



PROJEKT “SPORAZUMEVANJE V SLOVENSKEM JEZIKU” SPECIFIKACIJE ZA UČNI KORPUS

Operacijo delno financira Evropska unija iz Evropskega socialnega sklada ter Ministrstvo za šolstvo in šport. Operacija se izvaja 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.

Avtorji: 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



1. Uvod

Specifikacije za učni korpus projekta "Sporazumevanje v slovenskem jeziku" v celoti sledijo sistemu označevanja korpusov, razvitih v okviru projekta "Jezikoslovno označevanje slovenščine",¹ nabor oznak učnega korpusa SSJ, tako za oblikoslovno kot za skladijsko označevanje, pa bo v celoti kompatibilen z oznakami t. i. korpusa jos100k, ki bosta tako dopolnjujoča se in združljiva vira. Celoten učni korpus SSJ v obsegu 400.000 bo izločen iz enomilijonskega korpusa jos1M, kjer so oblikoslovne oznake že delno ročno preverjene. Proces izločanja učnega korpusa SSJ iz korpusa jos1M bo upošteval predvsem kriterij uporabnosti besedil (stavkov, odstavkov) za namen učenja statističnih označevalnikov in razčlenjevalnikov, med drugim razčlenjevalnika, ki bo razvit v okviru projekta SSJ. To pomeni, da se izločijo tisti deli besedil, ki so v procesu pretvobe iz izvorne oblike v XML zaradi različnih razlogov (kode html, radijski ali tv programi itd.) izkažejo kot neprimerni za učenje statističnih označevalnikov in razčlenjevalnikov. Proces izločanja izvajata dr. Tomaž Erjavec in Nanika Holz.

Učni korpus predstavlja en del jezikovnotehnološkega trojčka: učni korpus – leksikon besednih oblik – (oblikoslovni) označevalnik / (skladijski) razčlenjevalnik. Sistem označevanja v učnem korpusu je kompatibilen z oznakami v leksikonu besednih oblik, leksikon in ročno označeni korpus pa kot učno množico uporabljata obe orodji za označevanje in razčlenjevanje besedil. Izdelava leksikona in učnega korpusa je predpogoj za uspešno učenje statističnega označevalnika in razčlenjevalnika ter predstavlja prvi korak na poti do avtomatske razpoznave morfoloških in skladijskih pojavov v slovenščini.

Lastna imena se v učnem korpusu ne označujejo ročno, proces prepoznavanja lastnih imen (*Named Entity Recognition*) bo izpeljan v okviru izdelave označevalnika in razčlenjevalnika. Edini segment, ki se neposredno nanaša na prepoznavo lastnih imen, so večbesedna lastna imena, zabeležena na predviden način v leksikonu besednih oblik.

¹ <http://nl.ijs.si/jos/>



OBLIKOSLOVNO OZNAČEVANJE

Cilj tega dela projekta SSJ je ročno označiti 400.000 izbranih besed iz korpusa jos1M. Pri označevanju se uporablja prenovljeni nabor oblikoskladenjskih oznak za slovenščino JOS. Tabele so nevedene v nadaljevanju. 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 oblikoslovno označevanje bo uporabljeno orodje za označevanje Lexical Annotation Workbench 0.7,² ki je bilo razvito za označevanje korpusov v okviru projekta Praška odvisnostna drevesnica.³

2. Seznam besednih vrst in njihovih oznak

2.1. BESEDNE VRSTE

3.1.1. Imena in okrajšave

1	samostalnik	S
2	glagol	G
3	pridevnik	P
4	prislov	R
5	zaimek	Z
6	števnik	K
7	predlog	D
8	veznik	V
9	členek	L
10	medmet	M
11	okrajšava	O
12	neuvrščeno	N

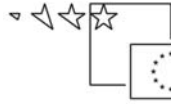
2.2. SAMOSTALNIK

3.2.1. Atributi in vrednosti

Table 2. Tabela atributov in vrednosti za samostalnik

² <http://www.ling.ohio-state.edu/~hana/law.html>

³ <http://ufal.mff.cuni.cz/pdt2.0/>



P	Atribut	Vrednost	K	Attribute	Value	C
0	besedna_vrsta	samostalnik	S	Category	Noun	N
1	vrsta	občno_ime	o	Type	common	c
		lastno_ime	l		proper	p
2	spol	moški	m	Gender	masculine	m
		ženski	z		feminine	f
		srednji	s		neuter	n
3	število	ednina	e	Number	singular	s
		dvojina	d		dual	d
		množina	m		plural	p
4	sklon	imenovalnik	i	Case	nominative	n
		rodilnik	r		genitive	g
		dajalnik	d		dative	d
		tožilnik	t		accusative	a
		mestnik	m		locative	l
		orodnik	o		instrumental	i
5	živost	ne	n	Animacy	no	n
		da	d		yes	y

3.2.1.1. Spremembe glede na FIDA/MULTEXT-East

1. Kategorija "živost" je prestavljena na 5. mesto

2.3. GLAGOL

3.3.1. Atributi in vrednosti

Table 3. Tabela atributov in vrednosti za glagol

P	Atribut	Vrednost	K	Attribute	Value	C
0	besedna_vrsta	glagol	G	Category	Verb	V
1	vrsta	glavni	g	Type	main	m
		pomožni	p		auxiliary	a
2	vid	dovršni	d	Aspect	perfective	e
		nedovršni	n		imperfective	p



		dvovidski	v		biaspectual	b
3	oblika	nedoločnik	n	Form	infinitive	n
		namenilnik	m		supine	u
		deležnik	d		participle	p
		sedanjik	s		present	r
		prihodnjik	p		future	f
		pogojnik	g		conditional	c
		velelnik	v		imperative	m
4	oseba	prva	p	Person	first	1
		druga	d		second	2
		tretja	t		third	3
5	število	ednina	e	Number	singular	s
		množina	m		plural	p
		dvojina	d		dual	d
6	spol	moški	m	Gender	masculine	m
		ženski	z		feminine	f
		srednji	s		neuter	n
7	nikalnost	nezanikani	n	Negativ	no	n
		zanikani	d		yes	y

3.3.2. Spremembe glede na FIDA/MULTEXT-East

1. Lastnost "vrsta":
 1. "polnopomenski" in "modalni" postenejo "glavni"
 2. "vezni" se preimenuje v "pomožni".
2. Ukinjena je kategorija "čas" (bivša št. 3)
 1. sedanja in prihodnja obl. glagola se prestavita v gl. obliko (št. 2)
 2. povednik se ukinja in namesto njega uporabi sedanjo obl. glagola.
3. Ukinjena je kategorija "način" (bivša št. 7)
 1. deležniki z lastnostjo "tvornik" ostanejo
 2. deležniki z lastnostjo "trpni_deležnik" se spremenijo v pridevnike vrste "deležniški".
4. Kategorija "vid" se prestavi na št. 2 (iz št. 14) in dobi lastnost "dvovidski".
5. Vse kategorije od kategorije "oblika" se prestavijo za mesto višje.



2.4. PRIDEVNIK

3.4.1. Lastnosti in vrednosti

Table 4. Tabela atributov in vrednosti za pridevnik

P	Atribut	Vrednost	K	Attribute	Value	C
0	besedna_vrsta	pridevnik	P	Category	Adjective	A
1	vrsta	splošni	p	Type	general	g
		svojilni	s		possessive	s
		deležniški	d		participle	p
2	stopnja	nedoločeno	n	Degree	undefined	u
		primernik	p		comparative	c
		presežnik	s		superlative	s
3	spol	moški	m	Gender	masculine	m
		ženski	z		feminine	f
		srednji	s		neuter	n
4	število	ednina	e	Number	singular	s
		dvojina	d		dual	d
		množina	m		plural	p
5	sklon	imenovalnik	i	Case	nominative	n
		rodilnik	r		genitive	g
		dajalnik	d		dative	d
		tožilnik	t		accusative	a
		mestnik	m		locative	l
		orodnik	o		instrumental	i
6	določnost	ne	n	Definiteness	indefinite	i
		da	d		definite	d

3.4.2. Spremembe glede na FIDA/MULTEXT-East

1. Pridevniki z lastnostjo "kakovostni" in "vrstni" postanejo "splošni".
2. Kategorija vrsta dobi lastnost "deležniški".
3. Ukine se kategorija "živost".
4. Pri kategoriji "stopnja" gre "elativ" v "nedoločeno".

2.5. PRISLOV

3.5.1. Lastnosti in vrednosti



Table 5. Tabela atributov in vrednosti za prislov

P	Atribut	Vrednost	K	Attribute	Value	C
0	besedna_vrsta	prislov	R	Category	Adverb	R
1	stopnja	nedoločeno	n	Degree	undefined	u
		primernik	r		comparative	c
		presežnik	s		superlative	s
2	deležje	ne	n	Participle	no	n
		da	d		yes	y

3.5.2. Spremembe glede na FIDA/MULTEXT-East

1. Ukine se kategorija "vrsta".
2. Pri kategoriji "stopnja" gre "osnovnik" v nedoločeno.
3. Pri kategoriji "stopnja" gre "elativ" v nedoločeno.
4. Uvede se kategorija "deležje", glede podrobnosti glej poglavje 3.5.

2.6. ZAIMEK

3.6.1. Lastnosti in vrednosti

Table 6. Tabela atributov in vrednosti za zaimek

P	Atribut	Vrednost	K	Attribute	Value	C
0	besedna_vrsta	zaimek	Z	Category	Pronoun	P
1	vrsta	osebni	o	Type	personal	p
		svojilni	s		possessive	s
		kazalni	k		demonstrative	d
		oziralni	z		relative	r
		povratni	p		reflexive	x
		celostni	c		general	g
		vprašalni	v		interrogative	q
		nedoločni	n		indefinite	i
		nikalni	l		negative	z
2	oseba	prva	p	Person	first	1
		druga	d		second	2
		tretja	t		third	3
3	spol	moški	m	Gender	masculine	m
		ženski	z		feminine	f



		srednji	s		neuter	n
4	število	ednina	e	Number	singular	s
		dvojina	d		dual	d
		množina	m		plural	p
5	sklon	imenovalnik	i	Case	nominative	n
		rodilnik	r		genitive	g
		dajalnik	d		dative	d
		tožilnik	t		accusative	a
		mestnik	m		locative	l
		orodnik	o		instrumental	i
6	število_svojine	ednina	e	Owner_Number	singular	s
		dvojina	d		dual	d
		množina	m		plural	p
7	spol_svojine	moški	m	Owner_Gender	masculine	m
		ženski	z		feminine	f
		srednji	s		neuter	n
8	oblika	klitični	k	Form	clitic	c
		navezni	z		bound	b

3.6.2. Spremembe glede na FIDA/MULTEXT-East

1. Ukinejo se vse kategorije od 8. naprej, vključno z živostjo.
2. Pri kategoriji 8. se poimenovanje spremeni v "oblika" in lastnosti "da" in "ne" zamenjata s "klitični" in "navezni"

2.7. ŠTEVNIK

3.7.1. Lastnosti in vrednosti

P	Atribut	Vrednost	K	Attribute	Value	C
0	besedna_vrsta	števniki	K	Category	Numeral	M
1	zapis	arabski	a	Form	digit	d
		rimski	r		roman	r
		besedni	b		letter	l



2	vrsta	glavni	g	Type	main	m
		vrstilni	v		ordinal	o
		zaimkovni	z		pronominal	p
		drugi	d		special	s
3	spol	moški	m	Gender	masculine	m
		ženski	z		feminine	f
		srednji	s		neuter	n
4	število	ednina	e	Number	singular	s
		dvojina	d		dual	d
		množina	m		plural	p
5	sklon	imenovalnik	i	Case	nominative	n
		rodilnik	r		genitive	g
		dajalnik	d		dative	d
		tožilnik	t		accusative	a
		mestnik	m		locative	l
		orodnik	o		instrumental	i
6	določnost	ne	n	Definiteness	no	n
		da	d		yes	y

3.7.2. Spremembe glede na FIDA/MULTEXT-East

1. Pod kategorijo "vrsta" se doda lastnost "zaimkovni".
2. Ukine se vrednost "množilni" pod kategorijo "vrsta", množilni števniki dobijo vrednost "drugi".
3. Ukine se kategorija "živost".
4. Kategorija "zapis" se prestavi na prvo mesto. Ostale kategorije se pomaknejo za mesto višje.

2.8. PREDLOG

3.8.1. Lastnosti in vrednosti

Table 8. Tabela atributov in vrednosti za predlog

P	Atribut	Vrednost	K	Attribute	Value	C
0	besedna_vrsta	predlog	D	Category	Preposition	S
1	sklon	imenovalnik	i	Case	nominative	n
		rodilnik	r		genitive	g
		dajalnik	d		dative	d



		tožilnik	t		accusative	a
		mestnik	m		locative	l
		orodnik	o		instrumental	i

3.8.2. Spremembe glede na FIDA/MULTEXT-East

1. Ukineta se kategoriji "vrsta" in "sestavljenost".

2.9. VEZNIK

3.9.1. Lastnosti in vrednosti

P	Atribut	Vrednost	K	Attribute	Value	C
0	besedna_vrsta	veznik	V	Category	Conjunction	C
1	vrsta	priredni	p	Type	coordinating	c
		podredni	d		subordinating	s

3.9.2. Spremembe glede na FIDA/MULTEXT-East

1. Ukine se kategorija "oblika".

2.10. ČLENEK

3.10.1. Lastnosti in vrednosti

P	Atribut	Vrednost	K	Attribute	Value	C
0	besedna_vrsta	členek	L	Category	Particle	Q

2.11. MEDMET

3.11.1. Lastnosti in vrednosti

P	Atribut	Vrednost	K	Attribute	Value	C
---	---------	----------	---	-----------	-------	---



0	besedna_vrsta	medmet	M	Category	Interjection	I
---	---------------	--------	---	----------	--------------	---

2.12. OKRAJŠAVA

3.12.1. Lastnosti in vrednosti

Table 12. Tabela atributov in vrednosti za okrajšava

P	Atribut	Vrednost	K	Attribute	Value	C
0	besedna_vrsta	okrajšava	O	Category	Abbreviation	Y

2.13. NEUVRŠČENO

3.13.1. Lastnosti in vrednosti

Table 13. Tabela atributov in vrednosti za neuvrščeno

P	Atribut	Vrednost	K	Attribute	Value	C
0	besedna_vrsta	neuvrščeno	N	Category	Residual	X
1	vrsta	tujejezično	j	Type	foreign	f
		tipkarska	t		typo	t
		program	p		program	p



3. Splošna načela za označevanje

3.1. Enoznačnost označevanja

Vsaka pojavnica v korpusu dobi le eno oznako. V primerih, kjer ni mogoče določiti neke lastnosti zato, ker referent besedni obliki ne obstaja, iz okolice ni mogoče ugotoviti, za katero lastnosti gre, oblikoslovno pa je možnih več lastnosti, pripišemo:

1. spol: a. srednji spol v primerih, ko gre v osnovi za pridevniške pojavnice (pridevniki, zaimkovni števniki, pridevniški zaimki):

*Iskreno ozračje groze , ki resnično krepi občutek **neznanega [Ppnser]** in skrivnostnega, , je ... v odprtem pismu slovenski javnosti je med **drugim [Kbzseo]** zapisala ... In kljub **temu [Zk-sed]** potem ni šla delat, temveč je prekinila svoje strokovno zdravniško izobraževanje*

b. moški spol v primerih, kjer gre za posamostaljene pridevnike

*Poudarek bo veljal razvoju gospodarstva, da bo **mladim [Ppnmmd]** omogočena zaposlitev.*

2. število: množina, če ni mogoče zanesljivo ugotoviti, ali gre za dvojino ali množino

*Ni razloga, da bi sodišče nadaljevalo postopek, za katerega nobena od **strank [Sozmr]** ne izkaže pravega interesa.*

3. če gre za pravo pomensko dvoumnost, pripišemo tisto vrednost, ki se nam po najboljši presoji zdi verjetnejša:

*Če ribe ne pokažejo zanimanja za muho ali blestivko, nam za **asa [Somdt]** v rokavu ostaneta rakec ali črv.*

*Preden pa se kdo preveč razgreje, bi želel napisati nekaj pojasnil. Nikakor ne zagovarjam vožnje zunaj utrjenih cest, da bi se približali objektu. To je ne le neetično, temveč tudi kaznivo. Vendar obstaja v Sloveniji dovolj makadamskih cest, pa tudi ob avtocesti je moč odkriti kako kanjo, ki ne bo zletela takoj, ko se bomo ustavili. Nazadnje pa bi rad napisal, da ne smemo pozabiti, da vozimo avto in smo dolžni upoštevati vse cestnoprometne predpise. To je sicer popolnoma očitno, a pri iskanju **živali [Sozmr]** lahko prehitro pozabimo na cesto in si uničimo prelep dan v naravi s prometno nesrečo.*

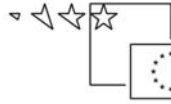
*V teh letih je tudi pravi čas, da uredite vsakršne hormonalne težave, ki lahko vplivajo na plodnost, na primer **sindrom [Somei]** policističnih ovarijev.*

3.2. Označevanje samostalnikov

Občna imena

- glagolnike označujemo kot samostalnike

*Tudi nasvet o **soljenju [Sosem]** zelja za kisanje sem povzela s seminarja Pridelovanje in kisanje vrtnin.*



- posamostaljene pridevnike označujemo kot pridevnike, ne kot samostalnike.

