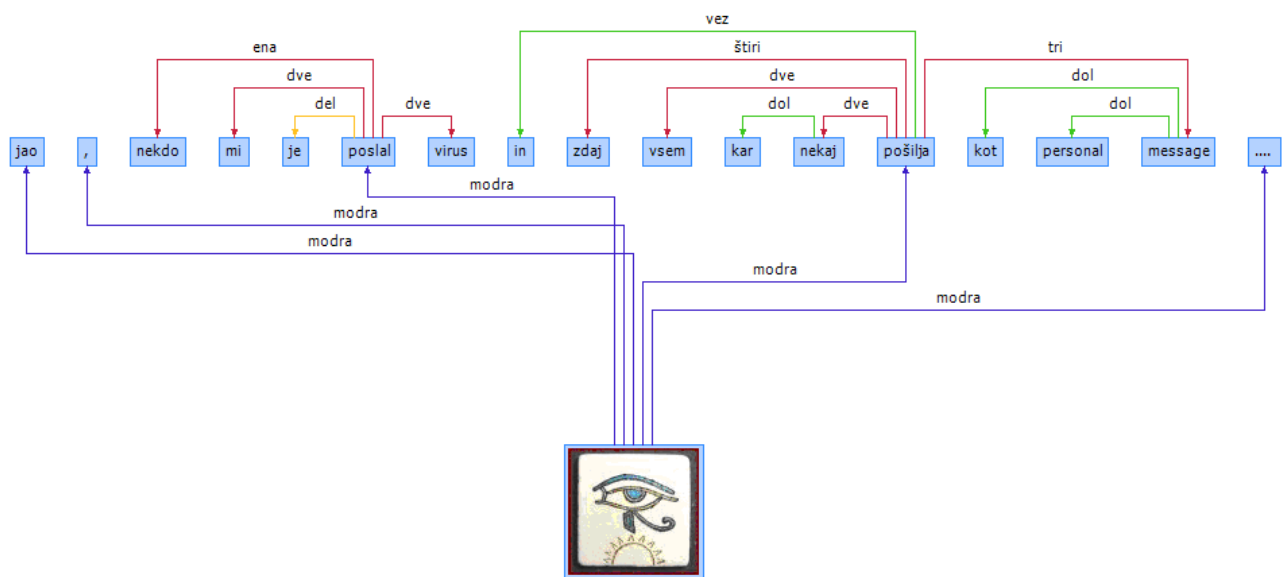


Tujejezične besede povežemo po analogiji s slovenskimi: *fake* povežemo kot pridevnik samostalnika *morje*; *deal* na drugem nivoju povežemo s povezavo *dve*, kot bi veljalo za slovenski samostalnik. Če naletimo na težave, se posvetujemo z drugimi označevalci.



