

SLOVENŠČINA 2.0: »SLOVNIČNE RAZISKAVE ZA JEZIKOVNI OPIS«

Špela ARHAR HOLDT

Center za jezikovne vire in tehnologije

Polona GANTAR

Filozofska fakulteta Univerze v Ljubljani

Vojko GORJANC

Filozofska fakulteta Univerze v Ljubljani

Robert GROŠELJ

Filozofska fakulteta Univerze v Ljubljani

Arhar Holdt, Š., Gantar, P., Gorjanc, V., Grošelj, R. (2018): Slovenščina 2.0: »Slovníčne raziskave za jezikovni opis«. Slovenščina 2.0, 6 (2): i–iv.

DOI: <http://dx.doi.org/10.4312/slo2.0.2018.2.i-iv>.

Letošnja tematska številka revije *Slovenščina 2.0* vsebuje sedem razprav o izbranih slovníčnih pojavih, ki so jih pripravili raziskovalke in raziskovalci z različnih slovenskih inštitucij. Poleg razprav je na voljo zapis strokovnega panela, na katerem so sodelujoči (Kozma Ahačič, Irena Krapš Vodopivec, Simon Krek, Marko Stabej, Rok Žaucer in Špela Arhar Holdt) razmišljali o bodočem slovníčnem opisu za slovenski jezik. Panel slikovito predstavlja duha časa, v katerem tematska številka nastaja: digitalna sodobnost zahteva ne samo vsebinske posodobitve slovníčnega opisa za slovenščino, ampak tudi odprto dostopnost, strojno berljivost in povezljivost slovníčnih podatkov, kot tudi njihovo premišljeno zasnovanost za različne vrste (strokovnih in uporabniških) potreb. Zato tudi ni naključje, da zbrani prispevki v luči novih izzivov na slovníčno problematiko gledajo aplikativno, empirično in metodološko inovativno. Prijazno vabljeni k branju.

Špela Arhar Holdt in **Jaka Čibej** predstavljata metodologijo za avtomatsko pridobivanje oblikoslovnih vzorcev iz leksikona besednih oblik Sloleks. Metoda razvrsti oblikoslovne paradigme glede na spremenljivi in nespremenljivi del oblike, z ročnim pregledom pa so vzorci urejeni in v strojno berljivi obliki pripravljene za nadaljnjo rabo. Dodana vrednost opisanega dela so natančno opredeljene nedoslednosti, pomanjkljivosti in napake leksikonskih podatkov, ki bodo ob prihajajoči posodobitvi tega izredno uporabnega jezikovnega vira lahko odpravljene.

Kaja Dobrovoljc piše o formulaičnih besednih nizih, tj. neprekinjenih nizih dveh ali več besed z izstopajočo pogostostjo pojavljanja, ne glede na njihovo strukturno, pomensko ali zaznavno celovitost. Avtorica predstavi metodologijo strojnega pridobivanja nizov, nato pa podatke kategorizira in opiše glede na njihovo zgradbo ter funkcijo. Prispevek je pomemben, ker utemeljuje jezikoslovni interes za tovrstno gradivo, ki je v preteklosti kljub visoki pogostosti v jezikovni rabi ostajalo na obrobju jezikovnega opisa.

Nataša Logar podprto z izčrpnim pregledom področja in na veliki količini korpusnih zgledov raziše pojem besedotvorne produktivnosti. Avtorica izluščene samostalniške kandidate za izbranih 10 samostalniških pripon ročno selekcionira, opredeli njihovo produktivnost po uveljavljeni Baayenovi formuli in jih analizira. Ključnega pomena so smernice, kako produktivost upoštevati pri slovničnem opisu, kot tudi konkretna opozorila na stilistične, zvrstne, fonološke in številne druge vidike, ki jih je pri pripravi nove slovnice smiselno upoštevati.

Agnes Pisanski Peterlin in **Tamara Mikolič Južnič** s primerjavo gradiva pisnega in govornega referenčnega korpusa raziskujeta primere jezikovne rabe, pri katerih se v vlogi osebka pojavlja osebni zaimsek. Čeprav je takšna raba za slovenščino netipična, se pojavlja tako v pisnem, še pogosteje pa govorjenem jeziku, in kot ugotavljata avtorici, pogosto služi posebnemu pragmatičnemu namenu, po večini za razdvoumljanje in poudarjanje. Analiza razkrije, da je

tematika rabe osebnega zaimka kot osebka v slovenščini mnogo kompleksnejša, kot bi lahko sklepali iz dosedanjih opisov.

Eva Pori in **Iztok Kosem** se slovnične tematike dotikata prek analize kolokacij, in sicer tistega dela kolokacijskega gradiva, ki nastopa v besednozveznih strukturah s prislovi. Prispevek predstavlja rezultate jezikoslovne evalvacije strojno izluščenega kolokacijskega gradiva in odpira empirično podprto razpravo o tem, kje so meje slovnično relevantne kolokacije kot koncepta. To vprašanje je ključno za nadgradnjo avtomatskega luščenja kolokacijskih kandidatov za slovenščino, ki bo neposredno podprla razvoj Kolokacijskega slovarja sodobne slovenščine in podobnih jezikovnih virov.

Mojca Smolej se posveča dvema skupinama odvisniških stavkov, ki so bili v dosedanjih obravnavah pogosto prezrti: izhodiščno nestavčnočlenskim odvisnikom in členkovnim odvisnikom. Razprava, ki s pomočjo korpusnega gradiva opiše navedeni skupini odvisnikov, opozarja na dejstvo, da tudi vprašanja, ki so v preteklosti uživala veliko jezikoslovne pozornosti, niso nujno v celoti odgovorjena. Celovit izkorpusni pristop k pregledu odvisnikov lahko naslovi identificirane vrzeli, ki so pereče tudi za jezikovno didaktiko.

Jerica Snoj podaja slovaropisni pogled na besednovrstno kategorizacijo leksikalnega gradiva, ki je v idealnem stanju skladna s slovnično teorijo, saj sta slovar in slovnica v prizmi leksikogramatike različni področji enega samega kontinuuma. Avtorica ob primerih za povedkovnik, členek in izdeležniške tvorjenke prikaže razvoj besednovrstnega kategoriziranja v izbranih slovenskih slovarjih in na osnovi tega podaja dragocene smernice za besednovrstno obravnavo v prihodnjih slovenskih leksikalnih virih.

SLOVENŠČINA 2.0: “GRAMMAR IN LINGUISTIC DESCRIPTION”

To delo je ponujeno pod licenco Creative Commons: Priznanje avtorstva-
Deljenje pod enakimi pogoji 4.0 Mednarodna.

This work is licensed under the Creative Commons Attribution-ShareAlike 4.0
International.

<https://creativecommons.org/licenses/by-sa/4.0/>