Poudarek bo veljal razvoju gospodarstva, da bo mladim [Ppnmmd] omogočena zaposlitev.

- živost označujemo samo pri samostalnikih moškega spola, kadar se pojavljajo v ednini v tožilniku
- kot občni samostalnik označujemo tudi citatno pisane samostalnice iz drugih jezikov. Držimo se načela: karkoli se sklanja kot samostalnik, uvrstimo med samostalnike.

Nastale so klasike Rozmarijin otrok, Izganjalec duhov in slasher [Somej] Noč čarovnic.

- kot občna označujemo kratična imena tipa WAP, GPRS, USB

Lastna imena – osebna ter zemljepisna

- kadar je v besedilu samostalnik, ki ga prepoznamo (iz oblike, konteksta, splošnega znanja) za osebno ali zemljepisno ime, ga vedno označimo kot lastno, pripišemo pa mu lemo z veliko začetnico; kadar je osebno ali zemljepisno lastno ime prekrivno z občnim samostalnikom, je lema pri lastnih imenih torej pisana z veliko začetnico, pri občnih pa z malo.
- gornje pravilo velja tudi za tuje samostalnice: če tujo besedo identificiramo za samostalnik (če se sklanja po samostalniških vzorcih) in je iz sobesedila jasno, da gre za osebno ali zemljepisno ime, označimo kot lastno ime, ne pa kot Nj. Če sobesedilo ali splošna izobrazba ne pravita drugače, tem primerom pripišemo moški spol.

trenutno najstarejši kaveljc je 80-letni Boris Mlinar, ki je tudi častni član društva → lastno ime, lema: Mlinar
pogovor znanega francoskega sociologa Pierra Bourdieuja z nemškim pisateljem → lastno ime, lemi: Pierre, Bourdieu
Francoski igralec Gerard Depardieu je v Cannes prišel zaradi filma Asteriks in Obeliks proti Cezarju → lastna imena, vse leme z veliko

- če iz sobesedila ni razvidno, katerega spola je lastno ime, mu pripišemo moški spol

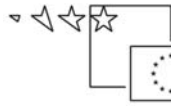
Saša [Slmei] ne reagira.

- priimkom pripišemo spol, kot ga ime, ki se pojavlja ob njem
- avtomatsko lematizirane (napačne) leme lastnih imen popravljamo

utiša medije in ustoliči generala Liamina [Liamin] [Liamine] Zerouala za predsednika

Lastna imena – stvarna

- kadar je v besedilu samostalnik, ki ga prepoznamo (iz oblike, konteksta, splošnega znanja) za stvarno ime, ga označimo kot lastno (lema z veliko začetnico) samo, če



nima prekrivne leme z občnim samostalnikom. Kadar je stvarno ime prekrivno z občnim samostalnikom, ga označimo za občni samostalnik, lema v tem primeru z malo

*Nastale so klasike Rozmarijin otrok, Izganjalec duhov in slasher **Noč [Sozei]** čarovnic.*

- gornje pravilo velja tudi za tuje samostalnike: če tujo besedo identificiramo za samostalnik (če se sklanja po samostalniških vzorcih) in je iz sobesedila jasno, da gre za stvarno ime, označimo kot lastno ime, ne pa kot Nj. Če sobesedilo ali splošna izobrazba ne pravita drugače, tem primerom pripišemo moški spol.

*Od temačnih, samotnih graščin na vrhu hriba, stolpov starih dvorcev, zaprašenih grobnic smo se postopoma preselili v manjše, navidez prijazne hiše, zgrajene na starih in davno pozabljenih pokopališčih (**Poltergeist [Slmei]**, Mačje pokopališče, **House [Slmei]**), nato še v obrobne, visokogorske in odtujene kraje (**Shining [Slmei]**, **Misery [Slmei]**, **Psiho [Somei]**).*

Problematični primeri

- kot lastna imena označujemo kratična imena:
 - .1. geografskih oz. političnih enot: *EU, RS, ZR* itd.
 - .2. političnih strank: *CDU, SD, SDS* itd.
 - .3. podjetij in organizacij (tudi njihovih delov): *BBC, BTC, DURS, DZ, RTV, SFOR, OŠ* itd.
 - .4. kratice zakonov: *ZDR, ZJU, ZSPJS* itd.
- kot lastnih imen ne označujemo delov kratičnih poimenovanj blagovnih znamk ali imen programskih orodij: *M30X, pocketPC, palmOS, LX, DNKA, (F 650) CS*
- lastnim imenom z ženskim in srednjim spolom dodajamo ustrezno lemo:

*Razstava bo letos gostovala še v Mehiki, ZDA, **Japonski**, Finski in Južni Afriki → ženski spol, lema: **Japonska***

*Po avanturi na **Japonskem** ste ponovno nosili majico Maribora. → srednji spol, lema: **Japonsko***

3.3. Označevanje glagolov

Vrsta

- pomožni glagol je samo glagol biti, bodisi kot glagol obstajanja ali v funkciji tvorjenja časov; ostale glagole označujemo kot glavne.



Najpogosteje je **[Gp-ste-n]** vir okužbe jagnjeta njegova mati,...

Predlog, ki bi na referendumu zmagal, potem ne **bi [Gp-g] bil [Gp-d-em]** nič več odvisen od dobre volje poslancev.

Glagolska oblika

- kot deležnike označujemo samo opisne deležnike na *-l* v funkciji tvorjenja časov. Trpne deležnike na *-n/-t* označujemo kot deležniške pridevnike.

Pakistanski predsednik je **dejal [Ggvd-em]**, da bi lahko njegova vlada **razmislila [Ggdd-ez]** o odnosih z Izraelom glede na zadnje premike v bližnjevzhodnem mirovnem procesu.

Akumulator je zelo velik, saj je **oblikovan [Pdnmein]** tako, da lahko namesto njega vstavimo štiri običajne baterije.

- pogojniki označujemo samo pri obliki *bi*.

Podobno **bi [Gp-g]** lahko v splošnem označili tudi izvedbo v drugem delu koncerta, ko je bila na sporedu Mahlerjeva Simfonija št. 1 v D-duru.

- deležja označujemo v okviru besedne vrste prislov.

Ko ga mati objame in mu zapoje pesmico ali pove kratko zgodbo, **nevede [Rnd]** otroku privzgaja občutek, da je to nekaj lepega.

Ali se je, ne **meneč [Rnd]** se za Vrano in Daniela, osredotočila na otroka pred seboj.

..., **začenši [Rnd]** z nenadnim spoznanjem, ki se je gotovo registriralo v Cauliffovih možganih, vedenje, da zares umira.

Nikalnost

- lastnost nikalnosti imajo samo trije glagoli: *biti*, *imeti*, *hoteti*
- nikalnost označujemo samo pri glagolskih oblikah sedanjika in prihodnjika;
- ne označujemo je pri neosebni glagolskih oblikah (deležniku, namenilniku in nedoločniku), pogojniki *bi* ter pri velelniku.

Sicer pa **nisem [Gp-spe-d]** edini Ljubljančan ali Olimpijin mož v moštvu.

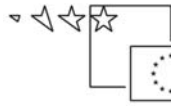
Nad natančnostjo izdelave **nimamo [Ggnspm-d]** pripomb.

Mizice se **nočejo [Ggnstm-d]** pogrnuti.

Vid

- Glagolu *biti* ne pripisujemo vida.

Problematični primeri



- Primere, pri katerih so glagoli na prehodu v druge besedne vrste, npr. *denimo*, *recimo*, *prosim* ipd., označujemo kot glagole oz. se posvetujemo o njihovi uvrstitvi.

3.4. Označevanje pridevnikov

- Pridevniki v besedilih nastopajo v prilastkovi ali v povedkovi rabi:

<i>Sledil je prelomni korak ...</i>	<i>pridevnik kot levi prilastek</i>
<i>Moj mož namreč uradno še ni rehabilitiran.</i>	<i>pridevnik kot povedkovo določilo</i>
<i>Navodila so videti zapletena.</i>	<i>pridevnik kot povedkov prilastek</i>
Tudi v primeru elipse: Pomembno : preden dvignete roke ... (= pomembno je, da preden dvignete roke ...)	

3.4.1 Lematizacija

- Vsi pridevniki imajo pridevniške leme. Če naletimo na deležnik z glagolsko lemo, to popravimo v pridevnik:

<i>hiše, zgrajene na</i>	<i>lema: zgrajen (ne zgraditi)</i>
<i>še ni rehabilitiran</i>	<i>lema: rehabilitiran (ne rehabilitirati)</i>

- Pridevnike, ki so del lastnih imen in se zato pišejo z veliko začetnico, lematiziramo z veliko začetnico:

<i>Tabithine oči</i>	<i>lema: Tabithin</i>
<i>Learbrookovega naročila</i>	<i>lema: Learbrookov</i>
<i>na Grablovičevi ulici</i>	<i>lema: Grablovičev</i>

3.4.2 MSD-ji

- Pridevnike opredeljujemo glede na naslednje kategorije: vrsta, stopnja, spol, število, sklon in določnost.

- **VRSTA: splošni (p), deležniški (d), svojilni (s)**

Splošni:



z **večjim** nožem, **zaporna** kazen, **kazenski** zakonik, **svetovna** javnost, nad **tragičnimi** prizori, kot **notranjo** zadevo, **vojaško** operacijo, **šiški** jug, za **naftne** pipe
speč klošar v travi, našobite se in si nadenite **trpeč** izraz, ima **izstopajočo** brčico, razširjen je v jezerih in toplejših odsekih **tekočih** vod, živel v temnem mraku **grozeče** nevihte

Deležniški:

predvidena kazen, tako **imenovano**, **novozvoljenega** predsednika, z **akcijami**, **uperjenimi** proti, med **ubitimi**, **prepovedanih** strank, strokovno **podprte** nasvete, o **uspelem** prebegu, **pridigana** moralna načela

Svojljni:

Rozmarijin otrok, **Tabithine** oči, **Learbrookovega** naročila, z **očetovim** dovoljenjem, brez **učiteljičine** pomoči

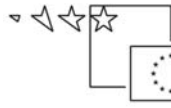
- Pridevnike, ki označujejo **splošno svojino**, npr. državni denar (= denar države), cerkvena posest (= posest cerkve), označimo kot splošne (p). Pridevnike na -ov/-ev načeloma označujemo za svojljne, tudi če pomensko izražajo splošno svojino

*ubrana in intonančno točna glasba zares prihaja iz korenin in **smrekovih** vej → **svojljni**
Začelo se je tavanje, napovedano v **torkovem** komentarju → **svojljni***

- **Deležniški** so tisti, ki imajo bližnjo povezavo s svojim glagolom, kar se tiče vezljivosti, kolokabilnosti in podobnih odnosov. Primer odločanja (prvih 10 kolokatorjev eno mesto desno po korpusu FidaPLUS):

	OPRIJETI brez OPRIJET	OPRIJET brez OPRIJETI
1	vzdevek	hlače
2	ime	majica
3	sloves	kavbojke
4	se	oblačilo
5	naziv	majčka
6	telo	obleka
7	rešilen	boksarica
8	oznaka	kroj
9	vsak	suknjič
10	rob	pulover

- Oprijet torej ni deležniški, ker ni odnosa -> glagolski predmet (oprijeti hlače, oprijeti se hlač?) in pridevniški levi prilastek oz. povedkovo določilo (oprijete hlače // hlače so oprijete). Hitri test za določanje deležniškosti pridevnika bi bil, če zlahka najdete tak predmet - npr. kršiti zakonodajo -> kršena zakonodaja -> zakonodaja je kršena. Primeri:



Deležniški: *poznan, napravljen, določen, omejen, ugotovljen, kršen, dodan, seznanjen, znižan, opustošen, pridobljen, urejen*

Splošni: *oprijet, preostal*

- Vse deležniške pridevnike, tvorjene iz zanikanih glagolov ali glagolskih zvez, označujemo kot splošne: *nedokončan, nedovoljen, neizgovorjen, neizplačan, nemoten, nenadzorovan, nenašminkan, neodgovorjen, neopažen, neorganiziran, neoviran, nepodeljen, nepopit, nepostavljen, neprekinjen, nerazčiščen, nerazodet, nespremenjen, neuskklajen, nevzdrževan, nezavarovan, nežgan, novoustanovljen* itd.
- **STOPNJA: nedoločena (n), primernik (p), presežnik (s)**
- Nedoločena stopnja označuje pridevnike, ki niso stopnjevani, torej »osnovnike«.
- Opisno stopnjevanje: ker nas na ravni morfosintaktičnega označevanja korpusa zanima vsaka beseda posebej, sta pri opisnem stopnjevanju prislova bolj in najbolj (tudi manj, najmanj) obravnavana posebej, pridevnikom pa pripišemo oznako nedoločeno (n):

za bolj prijazen [**Ppnm**etn] jutri – pridevnik je (samo) prijazen

- Kategorija stopnja je pri svojilnih pridevnikih vselej nedoločeno (n), saj teh pridevnikov ne stopnjujemo.

... brez **predsednikovega** [**Psn**mer] privoljenja ...

- Kategorije elativ ni več, zato se besede, prej označene kot elativ, sedaj označujemo kot nedoločene

spoznanje o pridnosti gospodinji, ki skrbijo za **prelep** [**Ppnm**etn] okras domala vsake hiše – nedoločena stopnja, lema: prelep.

- **SPOL: moški (m), ženski (ž), srednji (s)**
- Primeri problemov pri določanju spola

vprašanju vlaganj na **tujem** → **srednji spol**



- **ŠTEVILO:** ednina (e), množina (m), dvojina (d)
- **SKLON:** imenovalnik (i), rodilnik (r), dajalnik (d), tožilnik (t), mestnik (m), orodnik (o)
- Primeri problemov pri določanju sklona

Po hrbtu je rjavkasto **olivne** barve. → **rodilnik**
Hrbtna plavut je razmeroma močno **zarezana** → **imenovalnik**
Po hrbtu je modrikaste barve, spodnji del bokov in trebuh sta **srebrna** → **imenovalnik**
Sama vrednost pa že spominja na **ameriških 4,3 odstotka** → **rodilnik**
Kaj bistveno **novega** ne more povedati → **rodilnik**

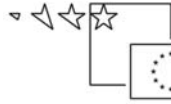
- **DOLOČNOST:** določno (d), nedoločno (n)
- Določnost označujemo pri vseh treh vrstah pridevnikov.
- Določnost označujemo samo pri pridevnikih moškega spola ednine v imenovalniku in tožilniku.
- tabela s tremi vrstami pridevnikov v pomoč:

OBLIKA	LEMA	MSD	Sklon + živost sam. jedra
policistov	policistov	Psnmein	im.
policistov	policistov	Psnmetn	tož. + neživo jedro
policistovega	policistov	Psnmet	tož. + živo jedro
omenjen	omenjen	Pdnmein	im., nedoloč. oblika
omenjeni	omenjen	Pdnmeid	im., določ. oblika
omenjen	omenjen	Pdnmetn	tož., nedoloč. oblika + neživo jedro
omenjeni	omenjen	Pdnmetd	tož., določ. oblika + neživo jedro
omenjenega	omenjen	Pdnmet	tož. + živo jedro
gospodarski	gospodarski	Ppnmeid	im.
gospodarski	gospodarski	Ppnmetd	tož. + neživo jedro
gospodarskega	gospodarski	Ppnmet	tož. + živo jedro

- primeri:

sledil je **prelomni [Ppnmeid]** korak
Prve priprave na **državni [Ppnmetd]** udar

3.4.3 Zahtevnejši primeri



- **Posamostaljene pridevnike** označujemo kot pridevnike (glede označevanja spola glej 3.1):

... ki resnično krepi občutek **neznane** [Ppnser] in **skrivnostne** [Ppnser]
... Tako smo v veliki meri sami krivi za **dobro** [Ppnset] ali **slabo** [Ppnset]
... Ker je bilo o njej **napisane** [Pdnser] že dovolj ...
... seznanjanje z **drugачnim** [Ppnseo] pa utrdi strpnost ...