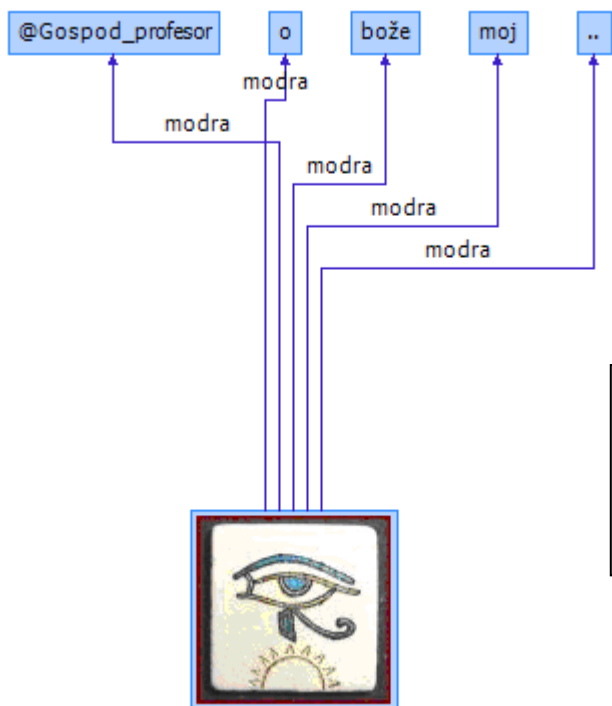
5.5.2.2 POVEZOVANJE DVOBESEDNIH ZVEZ



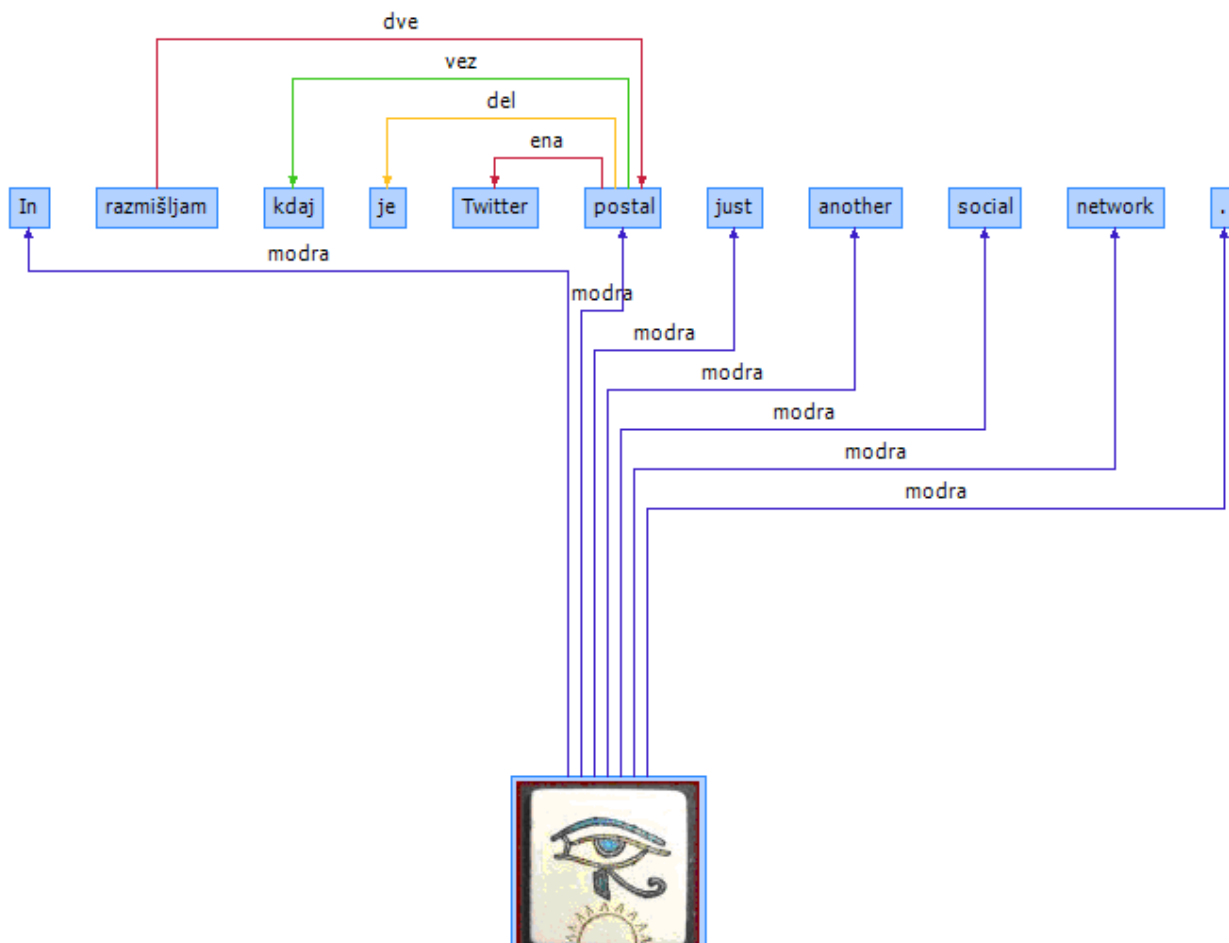


V odvisnostna razmerja povežemo zveze samostalnika s predlogom (*pod hitno*), pridevnika in samostalnika (*personal message*) in dve samostalnikov – v gornjem primeru gre za lastna imena (*Flappy Bird*, *Candy Crush*). Pri vseh tovrstnih kombinacijah gre za uporabo povezave *dol*. Če naletimo na primere, ki bi zahtevali drugačno obravnavo, se posvetujemo.

