Baština Akademije nauka i umjetnosti Bosne i Hercegovine

Leksikologija i leksikografija I: Leksikografski postupak

Halilović, Senahid

2022-12-31

Akademija nauka i umjetnosti Bosne i Hercegovine

https://bastina.anubih.ba/items/6ca51dde-2a77-4899-84dc-e2515e45155a

Preuzeto s Baštine Akademije nauka i umjetnosti Bosne i Hercegovine

https://bastina.anubih.ba/

DOI: 10.5644/PI2022.206.03

RASTOČI SLOVAR SLOVENSKEGA KNJIŽNEGA JEZIKA (eSSKJ) – ORGANIZACIJA IN PRIKAZ JEZIKOVNIH PODATKOV*

Duša Divjak Race, Nataša Gliha Komac
ZRC SAZU, Inštitut za slovenski jezik Frana Ramovša
Novi trg 4, Ljubljana, Slovenija
dusa.race@zrc-sazu.si
natasa.gliha-komac@zrc-sazu.si

Apstrakt

V prispevku je predstavljena tretja izdaja *Slovarja slovenskega knjižnega jezika* (eSSKJ), tj. spletni rastoči slovar sodobnega slovenskega knjižnega jezika, ki od leta 2016 nastaja na Inštitutu za slovenski jezik Frana Ramovša ZRC SAZU. Slovar se letno nadgrajuje na inštitutskem slovarskem portalu Fran. Od septembra 2021 je na šolskem portalu Franček dostopna tudi njegova prilagojena šolska izdaja.

eSSKJ poleg semantičnih, osnovnih formalnih (zapis, izgovor) in slovničnih lastnosti (tradicionalno navedenih v zaglavju, npr. besednovrstna umestitev) ter pragmatičnih posebnosti eno- in večbesedne leksike vsebuje še tipične kolokacije, razvrščene po skladenjskih vzorcih, celotne oblikoslovne paradigme, ki so jakostno in tonemsko naglašene, frazeološke enote in njihove variante, etimološke osvetlitve, po potrebi pa tudi normativna pojasnila in besedotvorne podatke. Označenost podatkov v slovarski bazi omogoča različne načine prikaza slovarske vsebine za različne uporabnike in uporabniške potrebe.

Zaradi možnosti, ki jih ponuja spletni medij, eSSKJ nadgrajuje klasičen način prikazovanja jezikovnih podatkov in dostopa do njih, hkrati pa vsebino povezuje z zunanjimi viri, npr. s *Slovarjem slovenskega znakovnega jezika*. Nekatere slovarske vsebine so uporabnikom na voljo tudi preko slušnega prenosnika, saj sta dodana zvočni posnetek izgovora iztočnice in stranske oblike ter sintetizirano branje celotnega slovarskega sestavka.

Vsebina eSSKJ je prikazana na dveh ravneh, ki ju narekuje portalska zasnova. Prva raven je seznam zadetkov, ki ga uporabnik dobi, ko v iskalnik vpiše želeno besedo. Na njem so iz eSSKJ prikazani osnovni podatki o iskani besedi (npr. formalne lastnosti besede, pomenska členitev, stalne besedne zveze in frazeološke enote), ki ob podatkih iz ostalih slovarjev (pravopisnih, zgodovinskih, narečnih ipd.) uporabniku omogočajo celosten, zaokrožen pogled na besedo z različnih vidikov. Druga raven je prikaz slovarskega sestavka, ki ga uporabnik dobi,

* Prispevek je nastal v okviru raziskovalnega programa *Slovenski jezik v sinhronem in diahronem razvoju* (P6-0038), ki ga financira Javna agencija za raziskovalno dejavnost Republike Slovenije.