- **Pridevniške zveze** z vezajem:

ob **iraško** [Ppnsei]-turški [Ppnzem] meji, vojaško [Ppnsei]-državnemu [Ppnmed] zboru

- interpretiramo tako, da prvi del z veznim samoglasnikom -o- obravnavamo kot pridevnik v imenovalniku ednine srednjega spola.
- **Ženski priimki** se označujejo kot pridevniki z lemo v moškem spolu

Strahotno žaluje za očetom, ji je rekla **Evansova** . – im. ed. ž. sp., lema Evansov
Olga **Andrejeva** in Elena **Pankova** – im. ed. ž. sp., lema Andrejev, Pankov

- nesklonljivi prilastki tipa **solo, neto, bruto** ipd. so samostalniki, kadar so sklonljivi (npr. solo), sicer so pridevniki
- **Sam (-a, -o)** označujemo kot pridevnik in ne zaimek

naj kar **sami** poiščejo vlagatelja
Namesto kreme lahko **sami** pripravite mešanico.
Čisto mogoče je tudi, da je Kajn v odgovor na božjo zavrnitev njegovega daru zavrnil in ubil **samega** (lema=sam; msd=!Ppnmetn) boga in da si je bog nemara to celo **sam** (lema=sam; msd=!Ppnmein) želel

3.5. Označevanje prislovov

- vprašalnice so prislovi – tudi kadar se pojavljajo na mestu veznika

Ker ne vemo, **kako** bi drugače dosegli... → **prislov**

- prislov v povedni (brezosebni) rabi je prislov – po SP je povedkovnik, vendar po sistemu **povedkovnika ne označujemo**

Gledano skozi zgodovino so se prizorišča... → **prislov**

Posebni primeri



- stopnjevalni prislovi **bolj/najbolj** ter **manj/najmanj** imajo leme *bolj/najbolj*, *manj/najmanj* ter so označeni kot Rnn. Podoben primer tudi prislova **največkrat** in **najprej**.
- **rad (rada, rado ...)** označimo kot Rnn z lemo *rad*
- po posvetu so bili za prislove označeni tudi naslednji problematični primeri:

*boste vsako jutro potrebovali dobre **pol** ure , da jih boste oblikovali v popolno ravno pričesko. Vetra ni dobesedno **nič**.
Devetletko naj bi **jeseni** poskusno uvedli na 46 šolah.
Tako ali **tako** se moram vrniti.
Zmerno pitje alkohola zmanjšuje tveganje za infarkt **približno** za tretjino.
Komaj so se posamezni igralci vživeli v igro pa so se **spet** razšli.
Imeti **prav**.
Na **splošno**.
Invencija je v ZDA **naravnost** kulturno razvita.*

Deležja

- kategorijo deležje načeloma dobijo vsa slovenska deležja. Vprašalnica za določanje deležij je **kako (se nekaj dogaja) + deležja** imajo predvidljivo **obliko**; deležje izraža hkratnost dogajanja – *shiramo, medtem ko čakamo; opravljam delo, medtem ko stojim*):

*- **na -č**: ...film po stripu, ki se zanaša na Hallejino miganje z boki ,medtem ko se **tepežkajoč** in **predeč** maščuje svojim morilcem.
- **na -aje**: ...Če neka firma propade, gredo delavci na borzo dela, kmetje pa lahko, **čakaje** na miloščino, le *shiramo*.
- **na -e**: ...Ob odprti strehi lahko krmar svoje delo opravlja **stoje**, v salonu pa je več svetlobe.
- **na -ši**: ...To **rekši** je usekal predikanta s pestjo po obrazu, da je bil takoj ves krvav.*

- previdnost je potrebna pri tistih besedah, ki so se bolj ali manj poslovile od svoje glagolskosti (= ne izražajo hkratnosti dogajanja, imajo pa še vedno deležijsko obliko): to **niso deležja**.

*tolarjev lahko celo slikajo s petmetrskim pitonom, ki je **baje** popolnoma nenevaren, razen če ste rejena podgana. → **členek**
Kot podružnica globalne korporacije je Unilever Slovenija tudi **glede** izobraževanja vezan na centralo. → **predlog***

Pridevnik vs. prislov

- če beseda v pridevniškem nizu določa samostalnik, je običajno pridevnik, če določa pridevnik, je prislov



Zakon o industrijski lastnini, **uradno** prečiščeno besedilo → kako prečiščeno? – uradno, z uradne strani → **prislov**; če bi šlo za pridevnik, ki tvori s samostalnikom zvezo uradno besedilo, bi se zveza glasila prečiščeno uradno besedilo, ker vrstni pridevniki v nizu nastopajo bliže samostalnika kot kakovostni.

Večji osebki imajo **značilno** podaljšano prvo plavutnico → kako podaljšano? – značilno (= tipično) → **prislov**; gre za zvezo, ki se v biologiji pogosto pojavlja, pomeni, da je plavut tipično podaljšana, ne pa, da je plavut sama po sebi tipična oz. značilna.

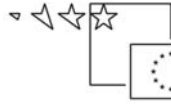
- ločevanje med posamostaljenimi pridevniki ter prislovi lahko olajšajo vprašalnice – če se v teh primerih lahko vprašamo s *kakšen?* gre najbrž za pridevnik

To je **neetično** ter **kaznivo**. → kakšno je? – **pridevnik**
V njem je jasno **zapisano**, da obstaja utemeljen sum → kakšno je? – **pridevnik**
Če bo prišel pravi scenarij, to ni **izključeno** → kakšno ni? - **pridevnik**

3.6. Označevanje zaimkov

Leksikon (leme):

CELOSTNI	KAZALNI	NIKALNI	OSEBNI	OZIRALNI	POVRATNI
oba	oni	<u>nič</u>	čezme	čigar	čezse
obadva	onile	nihče	jaz	čigarkoli	medse
oboj	ta	nikak	jest	čigarsižebodi	nadse
oboje	tadva	nikakršen	medme	čigaver	nase
obojen	taisti	nikakšen	nadme	čigaverkoli	obse
<u>ves</u>	tak	nikdo	name	čigavršen	podse
vsak	takisti	nikogaršen	obme	čigavršnji	predse
vsakateri	takle	nikogaršnji	on	kakršen	se
vsakdo	takošen	nobeden	podme	kakršenkoli	skozse
vsakogaršen	takov	noben	pome	kakršensižebodi	svoj
vsakogaršnji	takšen		predme	kar	vase
vsakršen	takšenle		skozme	karkoli	zase
vsakteri	tale		ti	karsibodi	
	talele		vame	karsižebodi	
	tisti		zame	katerikoli	
	tistile			katerisibodi	
	tolik			kdor	
	toliker			kdorkoli	
	tolikšen			kdorsibodi	
	tolikšnji			kdorsižebodi	
	toti			kolikršen	
SVOJILNI	VPRAŠALNI	NEDOLOČNI			
moj	čigav	bogsigavedikakšen	kdovekakšen	marsikdo	nekolik
najin	<u>kaj</u>	bogvedikaj	kdovekateri	mnog	nekolikšen
naš	kak	bogvedikateri	kdovekdo	mnogokaj	nekolikšnji
njegov	kakošen	bogvekaj	kdovekolik	mnogokateri	nekov



njen	kakov	bogvekakšen	malokaj	mnogokdo	nekšen
njihen	kakšen	bogvekateri	malokak	negdo	premarsikateri
njihnji	kateri	bogvekolik	malokakšen	nek	premnog
njihov	kdo	bogvekolikšen	malokateri	nekaj	prenekaj
njun	koji	enak	malokdo	nekak	prenekateri
tvoj	kolik	enaki	marsičigav	nekakov	prenekdo
vajin	koliker	enakšen	marsikaj	nekakšen	redkokateri
vaš	kolikšen	isti	marsikak	nekateri	redkokdo
		kdovekaj	marsikakšen	nekdo	
		kdovekak	marsikateri	neki	

- **Kot zaimke označujemo le pojavnice v okviru osnovnih oblik, ki so navedene v leksikonu.**

- prislovne vprašalnice oz. vprašalni prislovni zaimki so vedno označeni kot prislovi (*čemu, kako, kam, kdaj, kje, kod, kolikič, koliko, kolikokrat, počemu, zakaj*)
- kot prislovi so označeni tudi drugi prislovni zaimki (*toliko, kdovekoliko, količkaj, kolikanj, itd.*)
- pri zaimkih z določno-nedoločnim parom v imenovalniku ednine je lema enaka imenovalniški obliki le pri določni obliki (oblika: *nek* -> lema: *nek*; oblika: *neki* -> lema: *neki*), pri ostalih oblikah je nedoločna (*nek, enak*)
- **vsi vprašalni zaimki, ki so po obliki enaki bodisi nedoločnim ali oziralnim, so označeni kot vprašalni, ne glede na njihov pomen ali vlogo**

.1. **nedoločni:** *čigav, kaj, kak, kakošen, kakov, kakšen, kdo, kolik, koliker, kolikšen*

.2. **oziralni:** *kateri, koji*

Ločevanje zaimkov od prislovov:

1. *kaj*: vse rabe razen prislovnih

... neumno plačevati za **kaj** takšnega. → **zaimek**
... vsaj imele **kaj** od nje. → **zaimek**
..., s katerim se **kaj** spravi v premikanje. → **zaimek**
Kaj takega pa še ne! → **zaimek**
Če bo **kaj** narobe, pokličita policijo. → **zaimek**
Zato je vprašal, **kaj** mu ima ponuditi naš prijatelj. → **zaimek**
V tistem času se mi ni zdelo, da bi bilo v njem **kaj** tudi zame. → **zaimek**

b. prislov

- sovsebni nikalni členek po SP (s "šele")



....pogledov na dogodke niso mogli vsiliti niti, **kaj šele** socialdemokratski Gerhard. → **prislov**

- v vprašalnih stavkih, posplošuje vprašanje

Kako je **kaj** tam spodaj? → **prislov**

- prislovni zaimek po SP

ne zna pošiljati faksa, stvar je velika in ne gre v žep, kot telefon pa ni nič **kaj** prida → **prislov**
Pred tekmo smo upali, da bomo dosegli **kaj** več. → **prislov**

1. nič: vse rabe razen prislovnih

...**nič** slabega ni, če se znotraj vlade pojavljajo.... → **zaimek**
... delala, kot da se **nič** ne vidi. → **zaimek**
..., moja skupina nima **nič** z mamili. → **zaimek**
O tem problemu še nisem **nič** slišal → **zaimek**
Vozil je na vse ali **nič**. → **zaimek**

b. prislov

- s prislovi, "bolj/manj" (merni prislov)

...razprave v parlamentu ne bodo **nič** manj burne. → **prislov**
...kot telefon ni **nič** kaj prida → **prislov**
...porabništva ni **nič** težje pojasniti kot... → **prislov**

- s "prav"

Vas ni prav **nič** strah? → **prislov**

2. ves (zaimek) proti vse (prislov)

- V oblikah zaimka "ves" se meša vrsta kategorij. Ker od avtomatskega označevanja ne moremo pričakovati zelo subtilnih zmožnosti razlikovanja, se osredotočamo na večjo razliko med prislovom (Rnn), ki ima lemo "vse" in celostnim zaimkom (Zc...), ki ima lemo "ves". Znotraj zaimka se naslanjamo predvsem na oblike. Med celostnim količinskim zaimkom (SP), ki je večinoma levi prilastek (vsi umazani avtomobili) in celostnim mernim (SP), ki je desni (avtomobili, vsi umazani), ne razlikujemo. Količinski zaimek ima samo množino, merni celo paradigmo. Temu je pridružen še celostni zaimek nečloveškosti (SP) - npr. "vse je bilo v redu", ki ima oblike srednjega spola in samo ednino. Če bi želeli razlikovati med vsemi tremi podkategorijami, bi moral imeti merni lemo "ves", količinski lemo "vsi" in "nečloveški" lemo "vse". Po sistemu vse skupaj pa združujemo pod lemo "ves", pri čemer pa nastane dilema, ali



"nečloveškega" interpretiramo kot elipso "količinskega" = ...da pove vse // da pove vse (stvari, glasove, besede...). V prvem primeru bi bil msd Zc-set, v drugem bi imeli Zc-mmt/Zc-zmt. Skušamo ostati čim bolj kosistentni, zato ga opredeljujemo s prvo možnostjo: Zc-set.

V neki reklami sem prebrala nekaj, kar mislim, da pove vse → **zaimek**, lema: **ves**

b. prislov

Za rakom na materničnem vratu zboleva vse več žensk → **prislov**, lema: **vse**
... je bil vse počasnejši. → **prislov**, lema: **vse**
vse do konca 90. let je živela ideja → **prislov**, lema: **vse**

3.7. Označevanje števnikov

- kot arabske številke označimo cifre, pa tudi vsako besedo, ki vsebuje cifre: oznake za datum, še rezultate, številke publikacij, številke za ure z dvopičjem/piko vmes in vsa druga arabska števila, ki nimajo 'regularno' postavljene vejice in pike

... v 8. mesecu brejosti prodam. Tel. **05/763825 [Kag]**.
... s predpisano količino soli (**1.7[Kag]–2.0 [Kag]** kg na **100 [Kag]** kg zribanega zelja) ...

- kot glavne arabske številke označimo tudi tiste besede, pri katerih je okrajšava pomotoma pisana stično s števkjo (kadar je nedvoumno, da gre za številnik)

... s predpisano količino soli (**1.7[Kag]–2.0 [Kag]** kg na **100kg [Kag]** zribanega zelja) ...

- kot rimske številke označujemo zapise z rimskimi oznakami za števila

Hrbtna plavut je razmeroma močno zarezana in ima **X, I**, 31 plavutnic, repna plavut je vbočena, predrepna plavut pa ima **II**, 12 plavutnic.

Glavni in vrstilni števniki

- tako kot pri vrstilnih števnikih tudi pri glavnih v arabskem/rimskem zapisu NE opredeljujemo spola, števila in sklona – msd Kag.
- pri besednih števnikih nad pet spola ne označujemo (črtica)
- pri vrstilnih števnikih, pisanih z besedo, je lema vrstilni številnik
- lema pri glavnih števnikih je moška oblika števnika, zato

Sejem knjig bosta spremljali **dve** razstavi. → lema: dva
V Zdravilišču Laško prirejajo vsako leto **tri** salonske večere → lema: trije



Zaimkovni števniki

- med zaimkovne števnike spadata samo leme *en* (ter varianta *eden*) ter *drug*, in sicer ne glede na to, ali je njun pomen v stavku zaimkovni ali števniki (en je v tem primeru po pomenu glavni, drug pa vrstilni števnik).

*Lastnik je Morsu izročil obroček s tremi ključi: enega [Kbzmet] za sobo 27, ...
Valter kot drugi [Kbzmeid] sin iz Janezovega že tretjega zakona res ni imel veliko možnosti za uspeh.*

- vse oblike, ki spadajo pod "drug", "drugi", imajo lemo *drug*, oblikoslovne oznake pa spremenimo v zaimkovne števnike.

... obtožnica pa jima med drugim [Kbzmeo] očita še vlom v ilirskobistriško trgovino z orodjem, ...

- oblika "drug" in "drugi" pri moškem spolu ednine imenovalnik in tožilnik dobita lastnost "določnost".

*Pronukleusa se približata drug [Kbzmein] drugemu in se zlijeta.
Drugi [Kbzmeid] je v tesni povezavi s prvim.*

- "poldrug" je pridevnik, lema je "poldrugi" (oblika "poldrug" in "poldrugi" pri moškem spolu ednine imenovalnik in tožilnik imata lastnost "določnost"):

Bil je majhen možak, nič več kot za poldrugi [Ppnmetd] meter ga je bilo,...

Drugi števniki

Predvsem:

- množilni števniki (*enojen, dvojen, trojen, četveren, ...*)
- vrstilni ločilni (*enoj, dvoj, troj, četver, peter, ...*)

Določnost

- opredeljujemo samo pri množilnih števnikih (msd **Kbdmein/Kbdmeid** (*dvojen/dvojni*)) in pri zaimkovnem števniku drugi (samo oblika **drug**, msd **Kbzmein**).

Posebni primeri



- posamostaljeni števniki, ki označujejo večkratnik od osnovnega števila (*tisoči, stotine, desetine*), so samostalniki. V tem primeru so leme množinske, da se jih loči od deležev.

... kjer so marca in aprila **stotisoči** Kurdov skušali pobegniti v Turčijo.
V obdobju, ko je trg zasičen s **tisoči** marketinških sporočil, je boj za pozornost kupcev neizprosna.