5.5.2.3 POVEZOVANJE DALJŠIH NIZOV

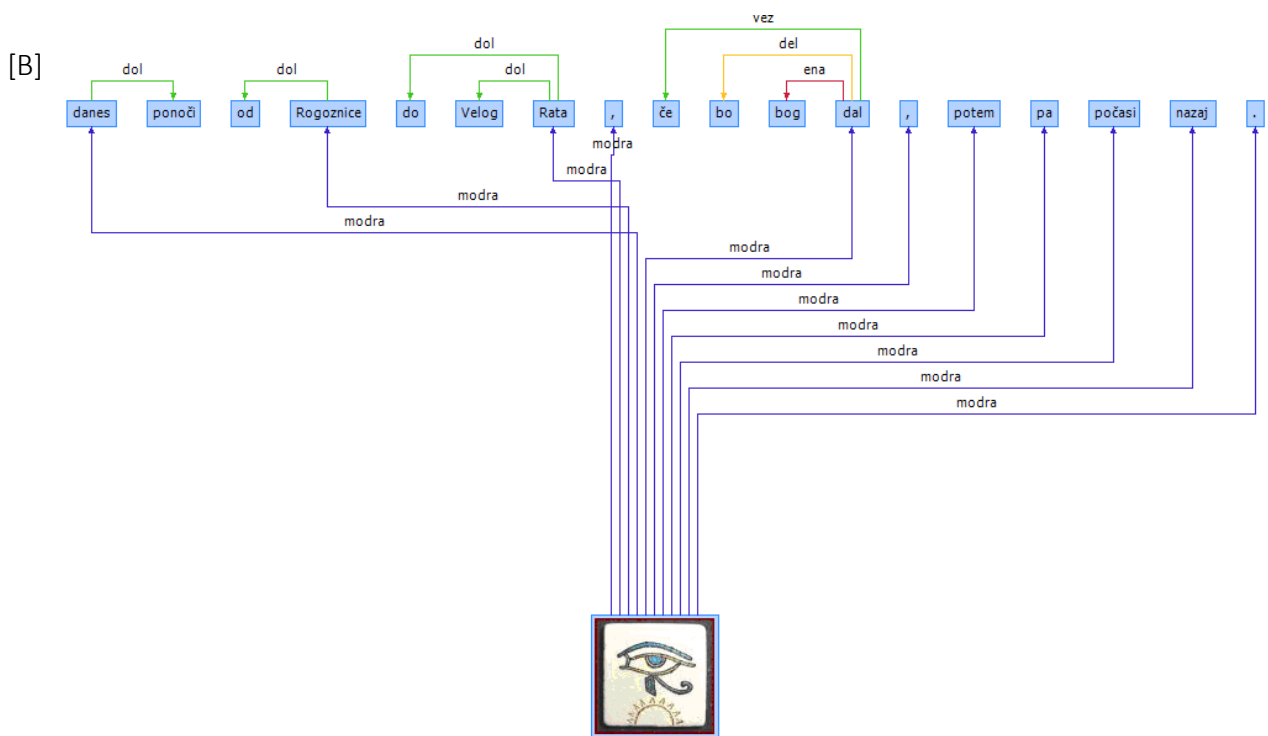
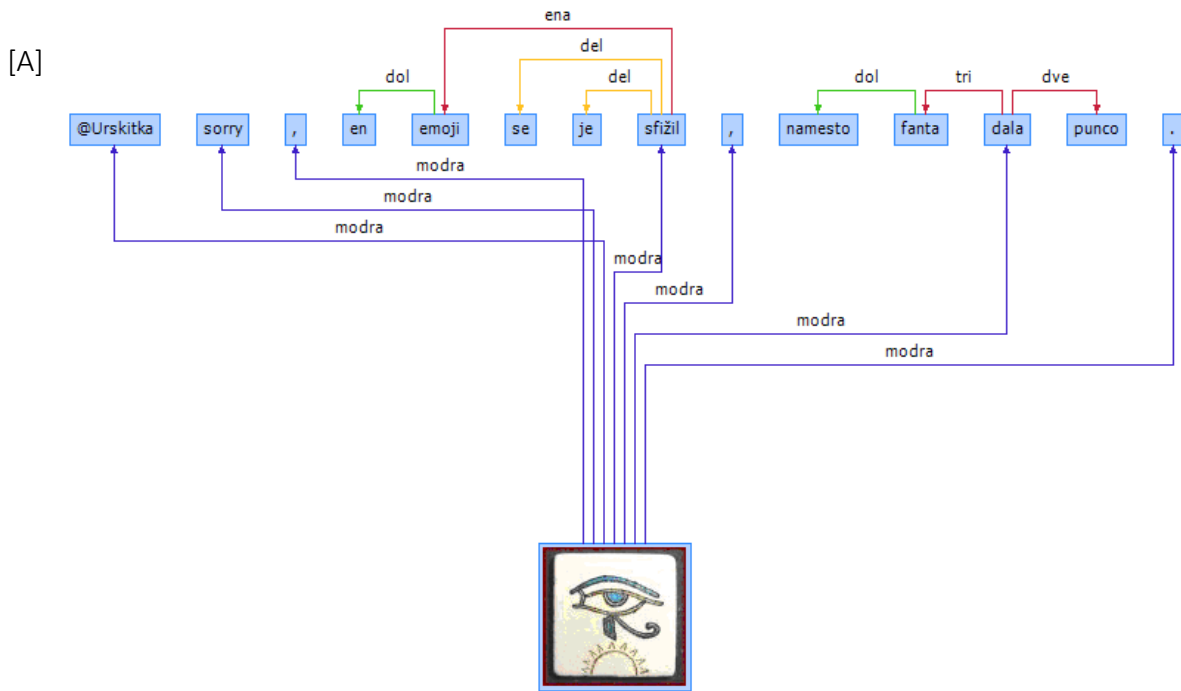


