

06 Odvisnostna skladnja JOS-SYN

Sistem JOS-SYN, ki je bil zasnovan v projektu Jezikoslovno označevanje slovenščine (Erjavec et al. 2010) in uporabljen v projektu Sporazumevanje v slovenskem jeziku (Krek et al. 2020), je namenjen označevanju skladenjskih odnosov v slovenskih povedih. Sistem sledi spoznanjem slovenskega jezikoslovja (zlasti slovnici Toporišič 2004), obenem pa temeljnim idejam, ki jih zarisujejo obstoječi sistemi odvisnostnega označevanja. Ključna lastnost sistema je, da upošteva informacije, ki jih prinašajo oblikoskladenjske oznake JOS oz. njihova sodobna različica MULTEXT-East v6 (Erjavec 2012). Na skladenjski ravni tako dodajamo samo informacije, ki jih še ni pokrila oblikoskladnja, kar omogoči robusten in pregleden označevalni sistem.

- [Predstavitev oznak](#)
- [Označevalne smernice](#)
- [Reference in povezave](#)

Predstavitev oznak

V tem poglavju so navedene in strnjeno predstavljene oznake sistema JOS-SYN. Podrobnejšo predstavitev najdete v smernicah v poglavju Označevalne smernice.

Oznaka	Opis povezave
dol	S povezavo dol povezujemo določujoči in določani del podredne besedne zveze, tj. jedro in različne vrste prilastkov. Izhodišče je jedro zveze, cilj pa beseda, ki jedro določa. Tipično jo uporabljamo pri samostalniških, pridevniških oz. prislovnih besednih zvezah ter za povezovanje delov povedka pri modalnih glagolih, povedkovem določilu in povedkovem prilastku. Na medstavčni ravni s to povezavo povezujemo prilastkove odvisnike z ustreznim glavnim stavkom.
del	S povezavo del povezujemo dele besednih zvez, pri katerih ne govorimo o prototipnem jedru ter določilu zveze, pač pa zgolj o delih zveze. Tipično jo uporabljamo za dele zloženega povedka, pri čemer je izvor osebna glagolska oblika oz. deležnik na -l, cilj so morfemi »ne«, »se«, »si«, »bi« oz. oblike glagola »biti« v sestavljenih nesedanjiških časih, tj. »bo«, »je« itd.
prir	S povezavo prir povezujemo dele priredij na besednozvezni ravni. Pri besednih zvezah povezujemo jedro prvega dela priredja z jedrom drugega dela priredja, pri čemer je izhodišče v levem delu, cilj pa v desnem delu priredja.
vez	Povezavo vez uporabljamo v kombinaciji s povezavo prir tako, da trije elementi, povezani s prir in vez, tvorijo trikotnik. S povezavo vez povezujemo jedro drugega dela besednozveznega priredja in priredni veznik ali priredno ločilo, če veznika ni.
skup	S povezavo skup povezujemo besede, ki imajo zelo močno tendenco po sopojavljanju in tvorijo neke vrste večbesedne enote, ne gre pa za klasične podredne ali priredne besedne zveze. S to povezavo tipično označujemo besede, ki imajo variantni zapis skupaj ali narazen, nekatere večbesedne veznike in podobne večbesedne enote.

Oznaka	Opis povezave
ena	S povezavo ena povezujemo dele stavkov oz. povedi, ki bi jih tradicionalno lahko opredelili kot osebek, vendar tako označena razmerja ne ustrezajo povsem klasični definiciji tega pojava. Na znotrajstavčni ravni povezuje povedkovo vozlišče z osebkom stavka, pri čemer je izhodišče razmerja jedro povedka, cilj razmerja pa jedro osebk. Na medstavčni ravni s to povezavo povezujemo osebkove odvisnike z ustreznim glavnim stavkom. Izhodišče je jedro povedka nadrednega stavka, cilj je jedro povedka v odvisniku.
dve	S povezavo dve povezujemo dele stavkov oz. povedi, ki bi jih tradicionalno lahko opredelili kot predmet stavka ali predmetni odvisnik, vendar tako označena razmerja ne ustrezajo povsem klasični definiciji teh dveh pojavov. Na znotrajstavčni ravni povezuje vozlišče povedka in predmet stavka. Izhodišče razmerja je jedro povedka, cilj je jedro predmeta. Na medstavčni ravni s to povezavo povezujemo predmetne odvisnike z ustreznim glavnim stavkom. Izhodišče je jedro povedka glavnega stavka, cilj je jedro povedka v odvisniku.
tri	S povezavo tri povezujemo dele stavkov oz. povedi, ki bi jih tradicionalno lahko opredelili kot prislovno določilo lastnosti ali lastnostni odvisnik, vendar tako označena razmerja ne ustrezajo povsem klasični definiciji teh dveh pojavov. Na znotrajstavčni ravni povezuje vozlišče povedka in prislovno določilo lastnosti. Izhodišče razmerja je jedro povedka, cilj je jedro prislovnega določila. Na medstavčnem nivoju s to povezavo povezujemo lastnostne odvisnike oz. polstavke z ustreznim glavnim stavkom. Izhodišče je jedro povedka glavnega stavka, cilj je jedro povedka v odvisniku.
štiri	S povezavo štiri povezujemo dele stavkov oz. povedi, ki bi jih tradicionalno lahko opredelili kot prislovno določilo ali prislovni odvisnik (razen prislovnega določila lastnosti ali lastnostnega odvisnika), vendar tako označena razmerja ne ustrezajo povsem klasični definiciji teh pojavov. Na znotrajstavčni ravni povezuje vozlišče povedka in prislovno določilo. Izhodišče razmerja je jedro povedka, cilj je jedro prislovnega določila. Na medstavčni ravni s to povezavo povezujemo odvisnike ustrezajočih tipov z ustreznim glavnim stavkom. Izhodišče je jedro povedka glavnega stavka, cilj je jedro povedka v odvisniku.
modra	S povezavo modra povezujemo abstraktno vozlišče stavka oz. povedi z elementi, iz katerih potekajo nadaljnje drevesnične povezave. S to povezavo so tipično povezani stavčni povedki, eliptični deli povedi brez povedka ali členkovni elementi znotraj stavka. S to povezavo so na vozlišče povezane tudi vse pojavnice (besede ali ločila), ki nimajo eksplicitne skladišne vloge v povedi.