- besede, ki označujejo kratnost, spadajo pod **a.** pridevnike (enkrat, dvakrat, trikrat, ...), **b.** prislove (enkrat, dvakrat, trikrat, ...)

pri **enkratnem [Ppnset]** navoju z ročko – lema: enkrat
Poleg tega rodiš **dvakrat**, morda **trikrat [Rnn]** v življanju,

- kadar v kombinaciji števnik z vezajem le-temu sledi sklonska končnica, označimo ta del kot Np, sicer pa kot predstavnika ustrežajoče besedne vrste

deset novink, bo treba sklepe iz Nice prilagoditi **25-članski** skupnosti → pridevnik
Po vsej državi je zagorelo kar **4.382–krat** → prislov

142-im so izrekli denarno kazen → Np

- variante izražanje časa, nekaj primerov

stanovalce pa so reševali na Šišenski cesti ob pol **osmi h[Kbg-mr]** zvečer → lema: osem
Prvo pravilo zanje je: po **dvajseti** uri nič več hrane → vrstilni števniki, mestnik, ženski spol, lema: dvajseti

3.8. Označevanje predlogov

Leksikon:

		k	Dd	nasproti	Dd	poleg	Dr	sprej	Dr
a	Dt	kljub	Dd	nasproti	Dr	povrh	Dr	spricho	Dr
blizu	Dd	koncem	Dr	navkljub	Dd	povrhu	Dr	sred	Dr
blizu	Dr	konec	Dr	navzlic	Dd	preblizu	Dr	sredi	Dr
brez	Dr	kraj	Dr	o	Dm	pred	Do	širo	Dr
čez	Dt	križem	Dr	ob	Dm	pred	Dt	tekom	Dr
do	Dr	krog	Dr	ob	Dt	prek	Dr	tostran	Dr
glede	Dr	med	Do	od	Dr	preko	Dr	v	Dm
iz	Dr	med	Dt	okoli	Dr	pri	Dm	v	Dt
izmed	Dr	mimo	Dr	okrog	Dr	proti	Dd	vdolž	Dr
iznad	Dr	na	Dm	onkraj	Dr	raz	Dr	via	Di
izpod	Dr	na	Dt	onstran	Dr	raz	Dt	vkjub	Dd
izpred	Dr	nad	Do	po	Dm	razen	Dr	vpričo	Dr
izven	Dr	nad	Dt	po	Dt	skoz	Dt	vrh	Dr
izvemši	Dt	namesto	Dr	pod	Do	skozi	Dt	vrhu	Dr
izza	Dr	napram	Dd	pod	Dt	spod	Dr	vsled	Dr



vštric	Dr	za	Do	zastran	Dr	znotraj	Dr	ž	Do
vzdolž	Dr	za	Dr	zavoljo	Dr	zoper	Dt	ž	Dr
z	Do	za	Dt	zmed	Dr	zraven	Dr		
z	Dr	zaradi	Dr	znad	Dr	zunaj	Dr		

- **Kot predloge označujemo le pojavnice, ki so navedene v leksikonu.**

- imenovalnik se uporablja pri predlogu "via"

*TV se mu zdi sicer prihodnost, toda ker ga začne hitro dolgočasiti, pusti, da ga CIA - **via** George Clooney - rekrutira za tajnega hitmana, ja, poklicnega morilca. → imenovalnik*

- v primeru, da je v stavku elipsa, določimo predlogu tak sklon, kot bi ga imel, če samostalnik ne bi bil izpuščen

*Na morju je za meter **do** dva "marete z juga", ki očitno piha nekje drugje → elipsa (meter do dva metra), predlog v rodilniku*

Ločevanje predlogov od veznikov

1. *rozen*: a. predlog pred samostalnikom v rodilniku

*Jugoslovanske oborožene sile se bodo povsem umaknile s Kosova, **rozen** graničarskih enot, ki bodo delovale samo v petkilometrskem pasu od meje. → predlog*

b. veznik v vseh ostalih rabah

*Ljudje s seboj nosimo množico podob sveta, ki ga še nikoli nismo videli, **rozen** morda po televiziji. ...nehamo nalivati, še preden pridejo kvasovke v kozarec (**rozen** pri nekaterih vrstah piva, pri katerih je sediment zaželen... → veznik*

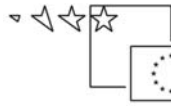
Ločevanje predlogov od prislovov

- za razliko od prislova predlog vedno stoji pred sklonljivo besedo (ali bes. zvezo)

*hotelirji se ravnaajo po pravilu, da je treba **okoli** svojega dvorišča postaviti ograjo → predlog
Nekaj časa so se skrivali **okoli**, nato so odšli → prislov*

- predlogi, ki imajo skupno lemo s prislovi, so: *čez, blizu, glede, križem, mimo, nasproti, okoli, okrog, poleg, povrh, povrhu, proti, skozi, širom, tostran, vštric, za, znotraj, zraven, zunaj*

- *glede* označujemo kot prislov v zvezi *glede na*



3.9. Označevanje veznikov

Leksikon:

		da	Vd	ki	Vd	od	Vd	toda	Vp
a	Vp	dasi	Vd	kjer	Vd	odkar	Vd	torej	Vp
ako	Vd	dasiprav	Vd	ko	Vd	odkoder	Vd	vendar	Vp
akoprav	Vd	dasiravno	Vd	koder	Vd	oziroma	Vp	vendarle	Vp
ali	Vp	dočim	Vd	kolikor	Vd	pa	Vp	zato	Vp
ampak	Vp	dokler	Vd	kot	Vd	preden	Vd	zatorej	Vp
bodisi	Vp	in	Vp	marveč	Vp	razen	Vp		
če	Vd	kadar	Vd	naj	Vd	saj	Vp		
čeprav	Vd	kajti	Vp	najsi	Vd	sicer	Vp		
čeravno	Vd	kakor	Vd	namreč	Vp	tako	Vp		
četudi	Vd	kamor	Vd	ne	Vp	temveč	Vp		
čim	Vd	ker	Vd	niti	Vp	ter	Vp		

- **Kot veznike označujemo le pojavnice, ki so navedene v leksikonu.**
- za ločevanje veznikov od predlogov glej poglavje Označevanje predlogov

Ločevanje veznikov od členkov

1. *da*: a. členek, ki izraža soglašanje

Ste ga poznali? To je ideja, da! → členek

b. podredni veznik

2. *naj*: a. členek (spodbujevalni, zadržkovni, domnevnostni itd.)

Na voljo naj bi imeli blizu milijardo tolarjev. → členek
S prvim januarjem 2003 naj bi sedanjo VVO nasledil /.../ → členek
Naj vam najprej povem, da je vaša revija THE BEST → členek

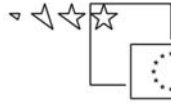
b. podredni veznik (kadar ga lahko zamenjamo z drugim veznikom – za ločevanje glej SSKJ)

Njegovi strokovnjaki so svetovali, naj oba fanta nekaj časa ne prideta v šolo → veznik
Naj gre za umetne ribe, vrteče ali valujoče blestivke, ribič mora po metu čutiti, kako se umetna vaba obnaša → Najsi gre za umetne ribe ... → veznik
dal mi je znamenje, da naj ga počakam → veznik

3. *ne*: a. členek (nikalni členek – vse rabe razen vezniške)

b. priredni veznik (samo kot del dvodelnega veznika ne – ne)

Tudi zato s tovrstnim nasiljem ne ravnatelj ne učitelji nimajo izkušenj → veznik



4. *niti*: a. členek (vse rabe razen vezniške)

b. priredni veznik (samo kot del dvodelnega veznika *niti* – *niti*)

Jadrnica s topovi v kupolah zato ne bi mogla **niti** streljati **niti** jadрати → **veznik**

Ločevanje veznikov od prislovov

1. *ali*: a. priredni veznik

Kako povečati verjetnost za zmago nad enakovrednim **ali** boljšim tekmečem? → **veznik**

b. prislov (vprašalnica – zamenljiva s "če", "mar", tudi v vezniški vlogi)

Ja, ampak **ali** je Jožef sledil tudi zvezdi? → **prislov**

2. *čim*: a. podredni veznik (kadar ga lahko zamenjamo z drugim veznikom: ko, takoj ko)

Čim je naredila taktično napako, je spregledala še taktično možnost nasprotnice in celo izgubila. → **veznik**

b. prislov (v vseh ostalih rabah)

...si najbolj želi poroke z njo, medtem ko si ona želi predvsem **čim** lepšo poroko... → **prislov**
Kapitan Picard bi rad Dato **čim** prej spravil na Enterprise, da bi izvedel o podrobnostih neljubega
→ **prislov**

3. *in*: a. priredni veznik

Mislite, da sta se Martin Krpan **in** Janez Vajkard Valvasor kdaj ustavila v Črnem Kalu? → **veznik**

b. prislov (na vprašanje "kje?")

Črnci so **in**. → **prislov**

4. *sicer*: a. priredni veznik (kadar ga lahko zamenjamo z drugim veznikom: "toda", "a", "če ne" + "in sicer")

Vendar moramo ob uživanju polnovrednih žit **in** tudi stročnic popiti veliko tekočine, **sicer** bo blato še bolj trdo.
V četrtek, **in sicer** uro po polnoči, je nekdo vlomil v trgovino Jeans corner store sport v Tyrševi ulici v Mariboru. → **veznik**

b. prislov (v vseh ostalih rabah)

Takrat jim je delodajalec **sicer** izplačal plače, ne pa tudi regresa



Odgovoril jima je poročevalec zakona, **sicer** predsednik parlamentarnega odbora Srednjeevropske pobude Felice Besostri: → **prislov**

5. **tako**: a. priredni veznik (samo v: "tako da" ter kot dvodelni veznik tako – kot, tako - kakor)

Kremljologija je danes bolj ali manj mrtva znanost, **tako** da so se morali tisti s talentom ozreti drugam.
Dejstvo pa je, da so samomori **tako** po svetu kot tudi pri nas podcenjeni.
→ **veznik**

- b. prislov (v vseh ostalih rabah)

U. J. iz Postojne je osumljen organiziranja mreže prekupčevalcev, ki so, **tako** kot on, prevažali, hranili in ponujali v prodajo mamila → **prislov**

6. **vendar**: a. priredni veznik (kadar ga lahko zamenjamo z drugim veznikom: toda, a)

isti del ne sodi pod njihovo pristojnost, **vendar** bomo radi tudi tam vrgli kakšno lopato soli → **veznik**

- b. prislov (v vseh ostalih rabah: "kljub temu", "nenazadnje")

...sam nisem nič vernika in **vendar** z največjim zadoščenjem uživam v tej razsežnosti Bachove glasbe → **prislov**

7. **zato**: a. priredni veznik

Lepo vreme se bo nadaljevalo še konec tedna, **zato** je prav, da ga preživimo prijetno na soncu. → **veznik**

- b. prislov (samo pred odvisnikom)

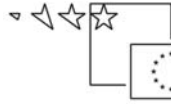
Golansko planoto so pripravljene zapustiti **zato**, ker so močni, ne **zato**, ker bi bili šibki. → **prislov**

8. **zatorej, torej**: a. priredni veznik (kadar jih lahko zamenjamo z drugim veznikom: in, kajti, vendar, saj)

Do sedaj teh razpisov še ni bilo, **zatorej** tudi sofinanciranja ne. → **veznik**
Pri nas ptice najintenzivneje prepevajo spomladi, **torej** v času gnezdenja. → **veznik**

- b. prislov (v vseh ostalih rabah)

Prva poteza je **zatorej**, da zavrnejo oktroirano "javno razpravo z vsiljenimi ideološkimi koordinatami. → **prislov**
Tudi vprašanje politike do drog je **torej** vrednostno vprašanje. → **prislov**



3.10. Označevanje členkov

Leksikon:

a	edino	malodane	no	še
baje	edinole	malone	pač	takorekoč
baš	itak	menda	prav	tudi
bojda	ja	morda	pravzaprav	več
bržčas	kajne	morebiti	predvsem	valjda
bržda	kajneda	naj	prejkone	vsaj
bržkone	kajpada	najbrž	ravno	vsekakor
celo	kajpak	najbrže	samo	zgolj
cirka	kakopak	ne	seveda	zlasti
češ	koli	nemara	skoraj	znabiti
da	kratkomalo	nikar	skorajda	že
domala	kvečjemu	nikarte	sploh	
	le	niti	še	

- Kot členke označujemo le pojavnice, ki so navedene v leksikonu.

Ločevanje členkov od prislovov

1. *prav*: a. členek (izpustljiv iz stavka, ni vprašalnice)

*zdi pa se mi, da tudi danes ni **prav** nič drugače ...
In **prav** to velikokrat pripelje do nepopravljivih napak.
Prav, pomagal ti bom. → **členek***

b. prislov (vprašalnica "kako?")

*Saj sem že ves čas slutil, da sem daljnega leta 1998 ravnal **prav**, ko sem... → **prislov***

2. *ravno*: a. členek (izpustljiv iz stavka, ni vprašalnice)

*Vendar pa se mi zdi verjetnost, da bi te civilizacije za svoj obisk izbrale **ravno** drobno Zemljo, zelo smešna.
→ **členek***

b. prislov (v vseh ostalih rabah)

*Junakinja sodobnega romana je zmerom božansko visoka in se zmerom drži **ravno** kakor sveča, z dvignjeno glavo. → **prislov***

3. *več*: a. členek



Otrok v pokritih peskovnikih slabo vreme ne bo **več** motilo ... → **členek**, lema: več, msd: L.
(izpustljiv iz stavka, ni vprašalnice)

b. prislov (v vseh ostalih rabah)

...komisija se namreč lahko odloči tudi za podelitev **več** enakovrednih nagrad ... → **prislov**,
lema: veliko, (**koliko** enakovrednih nagrad? - več)

c. števnik

Natanko enajst mesecev po hudi poškodbi in **večih** operacijah kolena ... → **števnik**, lema: več,
msd: K.
(kot števnik opredeljujemo SKLONSKE OBLIKE; kadar je oblika imenovalniška, ni števnik.)

O priročnikih:

- SSKJ ne ločuje med prislovi ter členki, včasih govori o členkovni rabi, ampak nedosledno. Je pa dokaj dobra referenca za ločevanje členkov (oz. prislovov) od veznikov oz. predlogov.

3.11. Označevanje medmetov

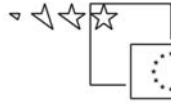
- kot medmet označujemo besede, ki so onomatopoetske v smislu, da posnemajo bodisi človeške neartikulirane glasove (smeh, krike itd.) ali kakšne druge zvoke iz okolja (živalski glasovi, zvoki strojev itd.)

sem in kar naenkrat mi je na klancu spodrsnilo in **aaaaa** sem se zapeljala ter na koncu tudi
sesedla
enega z naslovom Happy Feet, kjer me boste, **haha**, lahko videli preoblečeno v pingvina.
In vendar je Bog zmotljiv. Zakaj? **Heh**, ker je ženska.

- kadar se odločamo, ali nekaj označiti za medmet ali kaj drugega, se odločimo za **drugo** oz. se posvetujemo; namig za ločevanje: medmeti so načeloma nepredvidljivi, zato jih priročniki večinoma ne zajemajo. So tudi izpustljivi iz stavka.

3.12. Označevanje okrajšav

- kot okrajšave označujemo vse okrajšane besede in kratice (npr., itd. št., ing., mag., CO2, Feb, Avs, M2, Pe, EUR, USD) – VENDAR: glede sklanjanjih primerov glej spodaj!



- kot okrajšave označujemo tudi kratice avtorjev člankov ipd. (ar)
- besedo označimo kot okrajšavo tudi, kadar ni zapisana v skladu s pravopisom (na primer, kadar namesto d.o.o. piše d.o.o)
- če je okrajšava sklanjana in v besedilu funkcionira kot samostalnik, jo označimo kot samostalnik (glej označevanje samostalnikov, občna imena)
- kratice (akronime) označujemo kot samostalnike (NUK, NAMA, Mors itd.)
- okrajšave niso nujno samostalniki (na primer, okrajšava je tudi dipl.)
-

3.13. Označevanje neznanih besed

- če za besedo iz sobesedila ni mogoče ugotoviti, kaj označuje, in je zato ne moremo uvrstiti v nobeno od besednih vrst, ji pripišemo oznako N
- oznako N načeloma pripišemo vsem besedam, ki so hibrid črk in števil, razen če so le-te sklanjane

*Sattelite M30X [N] premore zadaj tudi dva...
foto 45b [N]: Z eno roko primemo stopalo (z levo ...*

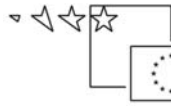
- številke, ure, datumi, negativna št., decimalna št., telefonske številke ipd. označujemo kot števnik (glej označevanje števnikov)

Tel. : 05/76 [Kag] 38 [Kag] 25 [Kag].

- če se v besedilu pojavljajo daljši citati v tujem jeziku, te besede označimo z oznako Nj; to ne velja za posamezne besede tujega izvora, ki funkcionirajo v slovenskem besedilu in imajo slovenska obrazila, te označujemo z oznako za ustrezno besedno vrsto)

*Ves stadion je glasno prepeval We [Nj] will [Nj] we [Nj] will [Nj] rock [Nj] you [Nj].
madam de [Nj] Berny
The Life of Brian«, »Handmade Films —> Brian je lastno ime. Ostalo Nj in leme z malo.
Druga imena: longnose [Nj] dace [Nj] (angl.), Amerikanischer [Nj] Schwarznase [Nj]
sploščen trup, podobno kakor R. rubilio [Nj]*

- če se v besedilu pojavi daljši citat v tujem jeziku, označimo celoten citat z Nj in pripišemo leme, enake besednim oblikam, tudi če navedek zaradi sorodnosti s slovenščino mestoma prinaša besedje, enako slovenskemu (npr. predlogi v hrvaščini).