Pri fragmentih, ki so daljši in kompleksnejši, povežemo vse besede tujejezičnega niza na metaelement. Za tuje besede se štejejo tudi besede, ki so enakopisne slovenskim, vendar jih je mogoče interpretirati kot del tujejezičnega niza (kot npr. beseda *o* v nizu *o bože moj*).



5.5.3 Elipse in fragmenti

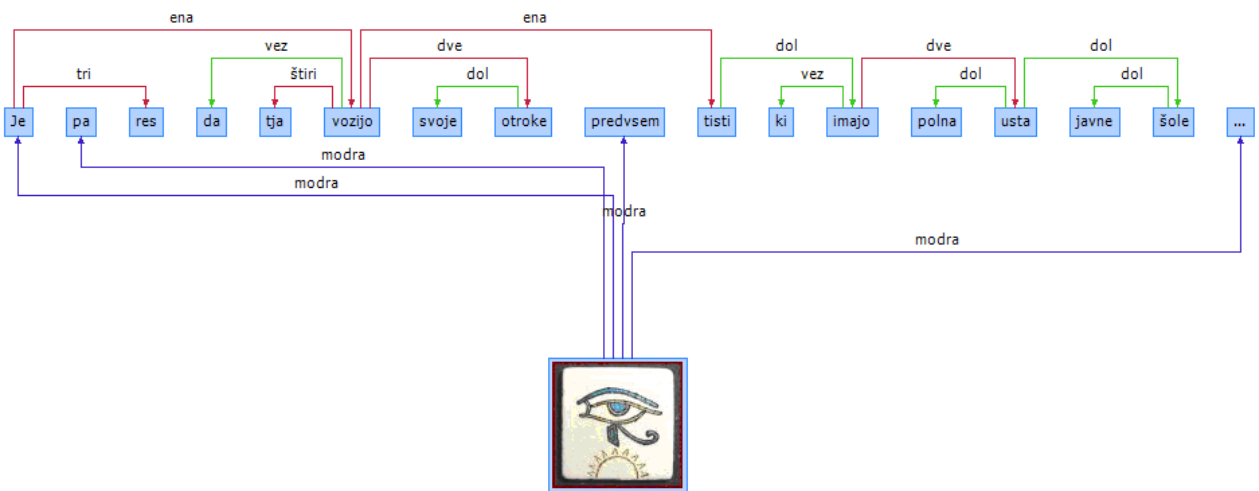
Za nestandardni jezik je značilna višja stopnja eliptičnosti in fragmentarnosti. Pri označevanju sledimo splošnemu načelu, da na prvi ravni povežemo besede v besedne zveze, na drugi pa stavčne člene. V primeru, da manjka del povedka, vežemo stavčne člene na del povedka, ki je na voljo [A]. V primeru, da manjka cel povedek, jedro vsakega od stavčnih členov povežemo na metaelement [B]. Če manjka kak drug stavčni člen, načeloma ni težav: povežemo, kar obstaja.