Označevalne smernice

V tem poglavju so zbrane označevalne smernice JOS-SYN. Smernice so razvrščene od nastarejše različice do zadnje, ažurne različice.

Različica 2.0 (02-2023)

projekt [Razvoj slovenščine v digitalnem okolju](#)

ARHAR HOLDT, Špela, TERČON, Luka, KREK, Simon, LEDINEK, Nina, MOŽE, Sara, SAKSIDA, Amanda, HOLZ, Nanika, 2023: *Navodila za skladiščno označevanje slovenščine po sistemu JOS-SYN*.

Različica 2.0. Rezultat projekta Razvoj slovenščine v digitalnem okolju. [\[DOCX\]](#) [\[PDF\]](#)

Različica 1.0 za nestandardno slovenščino (21-12-2016)

projekt [Janes - Viri, orodja in metode za raziskovanje nestandardne spletne slovenščine](#)

HOLOZAN, Peter, KREK, Simon, PIVEC, Matej, RIGAČ, Simon, ROZMAN, Simon, VELUŠČEK, Aleš, ARHAR HOLDT, Špela, 2016: *Smernice za označevanje z odvisnostnim sistemom JOS: nestandardna slovenščina*. Različica 1.0. Rezultat projekta Viri, orodja in metode za raziskovanje nestandardne spletne slovenščine. [\[DOCX\]](#) [\[PDF\]](#)

Različica 1.0

projekt [Sporazumevanje v slovenskem jeziku](#)

HOLOZAN, Peter, KREK, Simon, PIVEC, Matej, RIGAČ, Simon, ROZMAN, Simon, VELUŠČEK, Aleš, 2008: *Specifikacije za učni korpus*. Različica 1.0. Kazalnik K2 projekta Sporazumevanje v slovenskem jeziku. [\[PDF\]](#)

Reference in povezave

V tem poglavju so zbrane relevantne reference in povezave na projekte, v katerih se je označevalni sistem razvijal in uporabljal.

Projekti, na katerih se je razvijal označevalni sistem

[Sporazumevanje v slovenskem jeziku](#)

[Janes: Viri, orodja in metode za raziskovanje nestandardne spletne slovenščine](#)

[Razvoj slovenščine v digitalnem okolju](#)

Učni korpus z ročno pregledanimi oznakami JOS-SYN

Arhar Holdt, Špela; Krek, Simon; Dobrovoljc, Kaja; Erjavec, Tomaž; Gantar, Polona; Čibej, Jaka; Pori, Eva; Terčon, Luka; Munda, Tina; Žitnik, Slavko; Robida, Nejc; Blagus, Neli; Može, Sara; Ledinek, Nina; Holz, Nanika; Zupan, Katja; Kuzman, Taja; Kavčič, Teja; Škrjanec, Iza; Marko, Dafne; Jezeršek, Lucija; Zajc, Anja, 2022, Training corpus SUK 1.0, Slovenian language resource repository CLARIN.SI, ISSN 2820-4042, <http://hdl.handle.net/11356/1747>.

Orodje Q-CAT za ročno označevanje s sistemom JOS-SYN

Brank, Janez, 2022, Q-CAT Corpus Annotation Tool 1.4, Slovenian language resource repository CLARIN.SI, ISSN 2820-4042, <http://hdl.handle.net/11356/1684>.

Reference

Arhar Holdt, Špela; Fišer, Darja; Erjavec, Tomaž in Krek, Simon. Syntactic annotation of Slovene CMC: first steps. Proceedings of the 4th Conference on CMC and Social Media Corpora for the Humanities, 27.–28. september 2016, Ljubljana, Slovenia, 2016, str. 3–6. <http://nl.ijs.si/janes/cmc-corpora2016/proceedings/>.

Erjavec, Tomaž; Fišer, Darja; Krek, Simon in Ledinek, Nina. 2010. The JOS Linguistically Tagged Corpus of Slovene. V: Proceedings of the Seventh conference on International Language Resources and Evaluation (LREC'10), Valeta, Malta, Maj. European Language Resources Association (ELRA). http://www.lrec-conf.org/proceedings/lrec2010/pdf/139_Paper.pdf

Erjavec, Tomaž. 2012. MULTEXT-East: morphosyntactic resources for Central and Eastern European languages. Language Resources and Evaluation, 46(1): 131–142.

Krek, Simon; Erjavec, Tomaž; Dobrovoljc, Kaja; Gantar, Polona; Arhar Holdt, Špela; Čibej, Jaka in Brank, Janez. The ssj500k training corpus for Slovene language processing. V: Fišer, D. in Erjavec, T. Jezikovne tehnologije in digitalna humanistika: zbornik konference: 24.–25. september 2020, Ljubljana, Slovenija. Ljubljana: Inštitut za novejšo zgodovino, 2020. Str. 24–33. http://nl.ijs.si/jtdh20/pdf/JT-DH_2020_Krek-et-al_The-ssj500k-Training-Corpus-for-Slovene-Language-

[Processing.pdf](#).

Toporišč, Jože. (2004): Slovenska slovnica. Maribor: Obzorja.