- Če je beseda **zapisana z nestandardnim zapisom**, jo opredelimo z ustrežno oznako in dodamo lemo v nestandardnem zapisu. Nestandardni zapis prepoznamo po tem, da se v rabi večkrat pojavlja, tipkarske napake so večinoma redkejše

v **Nedeljskemu [Ppnmed]** dnevniku, v **zadnem [Ppnmem]** času (lema: zadni)
dober poljudni članek pa je dosti **teže [Rrn]** napisati kot strokovnega (varianta teže/težje, lema: težko)

- če gre pri besedi za tipkarske napake, jo označimo z oznako Nt

z **asbolutno.[Nt]** večino
Plavuti so razporejene podobno kot pri **atlantskih [Nt]** sledih

- če gre pri besedi za napačno tokenizacijo, jo označimo z oznako Np

Vse potrebne informacije najdete na **www. [Np] siol.net [Np]**. Za nadaljnja vprašanja smo na voljo na naslovu **info@ [Np] siol. net[Np]**.
Navjalna hitrost je med 0,4 m do **o,9 m /.../«** (namesto številke »0« je črka »o«) → napačna tokenizacija
SVIZ-ove [Np] primerjave plač

4. Problematični primeri

4.1. Večbesedne enote

- označujemo vsako besedo ločeno, ne glede na to, ali je del stalne besedne zveze ali ne

Prav zato na ministrstvu pripravljajo /... → **členek + veznik**
Medtem ko so krave na Veliki planini veselo čofotale po mlakah in ... → **prislov + veznik**
Prav te **glede na** potrebe raziskav izberejo strokovnjaki ... → **prislov + predlog**

4.2. Skladenjske napake v stavkih

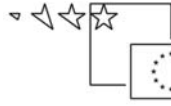
- v primeru skladenjskih napak (napačnih sklonov, napačnega nanašanja ipd.) skušamo problem označevanja rešiti na ravni minimalnega konteksta, tako da izberemo interpretacijo, ki je vsaj deloma ustrezna

*Ib Sejersen je dejal, da slovenska podjetja večinoma že ravnajo v skladu z evropskimi predpisi, čeprav ta za zdaj še niso **zavezujoča**. → predpisi niso zavezujoči, napaka upoštevanje nanašanje na **TA NISO**, torej množina, srednji spol: **Ppnsmi**.*

*Ko kateri od **testnimi ljudmi** zapusti sobo → napačna raba sklona označimo, kot da je izbira sklona ustrezna, m. spol, množina, orodnik: **Ppnmmo**.*

5. Priporočeni referenčni viri

- Slovar slovenskega knjižnega jezika



<http://bos.zrc-sazu.si/sskj.html>

- Slovenski pravopis 2001
<http://www.zrc-sazu.si/pravopis/>
- Amebisov besedni zaklad
<http://www.amebis.si/izdelki/abz/>
- Kaj početi s korpusom FidaPLUS
http://www.fidaplus.net/Files/Kaj%20poceti%20s%20korpusom%20FidaPLUS_navadna%20dvostranska.pdf



SKLADENJSKO OZNAČEVANJE

Cilj tega dela projekta SSJ je ročno označiti 400.000 izbranih besed iz korpusa jos1M. Pri označevanju se uporablja sistem skladijskega označevanja, razvit za slovenščino v okviru projekta JOS. 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 bo uporabljeno orodje za označevanje, ki je bilo razvito v okviru projekta JOS, avtor programa je mag. Janez Brank.

1 Sistem označevanja

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

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

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 skladijske konstrukcije.

Razlaga sistema je seveda podkrepljena s primeri, označenimi v Označevalniku. Na mestu je opozorilo, da se **označenost besednih zvez⁴ 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 kaki globlji označevalni logiki.

⁴Tj. svetlomoder okvirček okrog besedne zveze, ki prinaša tudi kodo zveze (npr. s1c1).

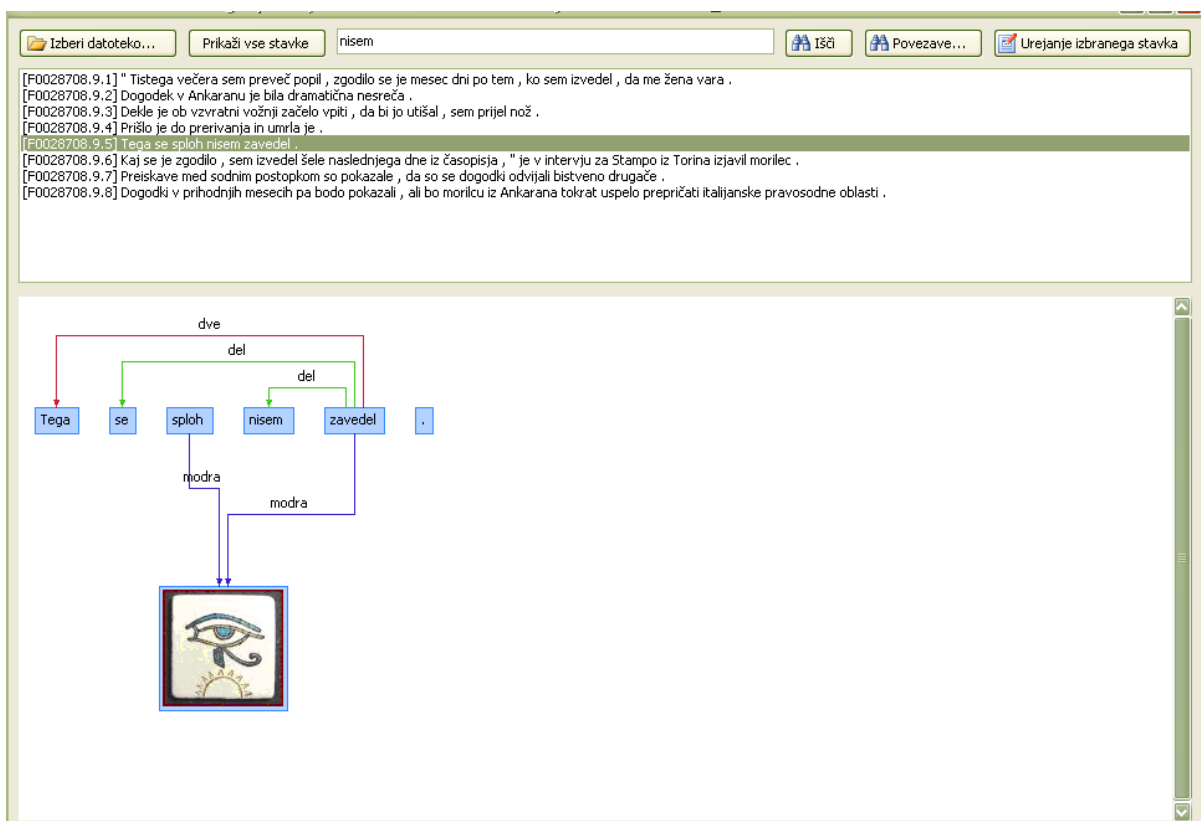


Dodatna literatura za označevalce je Slovenska slovnica (Toporišič 2004, v nadaljevanju SS), ki jo uporabljamo kot osvežitev znanja s področja jezikoslovnih kategorij tradicionalnega tipa. Primeri, ki jih slovnica navaja, so lahko v pomoč pri odločanju, kako obravnavati avtentične primere jezikovne rabe, 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).

3 Program za označevanje

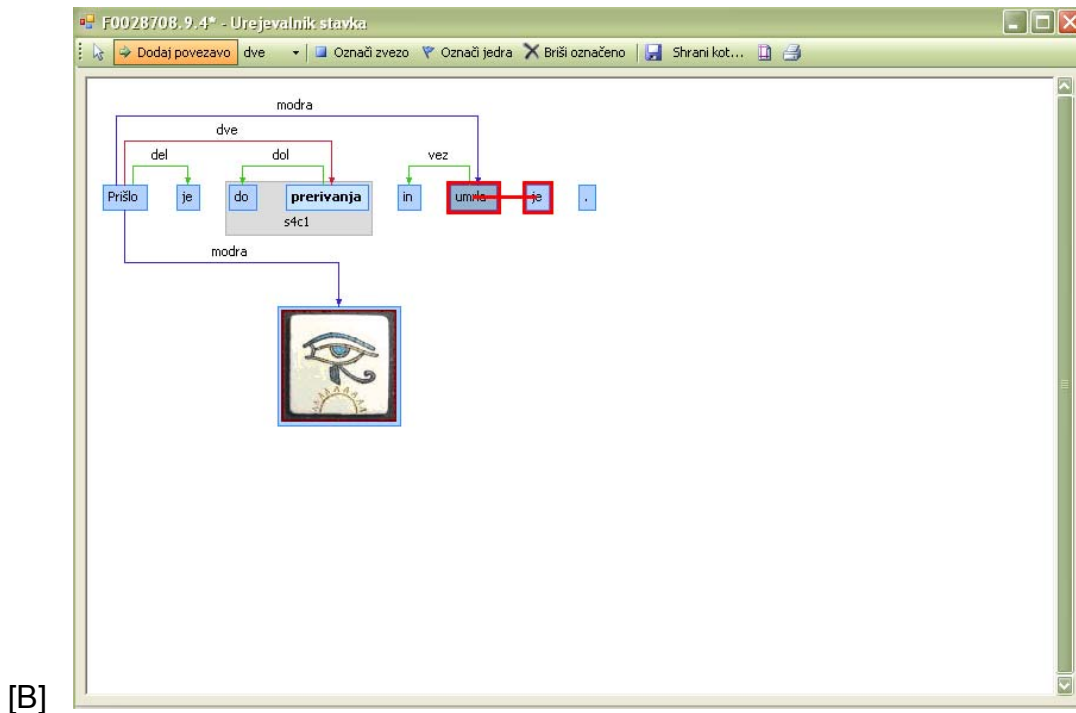
Predvideno je, da se označevalci s programom za označevanje seznanijo v živo, zato na tem mestu samo hiter povzetek funkcij vmesnika.

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].



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 jedrih 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.

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.



Povezave

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

- **zelena** je načeloma namenjena povezavam prvega nivoja, tj. znotraj besednih zvez⁵,
- **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.

4.1 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

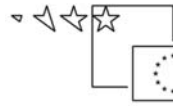
4.1.1 Povezava DOL

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 pregledati.

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

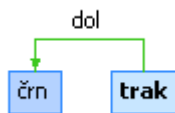
- **samostalniških besednih zvez** (jedro je samostalnik)

⁵Pojme *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.



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

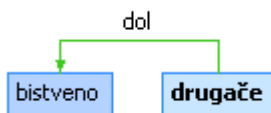
Vprašalnice, s katerimi si pomagamo pri določanju jedra, so torej *kakšen/čigav/ kateri* + jedro.



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?)

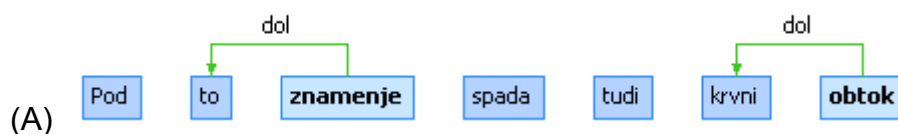


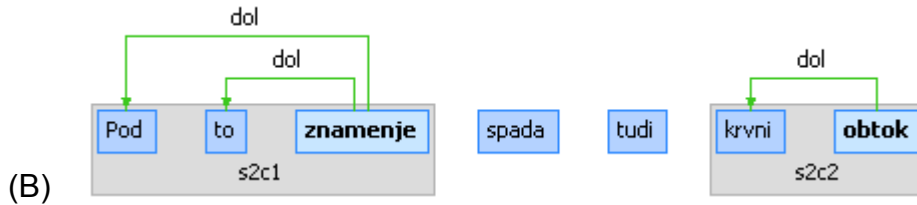
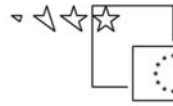
Povezava *dol* se uporablja tudi pri označevanju **povedkovega določila** in povedkovega prilastka (jedro je glagol *biti* oz. polnopomenski glagol), čemur je posvečeno poglavje 4.4.5.

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.





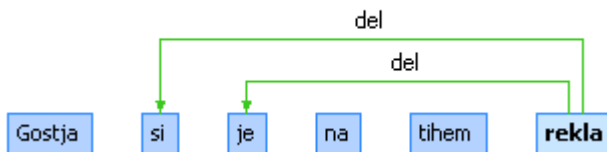
Pogosti v jeziku so primeri, kjer pridevnik ne določa le samostalniškega jedra, ampak **celotno samostalniško besedno zvezo na svoji desni** (tj. tudi pridevnike na svoji desni). V teh primerih postavimo izhodišče puščice v pridevnik, ne samostalnik. Pri odločanju, ali pridevnik določa le samostalnik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



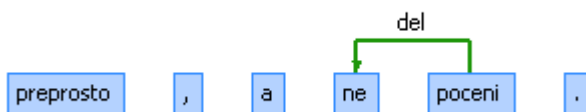
4.1.2 Povezava DEL

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 -l; 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:



4.1.3 Povezavi PRIR in VEZ



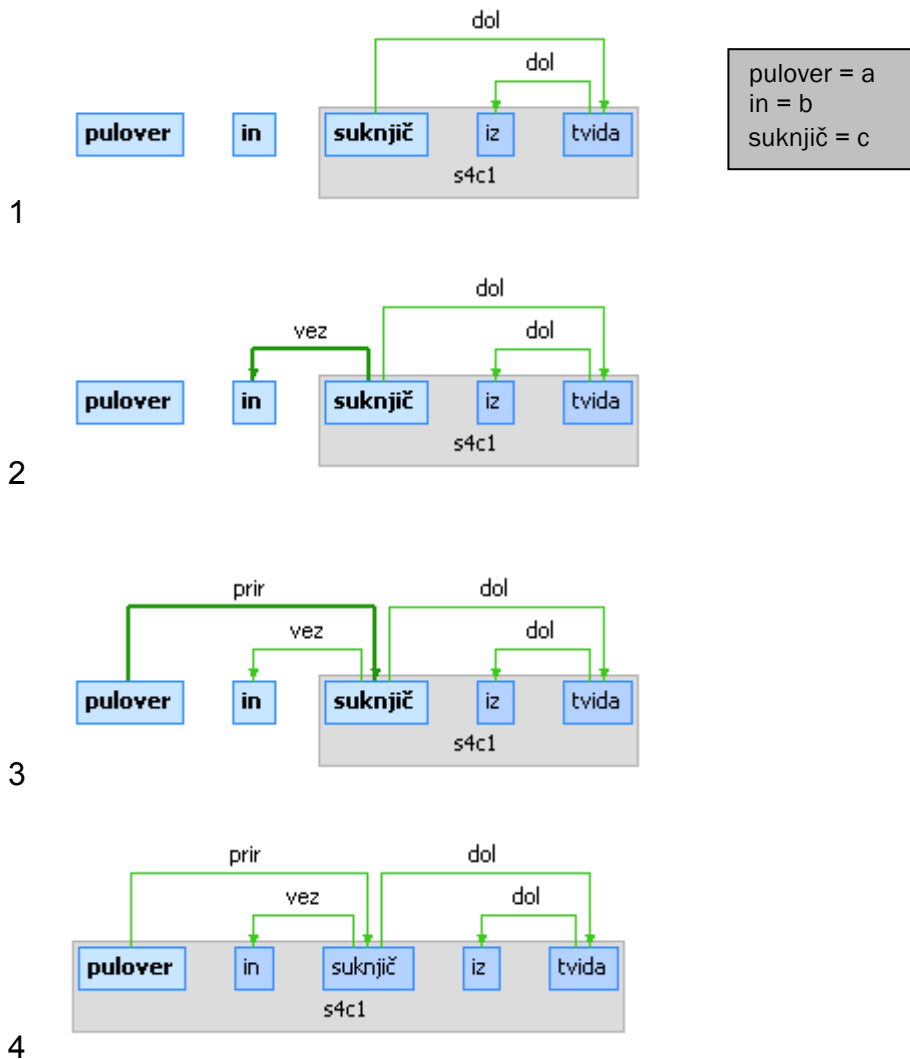
Na prvem nivoju povezav se povezavi *prir* ter *vez* pojavljata v kombinaciji, zato ju obravnavamo na istem mestu. Uporabljamo ju za označevanje prirednih besednih zvez.

Stavčna priredja označujemo drugače, o tem je govora v poglavju 4.3.