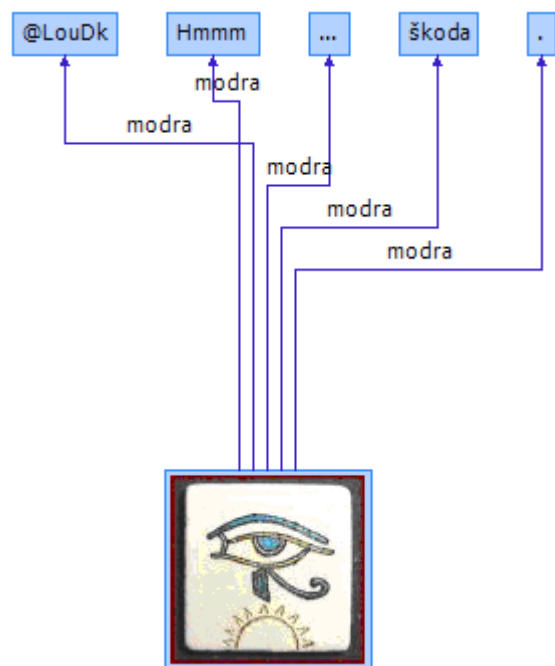
5.5.4 Nestandardna raba ločil

Čeprav napačna stava ločil, zlasti vejice, vpliva na natančnost avtomatskega razčlenjevanja, se pri označevanju oz. popravljanju označenih podatkov z ločili ne ukvarjamo. Če so vejice odveč, jih pustimo na miru, če manjkajo, jih ne dopolnjujemo [A]. Prav tako se ne ukvarjamo z drugimi nestandardnimi ločili, npr. pretirani rabi tropičja ali dveh pik namesto vejice [B]. Kar se tiče določanje meje povedi, sledimo segmentaciji, ki je bila določena na nižjih ravneh označevanja.

[A]



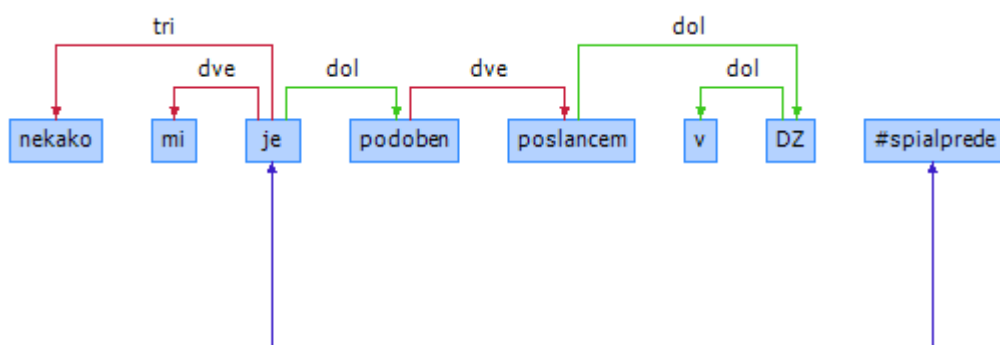
[B]



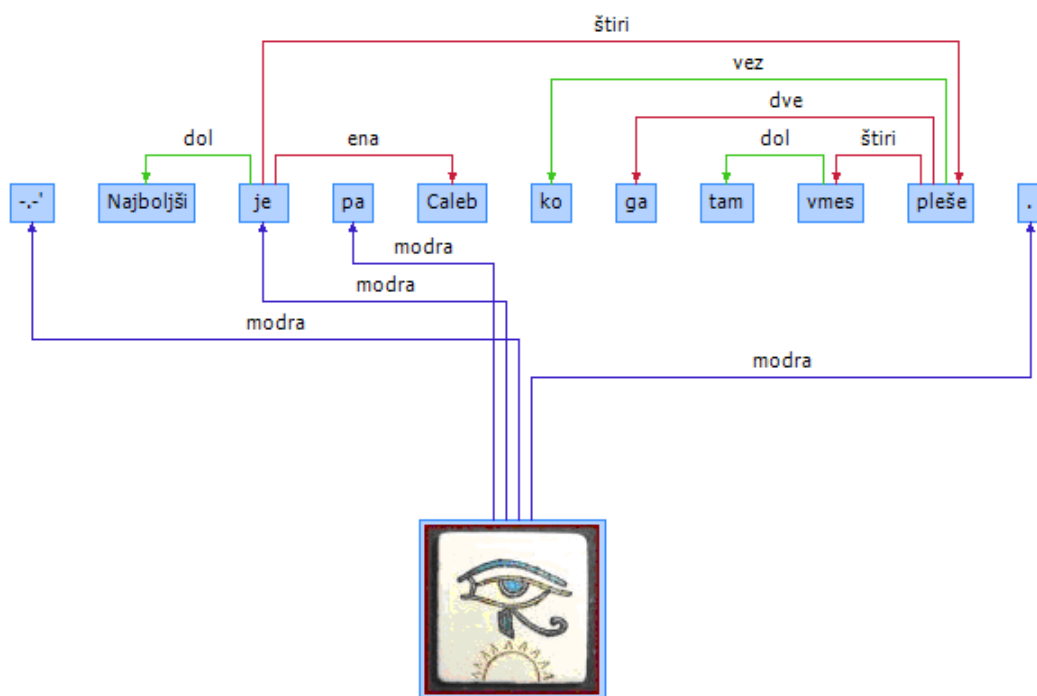
5.5.5 Druge nestandardne jezikovne značilnosti

