

# Uvod

Kolokacija kot tipična sopojavitev vsaj dveh leksikalnih enot je jezikovni pojav, ki je že dolgo predmet raziskav na različnih področjih, od korpusnega jezikoslovja in leksikografije do učenja in poučevanja jezikov. Pa vendar izzivov za jezikoslovce ne zmanjka – korpusi postajajo vse večji, kolokacijskih podatkov je posledično vse več in iščejo se metode za čim bolj učinkovito prepoznavanje, spremljanje in selekcijo kolokacij za različne namene. Ob naraščajočem vključevanju kolokacij v slovarje in druge leksikalne vire pa se zastavljajo tudi vprašanja o odnosu uporabnikov do tovrstnih podatkov.

Pričujoča monografija, ki je nastala kot rezultat temeljnega raziskovalnega projekta KOLOS (*Kolokacije kot temelj jezikovnega opisa: semantični in časovni vidiki*), prinaša v znanstveni prostor pomembna nova teoretska in metodološka spoznanja o različnih vidikih kolokacij v slovenskem jeziku. Hkrati so teoretska razmišljanja, rezultati opravljenih analiz in evalvacij ter novo nastali viri in orodja relevantni za vpenjanje slovenske jezikoslovne skupnosti v mednarodni prostor.

V prvem prispevku Polona Gantar, Simon Krek in Iztok Kosem predstavijo opredelitev kolokacije za namene izdelave digitalnih slovarskih virov za slovenščino. Pri tem opredelijo tri ključne kriterije, ki določajo kolokacijo, nato pa prikažejo mesto kolokacije v odnosu do drugih tipov večbesednih enot. Prispevek izpostavi koncept slovarske relevantnosti kolokacij in merila za njihovo vključevanje v slovarske vire.

Eva Pori in Iztok Kosem v svojem prispevku opišeta rezultate in ugotovitve obsežne evalvacije avtomatskega luščenja kolokacijskih podatkov z uporabo orodja Sketch Engine z jezikoslovnega in metodološkega vidika. Temeljni namen prispevka je predstaviti glavne prednosti in slabosti postopka prepoznavne kolokacij, ki se uporablja

pri izdelavi številnih slovarskih in drugih jezikovnih virov, ter predlagati izboljšave.

S statistično problematiko kolokacij se ukvarjajo Iztok Kosem, Nataša Logar, Kaja Dobrovoljc in Nikola Ljubešič, ki predstavijo rezultate treh ločenih preizkusov, od katerih se dva ukvarjata z razvrščanjem kolokacij (besedne vložitve in deltaP), eden pa s statistično prepoznavo slovarsko nerelevantnih kolokacij (razpršenost). Prispevek je inovativen v tem, da gre za prvo sistematično evalvacijo treh statističnih metod na slovenskih podatkih.

Prispevek Špele Arhar Holdt zaokrožuje analize statističnega vidika kolokacij s predstavitvijo rezultatov anketne raziskave med 457 uporabniki, ki so podali mnenje o najbolj optimalnem razvrščanju kolokacij in tako imenovanem jagodnem izboru, tj. kolokacijskih podatkih, ki so v kolokacijskem slovarju uporabnikom predstavljeni najprej oz. na najbolj vidnem mestu.

Semantični vidik kolokacij je naslovljen v prispevku Iztoka Kosma in Eve Pori, ki predstavita ontologijo semantičnih tipov za samostalnike SLONEST-sam na podlagi slovenskega gradiva. Ontologija bo služila kot osnova pri opredeljevanju pomenskih konceptov v digitalni slovarski bazi, pa tudi pri (avtomatskem) gručenju kolokacij in prepoznavanju posameznih pomenov. Avtorja podrobno opišeta vse krovne kategorije in njihove podkategorije ter izpostavita podobnosti oz. razlike glede na obstoječe ontologije, ki se uporabljajo v leksikografske namene, v mednarodnem in slovenskem prostoru.

Prispevek Iztoka Kosma in Jake Čibeja prinaša v slovenski prostor pomembne osvetlitve o diahroni rabi oz. časovnih trendih kolokacij v slovenščini. Avtorja prepoznavata različne vzorce časovnih trendov z vidika relevantnosti za jezikovni opis, pa tudi z vidika (nadaljnjih) analiz obnašanja kolokacij glede na družbene trende, besedilne žanre, socio-politično situacijo itd. V prispevku prikažeta metode zaznavanja naraščajoče oz. padajoče rabe kolokacij in z njimi povezanih pomenov. Metoda zaznavanja trendov pa omogoča tudi prepoznavanje korpusnega šuma in slovarsko nerelevantnih kolokacij, kar je metodološko uporabno tudi pri izdelavi in posodabljanju jezikovnih virov.

Eva Pori, Iztok Kosem, Jaka Čibej in Špela Arhar Holdt predstavijo rezultate kvalitativne jezikoslovne analize uporabniške evalvacije vmesnika Kolokacijskega slovarja sodobne slovenščine, pri kateri je bila uporabljena metoda glasnega razmišljanja. Prispevek prinaša številna nova spoznanja glede organizacije podatkov in njihove predstavitve v spletnem vmesniku, pa tudi o odnosu uporabnikov do avtomatsko pridobljenih podatkov. Prispevek ponudi tudi konkretna priporočila za izboljšavo slovarskega vmesnika pri nadaljnjih posodobitvah vira.

Podrobne analize uporabniških spoznanj se loteva tudi prispevek Špele Arhar Holdt, ki predstavlja rezultate uporabniške evalvacije kolokacijskih podatkov v Slovarju sopomenk sodobne slovenščine. Avtorica v prispevku uporabniško oceno podrobno analizira z jezikoslovnega vidika, sinergijo obeh vidikov pa združi v konkretna priporočila za izboljšavo metodologije luščenja in načina vključevanja kolokacijskih podatkov v Slovar sopomenk sodobne slovenščine.

Rdeča nit prispevkov v monografiji je njihova izhodiščna dvojna naravnost: po eni strani se ukvarjajo z jezikoslovno analizo kolokacij z več vidikov ter oceno njihove jezikoslovne ali slovarske relevantnosti, po drugi strani pa ponujajo konkretna metodološka priporočila ali rešitve za izboljšavo postopkov pri pripravi slovarskih virov. Poleg temeljnega namena, ki smo ga zasledovali pri pripravi monografije, namreč prikazati rezultate projekta, nas je pri izboru prispevkov vodila tudi želja po predstavitvi problematike, ki odpira in spodbuja nadaljnje raziskave tega v jeziku tako pomembnega pojava. Vsako novo spoznanje na tem področju ima namreč vrednost tudi za nadaljnjo izgradnjo jezikovne infrastrukture, ki se neposredno kaže v možnosti izboljšanja obstoječih in nastanku novih leksikalnih virov za slovenščino.

Ob izidu monografije bi se rad zahvalil vsem recenzentom, ki so pripomogli k izboljšanju prispevkov, tehničnim sodelavcem, brez katerih priprava podatkov ne bi potekala tako hitro oz. sploh ne bi bila mogoča, in številnim sodelujočim v uporabniških raziskavah.

Še posebna zahvala gre sodelavki Evi Pori, ki mi je priskočila na pomoč v zaključnih korakih, pa tudi dajala prepotrebno spodbudo pri spopadanju z uredniškimi izzivi.

Rad bi se zahvalil tudi Znanstveni založbi Filozofske fakultete za vso potrpežljivost in pomoč pri procesu dokončne priprave monografije.