Pri označevanju priredij na ravni besednih zvez iščemo tri elemente:

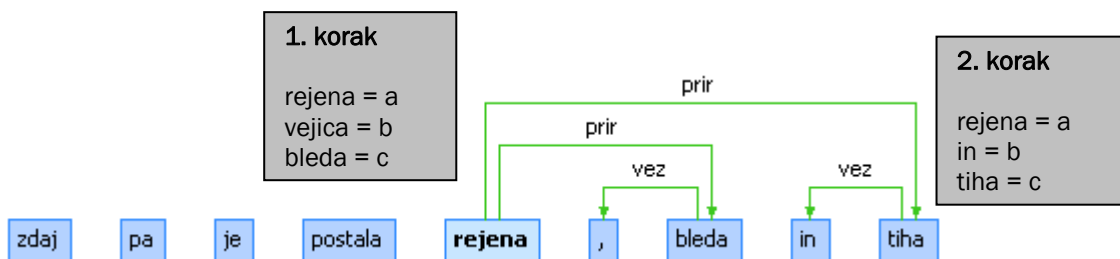
- jedro prvega dela priredja (a)
- priredni veznik ali priredno ločilo, če veznika ni (b)
- jedro drugega dela priredja (c)

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):



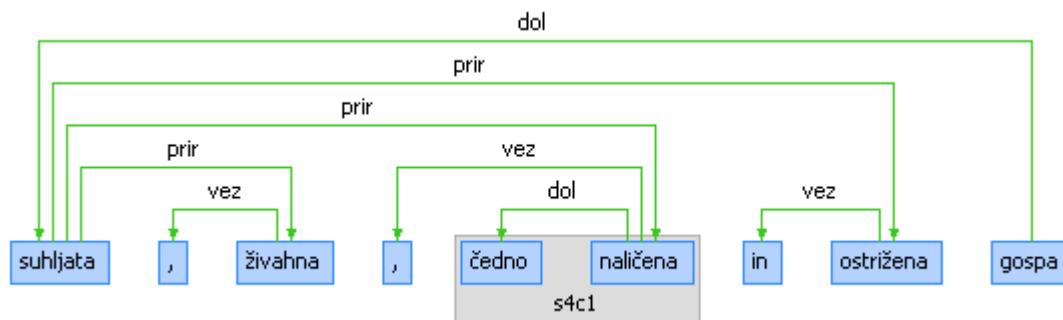
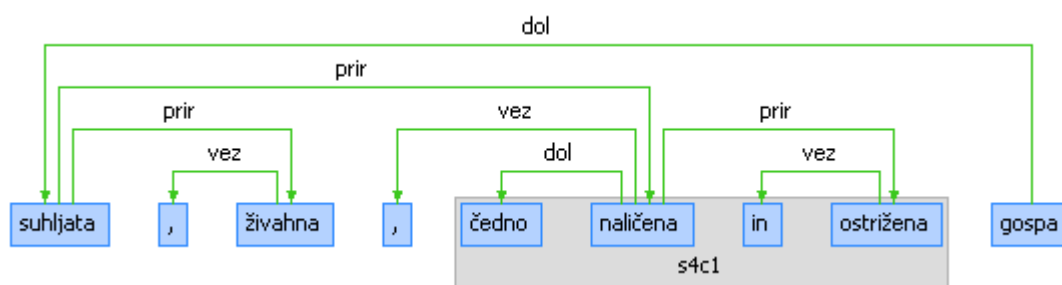


Tudi če je elementov priredja več, je jedro prvega dela priredja vedno tisto, ki bo na koncu povezovanja jedro celotne zveze. Ta del je torej stalnica v označevalnem trikotniku. Spodnji primer prikazuje zvezo, kjer so v prirednem odnosu trije pridevniki. Povežemo jih na način, kot je opisan zgoraj:



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 dvoumnosti smiselno predvideti.





4.1.3.1 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 (4.2), polstavčnih desnih prilastkov (4.4.8), primerjav (4.4.10) 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 metaelement⁶ s povezavo tretjega nivoja (4.3).

4.1.3.2 Povezovanje vejice

Kadar v stavku primerne veznika ni, stavčno razmerje pa nakazuje ločilo (običajno je to vejica), povežemo s povezavo *vez* vejico. V vezalnoprirednih razmerjih vejica npr. nadomešča veznik *in*. Pri nekaterih skladijskih konstrukcijah dosledno povezujemo vejico, npr. pri polstavčnih strukturah (4.4.8).

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* povezujemo besedo *pa* le takrat, ko je popolnoma zamenljiva z veznikom *in*, sicer besedo *pa* večinoma povezujemo s povezavo *modra* barve na metaelement – s povezavo *vez pa* v tem primeru označimo vejico.

4.1.4 Povezava SKUP

S povezavo *skup* povezujemo (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* povezujemo 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

⁶Kot 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.



skup in jo dodati na že pripravljen seznam, se posvetujemo z drugimi označevalci oz. s pripravljenci 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:



4.2 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:

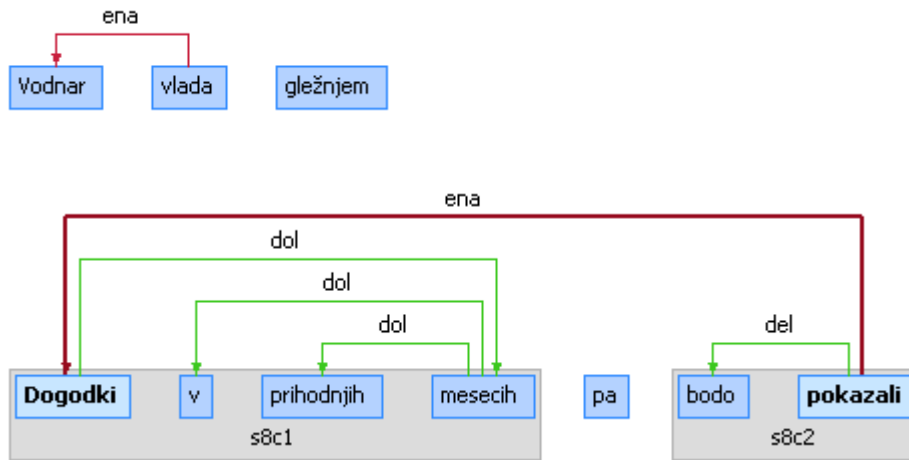
ena	osebek stavka ← povedek stavka
dve	predmet stavka ← povedek stavka
tri	prislovno določilo lastnosti ← povedek stavka
štiri	ostala prislovna določila ← povedek stavka

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

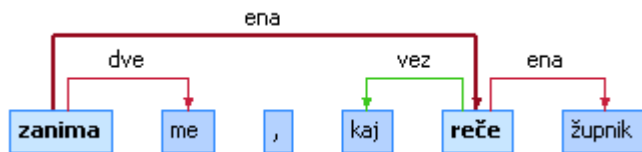
Povezave te kategorije pripisujemo v večini primerov znotraj stavčnih meja, v določenih primerih pa tudi preko le-teh (kadar je v vlogi stavčnega člena celoten stavek, kakor npr. pri osebkovih, predmetnih, prilastkovih odvisnikih – ti primeri so razloženi sproti).

4.2.1 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 osebkov:



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

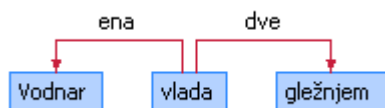


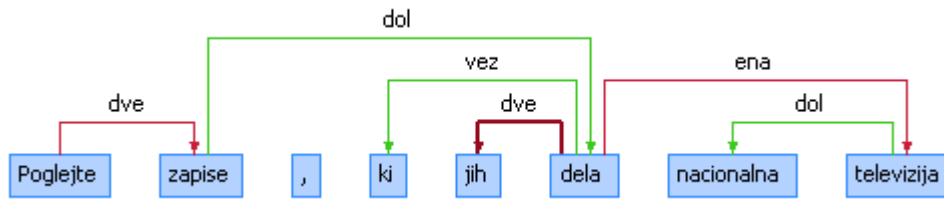
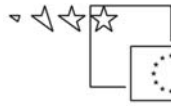
Logičnega osebk ne označujemo.

V primerih tipa *Lužinovi se je zdelo* bi bil samostalnik *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 (sledječ vprašalnici *komu – čemu se je zdelo?*). Kot osebki so lahko označeni samo zanikani osebki v roditelniku in partitivni roditelniki.

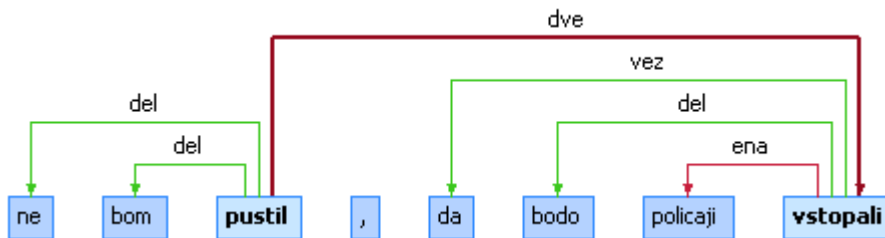
4.2.2 Povezava DVE

Na znotrajstavčni ravni povezava *dve* načeloma ustreza razmerju med povedkom ter nestavčnim predmetom stavka: pri določanju si pomagamo z vprašalnico za neimenovalniške skloni (*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 povezujemo 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):

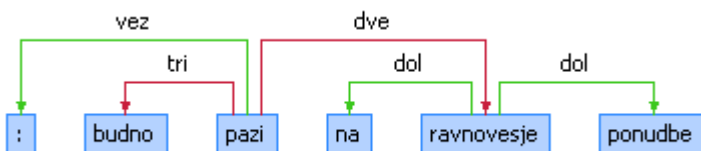


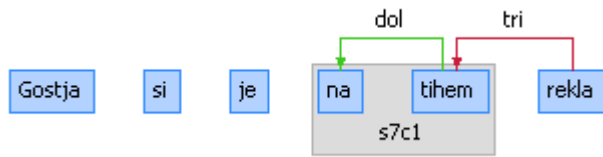
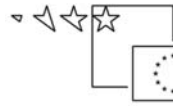
4.2.3 Povezava TRI

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):

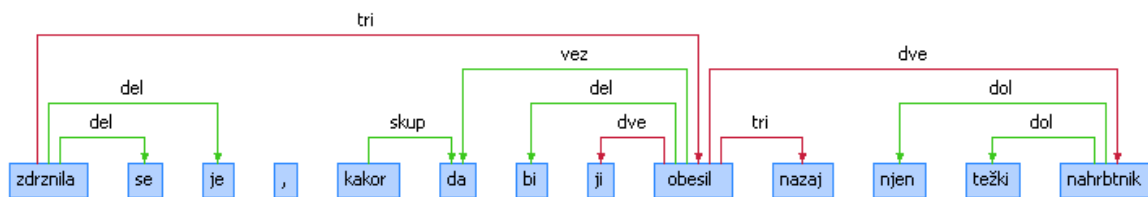
prislovna določila lastnosti identificirajo vprašalnice: *kako (na kakšen način), koliko, kolikokrat, kolikanj, kolikič, katerikrat, čemu podobno, kako (s čim), do katere mere (kako zelo), s kakšnim rezultatom oz. uspehom /.../*

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



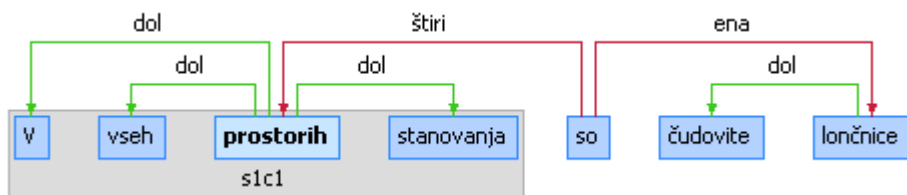
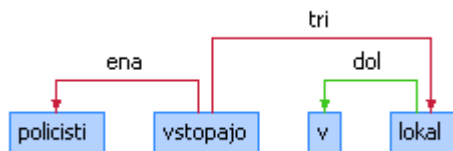


Na medstavčnem nivoju s to povezavo povezujemo tudi lastnostne odvisnike oz. polstavke z ustreznim glavnim stavkom: izhodišče puščice je jedro povedka glavnega stavka, cilj puščice je jedro povedka odvisnika (o lastnostnih odvisnikih piše Slovničica 2004 na straneh 641–643):

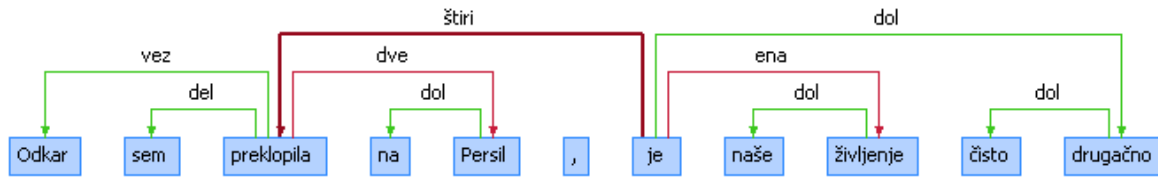
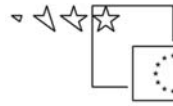


4.2.4 Povezava ŠTIRI

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:

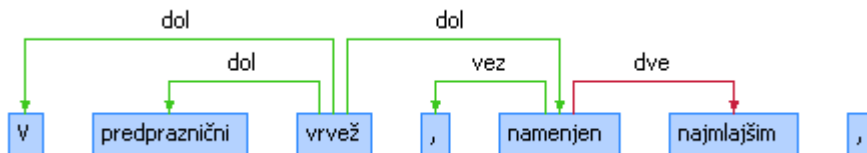


Na medstavčni ravni s to povezavo povezujemo tudi odvisnike ustrežajočih tipov z glavnim stavkom: izhodišče puščice je jedro povedka glavnega stavka, cilj puščice je jedro povedka odvisnika (o odvisnikih piše Slovničica 2004 na straneh 640–645):



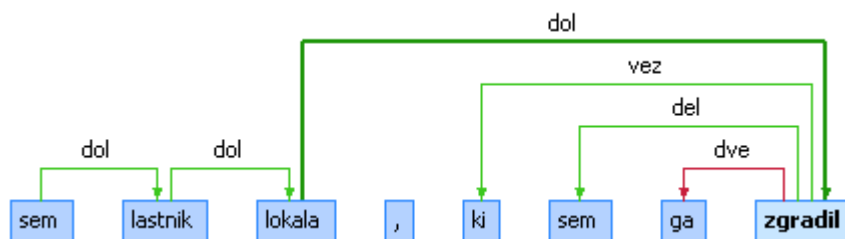
4.2.5 Polstavki

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 3.4.5 Polstavčni desni prilastki.



4.2.6 Prilastkovi odvisniki

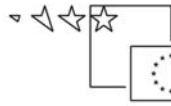
Prilastkove odvisnike povezujemo s povezavo **dol**: izhodišče puščice je beseda, na katero se prilastkov odvisnik nanaša, cilj puščice je jedro povedka odvisnika (o prilastkovem odvisniku piše Slovnica 2004 na straneh 645–646):



4.2.7 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?* Slovnica 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).

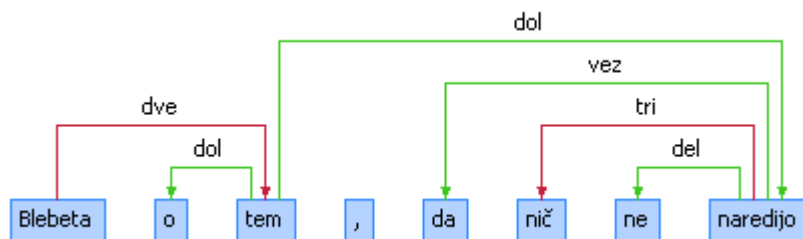


Kadar smo torej v dilemi, je varnejša odločitev za prislovno določilo, tako vsaj pravi Slovničar. Zelo ohlapna pomoč je lahko občutek, da so orodniške ter mestniške zveze pogosteje prislovna določila, roditeljske ter tožilniške bolj predmeti – vendar ti podatki še niso preverjeni na večji količini realnega 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 skladijsko in ne pomensko (tudi primera *kapljati na zabavo* ne bi označili z DVE).

4.2.8 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 4.2.6):



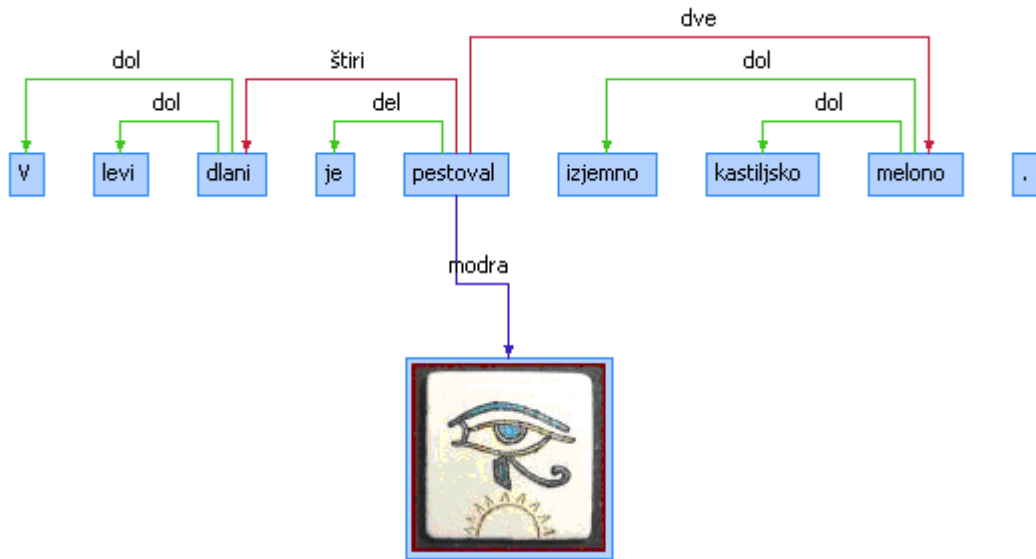
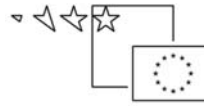
4.3 Povezave tretjega nivoja

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**.