Nestandardni jezik prinaša druge značilnosti, ki jih je mogoče opaziti pri skladijskem označevanju, npr. atipični besedni red, nestandardna raba veznikov, nestandardna raba slovničnega števila, sklonov, podspola živosti in določnosti, prisotnost jezikovnih elementov, tipičnih za govorno komunikacijo, in podobno. Vse te značilnosti pa ne vplivajo na način skladijskega označevanja.

V spodnjem primeru je prikazano označevanje povedi, v kateri se pojavita dva predmeta v dajalniku: nekako mi je podoben poslanecem v DZ: (I) podoben mi je (v pomeni 'podoben se mi zdi') in (II) podoben je poslanecem v DZ. Glede na smernice označimo dve povezavi vrste dve.



Še en primer je raba osebnega zaimka v kombinaciji z glagolom v zvezi plesati ga. Primer označimo glede na smernice in čimbolj analogno s pogostejšimi, tipičnimi primeri – besedo ga povežemo v strukturo s povezavo dve, kot kaže spodnja slika.



I. Priloga A – tabela bližnjic na tipkovnici

Namesto izbiranja imena povezave z visečega seznama lahko uporabljamo bližnjične tipke:

POVEZAVA	TIPKA
dol	o
del	e
vez	v
skup	s
prir	p
ena	1
dve	2
tri	3
štiri	4
modra	m

II. PRILOGA B – informativni seznam glagolov, ob katerih (ustrezne) samostalniške zveze označujemo z odnosom DVA, pridevniške pa z DOL

Simbolizacija	Primer
1. Sam ₁ Glag Sam ₁ /Prid ₁ Glagoli: <i>ostati, postati, izpasti, zdeti se, počutiti se, delati se;</i>	<i>Janez je postal smučarski učitelj/suh</i>
2. Sam ₁ Glag se kot Sam ₁ /Prid ₁ Glagoli: <i>odlikovati se kot, pokazati se kot, izkazati se kot;</i>	<i>Izkazala se je kot uspešna menedžerka</i>
3. Sam ₁ Glag se za Sam ₄ /Prid ₄ Glagoli: <i>izkazati se;</i>	<i>Izkazala se je za učinkovito;</i>
4. Sam ₁ /Sam ₄ / Glag za Sam ₄ /Prid ₄ Glagoli: <i>imeti, imeti se, veljati;</i>	<i>Velja za poštenjaka/za pošteno</i>
5. Sam ₁ Glag se Sam ₄ /Prid ₄ Glagoli: <i>čutiti se, delati se;</i>	<i>Čutila se je odgovorno</i>
6. Sam ₁ Sam ₄ Glag za Sam ₄ /Prid ₄ Glagoli: <i>izvoliti za, imenovati za, poklicati za, smatrati za, določiti za, spoznati za, sprejeti za, predlagati za, označiti za, vzeti za, jemati za, identificirati za, izglasovati za, potrditi za, predlagati za, izbrati za, priznati za;</i>	<i>Izberimo jo za vodičko</i>
7. Sam ₁ Sam ₄ Glag kot Sam ₄ Glagoli: <i>ceniti kot, poznati kot, identificirati kot, omeniti kot, sprejeti kot, označiti kot itd.;</i>	<i>Cenijo jo kot delavno</i>

(Simbolizacija: Sam = samostalnik; 1, 2, 3, 4, 5, 6 = sklon samostalnika; Prid = pridevnik; Glag = glagol)

III. PRILOGA C – informativni seznam zvez, ki jih povezujemo z odnosom

SKUP

brž ko
kakor/kar itd. + koli
kljub temu da
medtem ko
namesto da
navzlic temu da
ne le da
od kod
od koder
potem ko
prej ko
samo da
še ko
sedaj/zdaj ko
šele ko
tačas ko
tako da
takoj ko
takrat ko
toliko da
zatem ko

čim bolj
čim prej
do sedaj
kar se da
kdo ve + kdaj/koliko itd.
na glas
na videz
ne nazadnje
po navadi
res da
tako rekoč