12 Jezikovni popravki (KOST)

Označevalni sistem KOST je bil zasnovan skupaj s korpusom slovenščine kot tujega jezika KOST (Stritar Kučuk 2022) in je namenjen kategorizaciji jezikovnih napak v besedilih, ki so jih napisali govorci slovenščine kot drugega oz. tujega jezika. Označevalni sistem je hierarhično urejen v dve stopnji: najprej so popravki opredeljeni glede na jezikovno ravnino, sledi pa opredelitev splošnega tipa popravka oz. določitev besedne vrste. Dvodelne oznake omogočajo robustno analizo, ki ji mora slediti podrobnejši ročni pregled.

- Predstavitev oznak
- Označevalne smernice
- Reference in povezave

Predstavitev oznak

V tem poglavju so strnjeno navedene oznake jezikovnih popravkov po sistemu KOST.

Oznaka	Jezikovna ravnina	Tip popravka/besedna vrsta
Z-LOC	zapis	ločilo
Z-CRK	zapis	črkovanje
Z-SN	zapis	skupaj/narazen
Z-MV	zapis	mala/velika začetnica
Z-KR	zapis	krajšave
B-SAM	besedišče	samostalnik
B-GLAG	besedišče	glagol
B-ZAIM	besedišče	zaimek
B-PRID	besedišče	pridevnik
B-PRISL	besedišče	prislov
B-PRED	besedišče	predlog
B-VEZ	besedišče	veznik
B-OST	besedišče	ostalo
O-SAM	oblika	samostalnik
O-GLAG	oblika	glagol
O-ZAIM	oblika	zaimek
O-PRID	oblika	pridevnik
O-PRISL	oblika	prislov
O-OST	oblika	ostalo
S-STR	skladnja	struktura
S-BR	skladnja	besedni red
S-IZP	skladnja	izpuščeni jezikovni elementi
S-ODV	skladnja	odvečni jezikovni elementi
POV	/	povezani popravki
[???]	/	nerazumljivo, nejasen popravek

Označevalne smernice

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

Različica 1.0 (04-2022)

projekt Razvoj slovenščine v digitalnem okolju

STRITAR KUČUK, Mojca, 2023: *KOST 1.0: Priročnik za označevanje napak, delovna verzija*. Različica 1.0. [PDF]

Reference in povezave

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

Projekti, na katerih se je razvijal označevalni sistem

Razvoj slovenščine v digitalnem okolju

Trenutna različica korpusa KOST

STRITAR KUČUK, Mojca, ŠTER, Helena, PISEK, Staša, PETRIC LASNIK, Ivana, KETE MATIČIČ, Jana, PIRIH SVETINA, Nataša, PREGLAU, Daniela, ARHAR HOLDT, Špela, KRSNIK, Luka, ERJAVEC, Tomaž, 2023, *Slovene learner corpus KOST 1.0,* Slovenian language resource repository CLARIN.SI, ISSN 2820-4042, http://hdl.handle.net/11356/1753.

Orodje CJVT Svala za označevanje korpusov z jezikovnimi popravki

ARHAR HOLDT, Špela, KOSEM, Iztok, STRITAR KUČUK, Mojca, KRSNIK, Luka, JOVAN, Leon Noe, 2022: *CJVT Svala* (Kazalnik projekta Razvoj slovenščine v digitalnem okolju), v1.0, https://orodja.cjvt.si/svala/, dostop 2. 3. 2023.

Literatura o korpusu KOST

STRITAR KUČUK, Mojca, 2022: *KOST med korpusi usvajanja tujega jezika*. Obdobja 41: Na stičišču svetov: slovenščina kot drugi in tuji jezik. 323–334. https://centerslo.si/wp-content/uploads/2022/11/Stritar-Kucuk_Obdobja-41.pdf

ARHAR HOLDT, Špela, KOSEM, Iztok, STRITAR KUČUK, Mojca, 2022: *Metode in orodja za lažjo pripravo korpusov usvajanja jezika*. Obdobja 41: Na stičišču svetov: slovenščina kot drugi in tuji jezik. 23–30. https://centerslo.si/wp-content/uploads/2022/11/Arhar-Holdt-et-al_Obdobja-41.pdf

STRITAR KUČUK, Mojca, 2020: *Modul Leto plus – prvi korak do korpusa slovenščine kot tujega jezika* . Zbornik konference Jezikovne tehnologije in digitalna humanistika 2020. 131–135. http://nl.ijs.si/jtdh20/pdf/JT-DH_2020_StritarKucuk_Modul-Leto-plus%e2%80%93prvi-korak-do-korpusa-slovenscine-kot-tujega-jezika.pdf