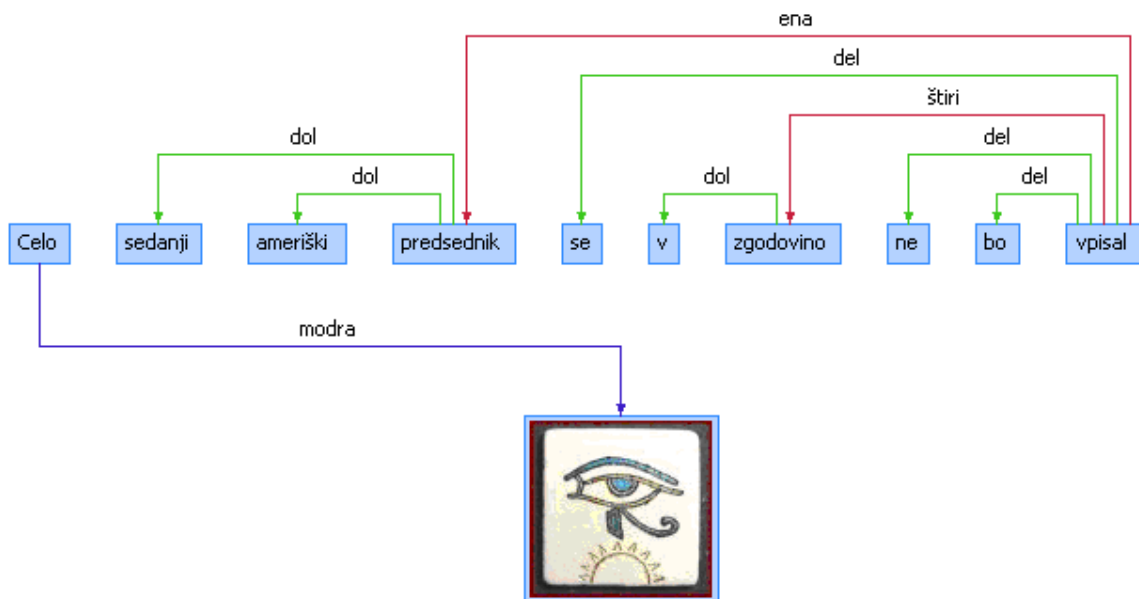
S povezavo *modra* povezujemo:

- 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 [D].

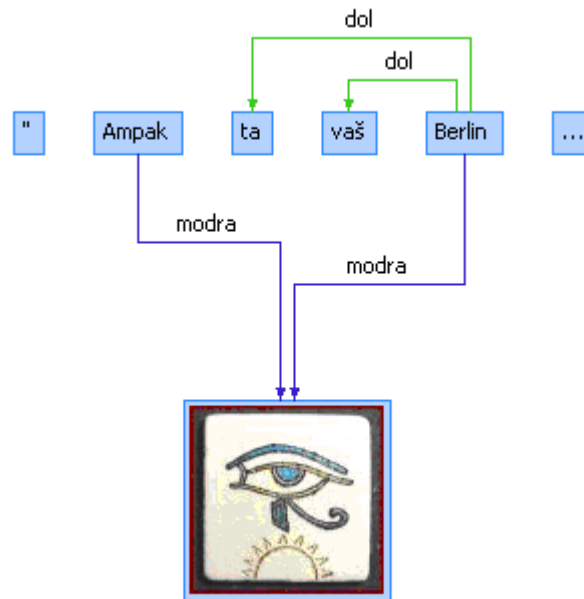
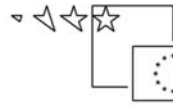
Kadar povezujemo nepovezane dele stavka z metaelementom, je izvor puščice vedno stavčni del, metaelement je cilj:



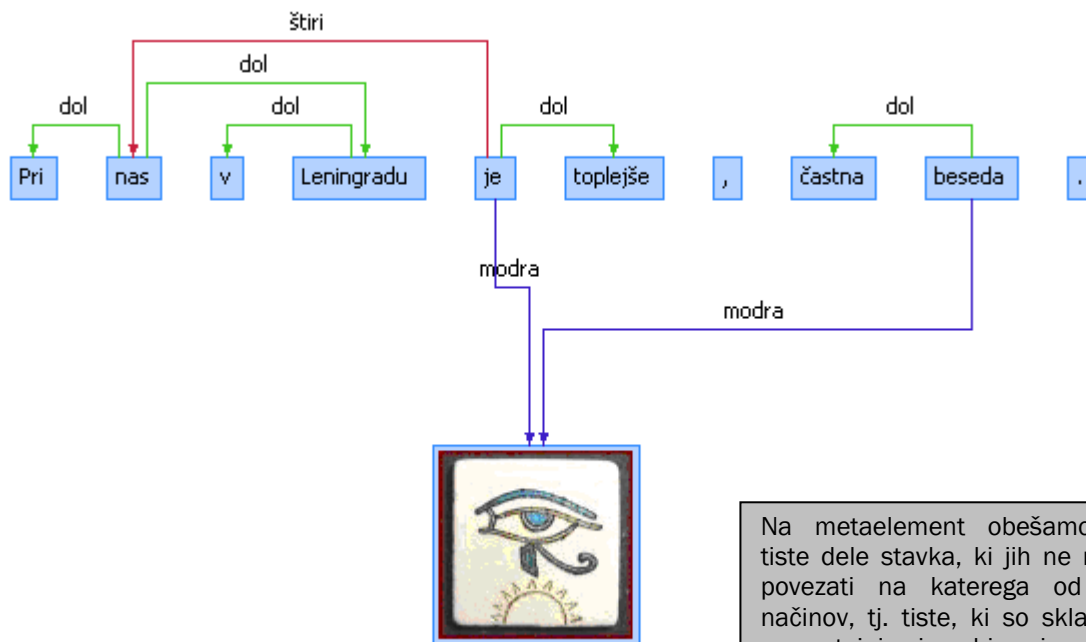
[A]



[B]



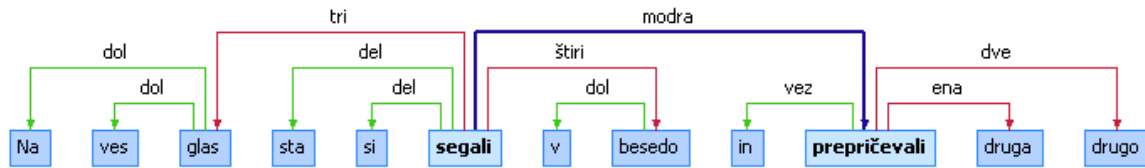
[C]



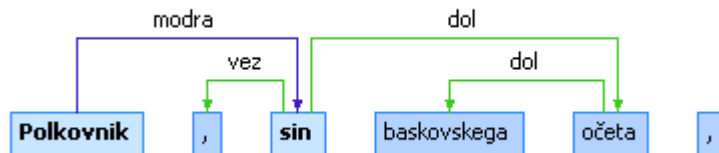
Na metaelement obešamo samo tiste dele stavka, ki jih ne moremo povezati na katerega od drugih načinov, tj. tiste, ki so skladijsko samostojni in bi sicer ostali nepovezani.

[Č]

Prav tako s povezavo *modra* povezujemo med sabo povedkovna jedra prirednih stavkov, in sicer po sistemu, ki velja za označevanja priredij znotraj stavka (4.1.3) – prvi element priredja je izhodišče puščice, naslednji je cilj. Ta tip povezovanja uporabljamo za vsa medstavčna priredja, ne le vezalnega.



Nazadnje s to povezavo povezujemo 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*):



4.4 Posebni primeri

4.4.1 Vezniki

Večdelne veznike povezujemo s povezavo:

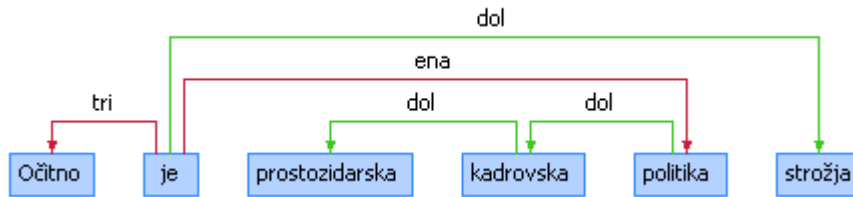
- **skup**, kadar gre za zvezo z vnaprej pripravljenega seznama skupov ali zvezo, sorodno tem, ki se na seznamu pojavljajo,
- **dol**, kadar gre za zvezo, kjer lahko določimo element, ki je v zvezi najbolj ključen oz. »vezniški« (npr. *da/ko* v kombinacijah *kakor da*, *vedno ko* itd.),
- **večdelnih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

4.4.2 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.

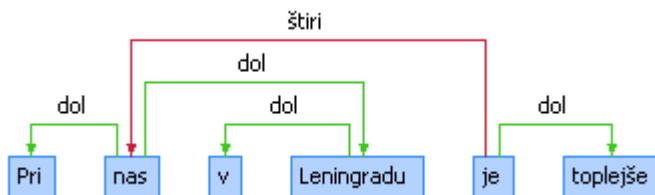
4.4.3 Izražanje mnenja ali ocene

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



4.4.4 Sestavljene predložne zveze

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

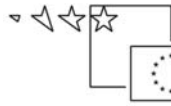


4.4.5 Povedkova določila in povedkovi prilastki

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.

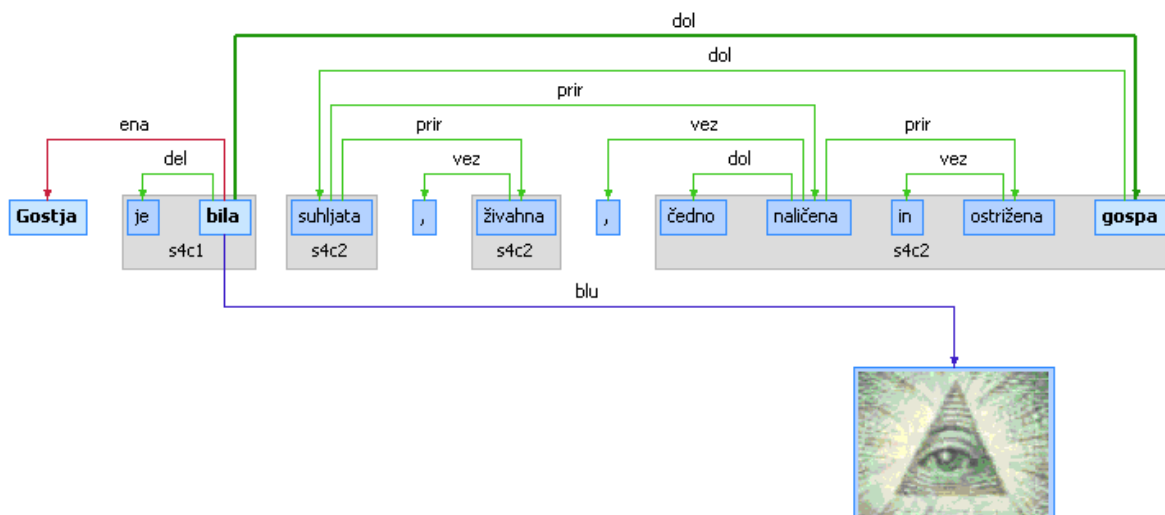
Glede na smernice sistema je označevanje nekoliko poenostavljeno v smislu, da lahko povedkovo določilo nastopa le ob (**nepolnopomenskem**) glagolu *biti* (*bil je ekonomist, nadarjena je bila, njegov načrt je bil, da bi napadli sovražnika, bili so vesele narave, popravljati napake je vaša naloga*), 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 osebko 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 oslabljenih glagolih«⁷ (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, Umrl je mlad* ipd.)

Tipično povedkovo določilo sestoji iz treh elementov, osebka (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 osebka razmerje *ena* ter z jedrom povedkovega določila razmerje *dol*.



4.4.6 Modalni glagoli⁸

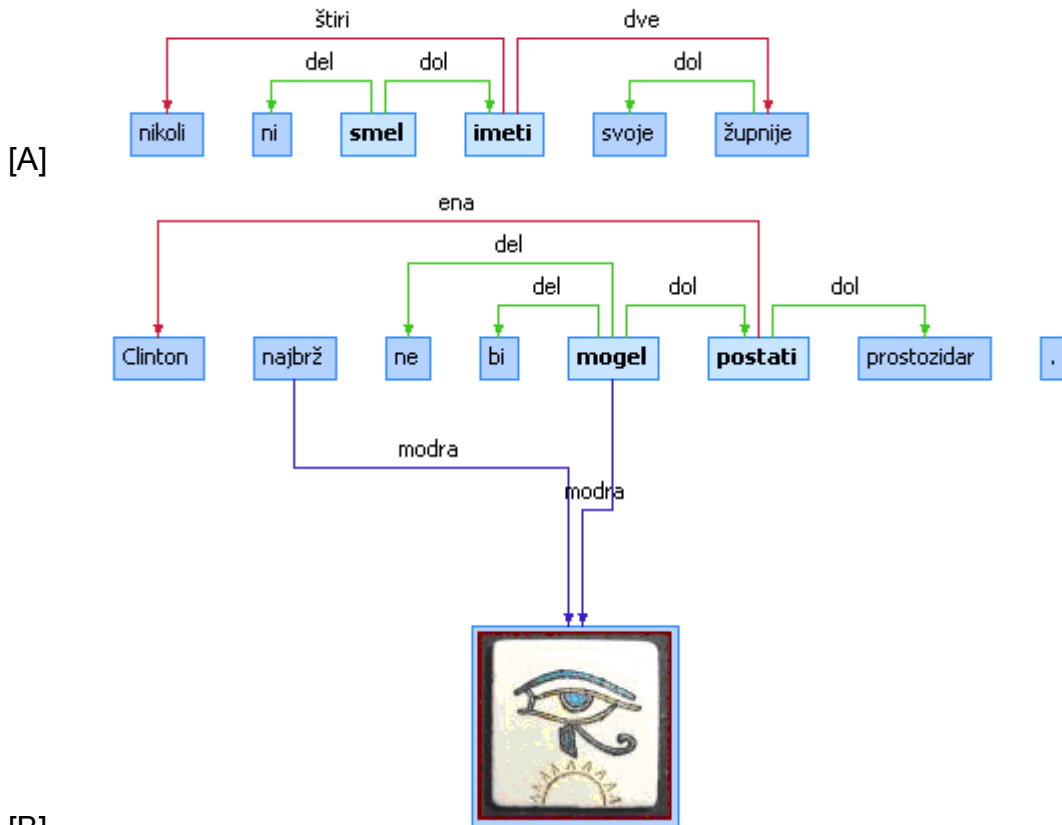
4.4.6.1 morati, smeti, hoteti itd.

⁷ V sistemu JOS so obravnavani kot polnopomenski glagoli.

⁸ Na enak način obravnavamo vse glagole, ki nastopajo v zvezah z nedoločniki in namenilniki.

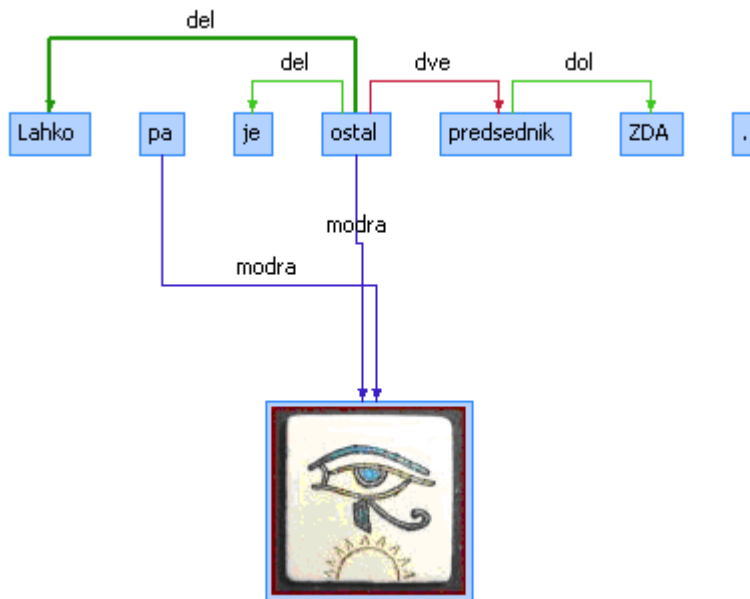
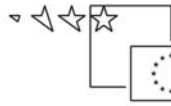


Konstrukcije z modalnimi glagoli označimo tako, da za izhodišče izberemo modalni glagol, nedoločnik oz. namenilnik pa označimo kot določilo modalnega glagola. Kadar konstrukciji sledijo drugi stavčni členi, jih običajno povezujemo iz nedoločnika – slednji je torej izhodišče za povezave ENA, DVE, TRI,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



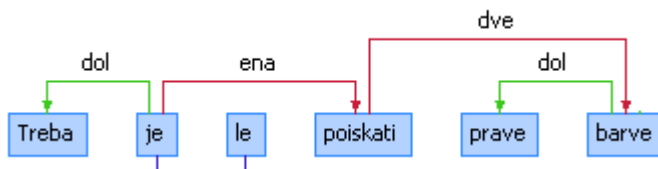
4.4.6.2 lahko je

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**:



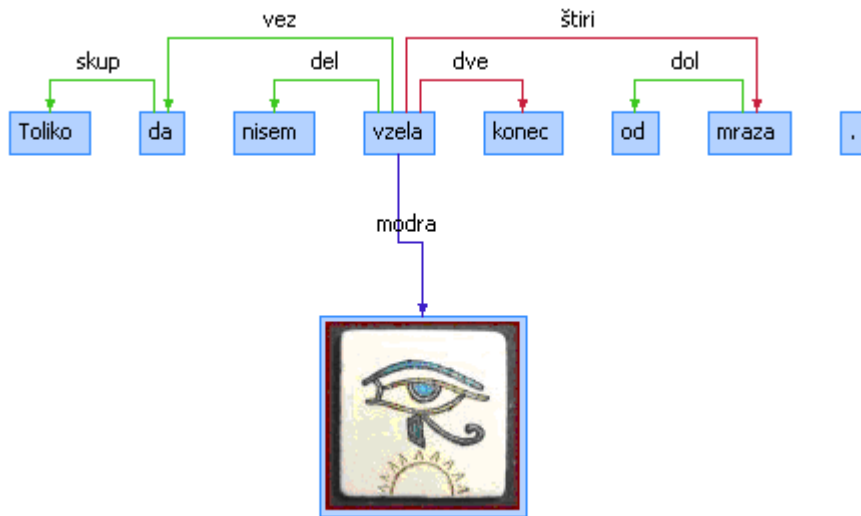
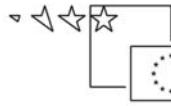
4.4.6.3 treba je

Na tem mestu še primer označevanja zvez tipa *treba je* + *nedoločnik*, ki jih označujemo podobno kot 4.4.6.1:



4.4.7 Idiomi

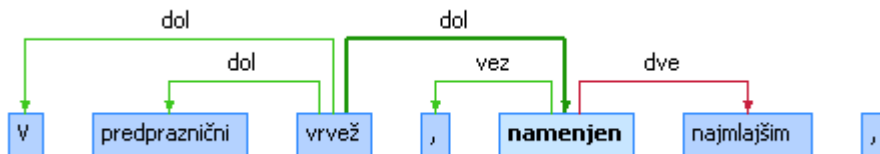
Obravnava idiomov se ne razlikuje od obravnave običajnih stavčnih konstrukcij, kar pomeni, da pri označevanju pomensko komponento glagolske kolokacije namenoma prezremo: zveze *vzeti konec* ali *vzeti usodo v svoje roke* npr. označimo na enak način kot *vzeti kladivo v svoje roke*. Pri določanju oznak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



Posebna pozornost je potrebna pri variantah, kjer je pomembna **skladenjska 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 4.2.7.

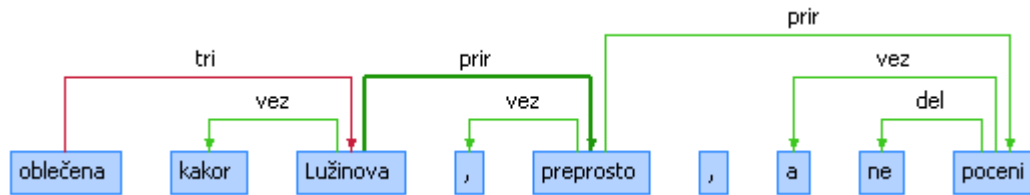
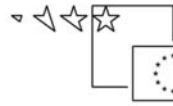
4.4.8 Polstavčni desni prilastki

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 (4.1.3) torej tudi tukaj iščemo trielementno razmerje:



4.4.9 Pojasnjevalne strukture

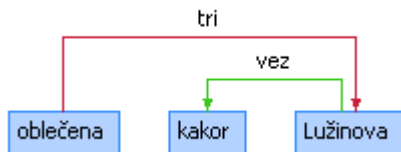
Pojasnjevalne strukture tipa *oblečena kakor Lužinova*, *preprosto, a ne poceni* interpretiramo kot pojasnjevalno priredje (*oblečena kakor Lužinova, IN SICER preprosto, a ne poceni*). Priredje označimo s povezavo *prir*, ki jo tukaj povežemo med prirednima deloma na podoben način, kot bi to storili 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.

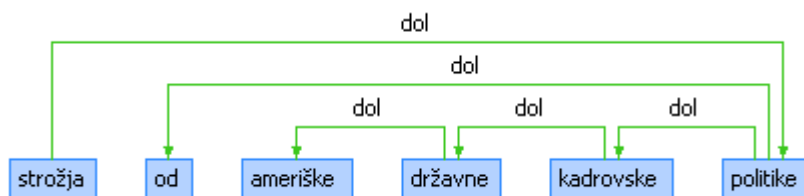
4.4.10 Primerjave

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*.

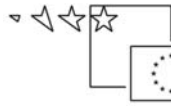
Kadar označujemo primerjavo tipa *strožja od ameriške politike*, jo označujemo s povezavami **dol**, kakor je prikazano spodaj:



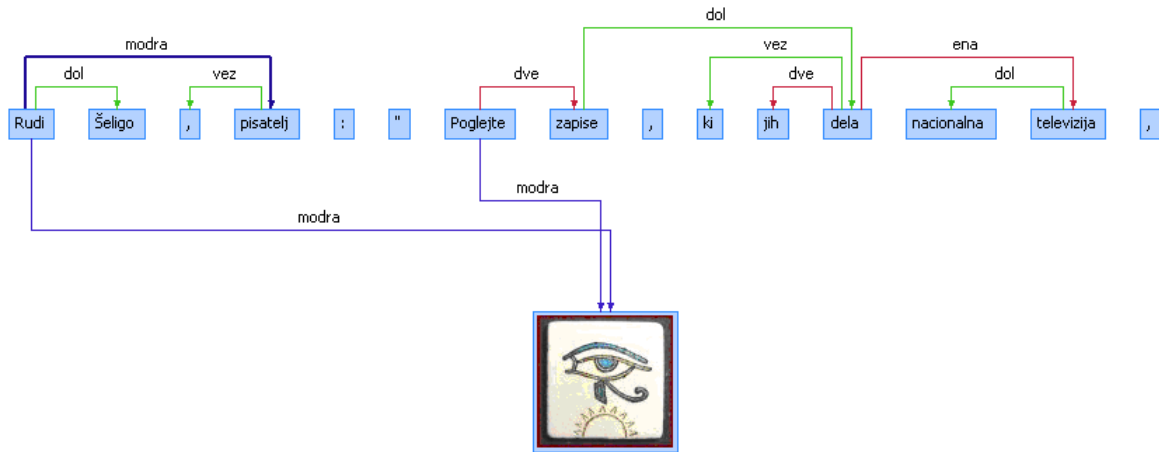
4.4.11 Premi govor

Za označevanje premega govora veljajo naslednja načela:

- narekovajev ne povezujemo v strukturo,
- enako velja za dvopičje,

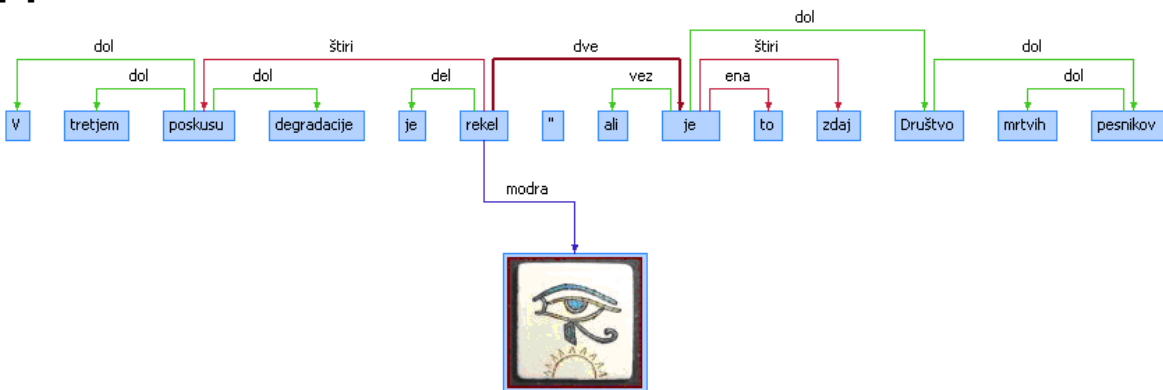


- narekovaji in dvopičja sugerirajo stavčno mejo, ki jo v določenih primerih upoštevamo [A], v drugih ne [B], odvisno od tega, kakšno vlogo igra v stavku navedek (ali ga interpretiramo kot samostojno enoto ali npr. predmetni oz. osebkov odvisnik – *koga ali kaj je rekel?*)
- pika je vsekakor meja, prek 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]

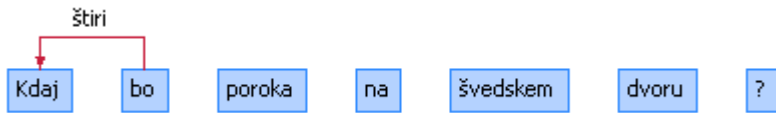


4.4.12 Vprašalnice

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**.

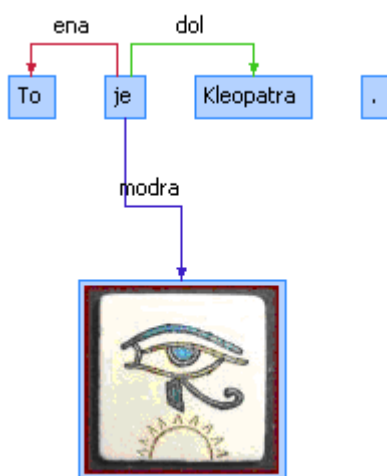


V primerih, ko te besede nastopajo kot vprašalnice (*kdaj bo poroka na švedskem dvoru?*), jih povezujemo s povezavami **drugega nivoja**, in sicer s tisto, ki pritiče stavčnemu členu, po katerem se z vprašalnico vprašamo:

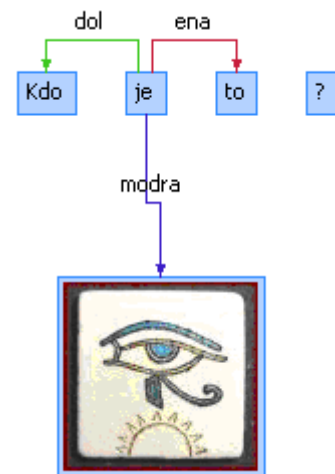


Kadar imamo z določanjem tipa povezav težave, pretvorimo stavek v trdilnega oz. **nadomestimo vprašalnico z ustrežajočo nevprašalno besedo**, označimo povezave, nato te povezave prevedemo nazaj na problematični vprašalni stavek – v gornjem primeru smo pri označevanju dela stavka *kdaj bo poroka?* izhajali 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]:



[A]



[B]

4.4.13 Pasivne konstrukcije

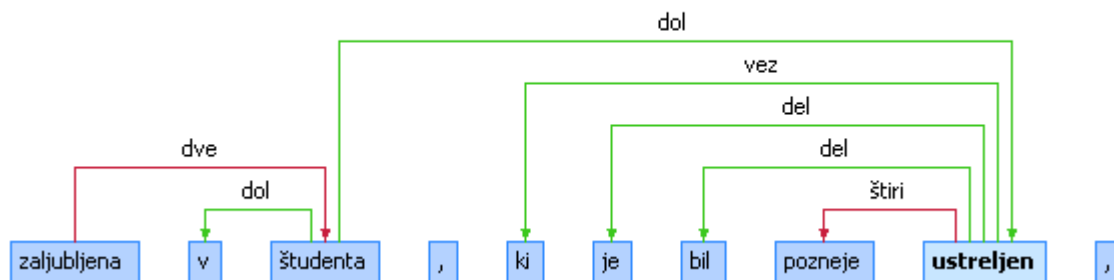
Kadar pri označevanju naletimo na pasivno konstrukcijo (npr. *študent, ki je bil pozneje ustreljen*), izhodišče povezav določimo analogno na tvorbo obliko – v gornjem primeru je izhodišče povezav *del, vez, ŠTIRI* torej beseda *ustreljen* (glej primer spodaj).

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.

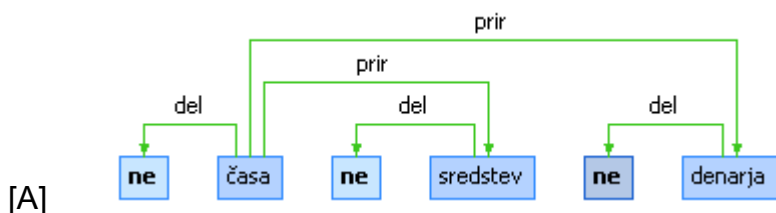
Če se po vsem tem (in posvetu z ostalimi označevalci) še vedno ne moremo odločiti, je varnejša izbira deležnik stanja (ker je pogostejši).



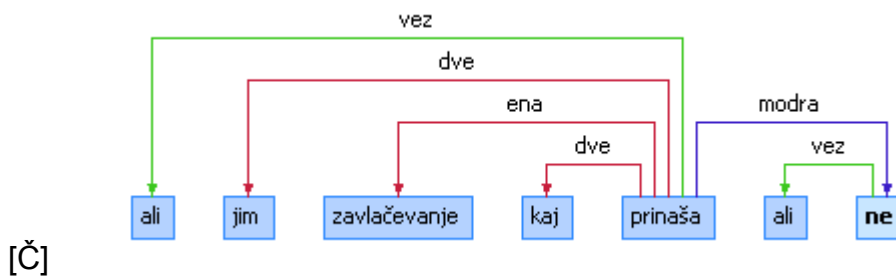
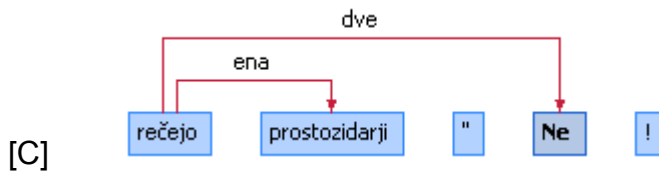
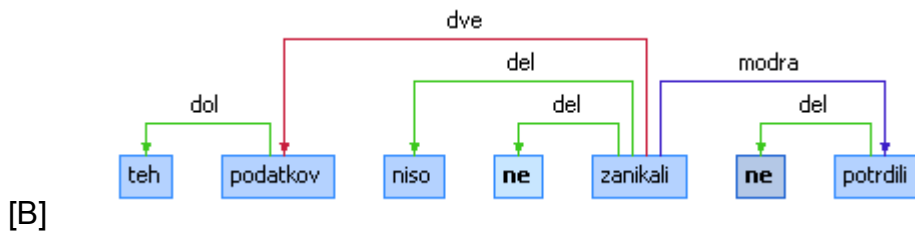
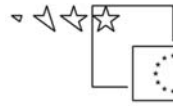
4.4.14 Členek NE⁹

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]

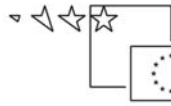


⁹ Na enak način obravnavamo tudi ustrezne zveze z *ni*.



4.4.15 Slovnice napake v besedilih

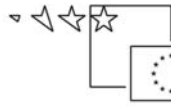
Z deli besedila, ki se ne skladajo z normo slovenskega jezika, se obremenjujemo čim manj. Označevanje nadaljujemo tako, kot da napake ne bi bilo (če gre le za odvečne vejice, jih ignoriramo ipd.), v zapletenejših primerih se posvetujemo z ostalimi označevalci.



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 <i>učitelj/suh</i> 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 <i>menedžerka</i> 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</i> itd.;	<i>Cenijo jo kot delavno</i>

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



6. 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