ko na seznamu zadetkov izbere želeni slovar. V primeru eSSKJ to pomeni celosten opis rabe posameznega leksema v sodobnem slovenskem knjižnem jeziku.
Poleg uporabnosti je cilj slovarja razumljivost. Slednji eSSKJ sledi z odpravo krajšav, značilnih za tiskane slovarje, naravnanostjo k eksplicitnemu navajanju podatkov, dodana so pojasnila v zaslonskih namigih, uvedeno je slovarsko metabesedilo za označevanje razdelkov idr. Novosti v interpretaciji slovarskih vsebin lahko spremljamo tudi v šolski izdaji eSSKJ na portalu Franček, kjer je še več pomagal in usmeritev, kot so npr. eksplicitna poimenovanja mikrostrukturnih elementov (npr. razlaga, zgled, različica), kratke razlage slovarskih oznak in slovničnih kategorij v zaslonskih namigih, hitre povezave na opise v slovnici, preimenovanje težje razumljivih terminov in oznak (npr. *kolokacije* v *tipične zveze*, oznaka *tudi*, ki uvaja neprednostno dvojnico, v *manj običajno*). Tako je šolajočemu se uporabniku omogočeno lažje razumevanje slovarskih vsebin, hkrati pa se seznanja s slovarsko strukturo nasploh in se uvaja v delo z različnimi slovarji.

Ključne besede: spletni rastoči slovar, slovenščina, prikaz jezikovnih podatkov, slovarski uporabnik

1. Uvod

Tretja izdaja *Slovarja slovenskega knjižnega jezika* (eSSKJ) je temeljni splošni razlagalni informativno-normativni slovar, ki od leta 2016 nastaja na Inštitutu za slovenski jezik Frana Ramovša ZRC SAZU. Nastaja na sodobnem pisnem jezikovnem gradivu, zlasti korpusu Gigafida, upošteva pa se tudi stanje v drugih korpusih, npr. Gigafida 2.0, slWac, KAS, in virih. Sledi smernicam, predstavljenim v *Konceptu novega razlagalnega slovarja slovenskega knjižnega jezika* (*Koncept*; Gliha Komac idr. 2015),[1] ki se jih glede na spoznanja sodobnih teorij in izkušnje pri delu z gradivom smiselno dopolnjuje. Je spletni rastoči slovar, ki se letno nadgrajuje na inštitutskem slovarskem portalu *Fran*. Z manjšim časovnim zamikom izhaja tudi e-knjiga prirastka, ki omogoča vpogled v proces nastajanja oz. rast slovarja ter spremljanje sprememb in novosti. Od septembra 2021 je na šolskem portalu *Franček*[2] dostopna šolska izdaja, prilagojena za uporabo v zadnjem triletju osnovnih šol in srednjih šolah.

2. Vsebina slovarja

eSSKJ opisuje jezikovno stanje in pojmovni svet sodobnega slovenskega knjižnega jezika, kot ju izkazuje jezikovna raba, potrjena predvsem v pisnih besedilih. Kot splošni razlagalni slovar je namenjen širokemu krogu slovarskih

[1] < https://fran.si/179/novi-slovar-slovenskega-knjiznega-jezika/datoteke/Potrjeni_koncept_NoviSSKJ.pdf >
[2] < https://www.franček.si/>

uporabnikov, t. i. splošnim in specializiranim jezikovnim uporabnikom (Gliha Komac 2018: 7–13), ki slovar uporabljajo pri tvorjenju besedil, njihovem sprejemanju, z raziskovalnimi nameni idr. Pri rabi slovarja je uporabnikom na voljo spremno besedilo *Kako brati slovar*,[3] ki prinaša kratek opis strukture slovarja in vodila za interpretacijo slovarskih podatkov. Slovarsko celoto zaokrožata *Predgovor* in *Kolofon* s povezavo do repozitorija Clarin.si,[4] kjer sta dostopni slovarski bazi eSSKJ in e-naslov za stik med slovarskimi uporabniki in redaktorji.

eSSKJ poleg semantičnih, formalnih (zapis, izgovor) in slovničnih lastnosti (tradicionalno navedenih v zaglavju, npr. besednovrstna umestitev) ter stilističnih in pragmatičnih posebnosti eno- in večbesedne leksike vsebuje še tipične kolokacije, razvrščene po skladenjskih vzorcih, celotne oblikoslovne paradigme, ki so jakostno in tonemsko naglašene, frazeološke enote in njihove variante, etimološke osvetlitve, po potrebi pa tudi normativna pojasnila in besedotvorne podatke.

3. Organizacija in prikaz slovarskih podatkov

eSSKJ izhaja iz tradicije tiskanih predhodnikov, *Slovarja slovenskega knjižnega jezika* (1970–1991; SSKJ) in druge, dopolnjene in deloma prenovljene izdaje *Slovarja slovenskega knjižnega jezika* (SSKJ^2),[5] ki sta pomembno vplivala na navade in pričakovanja slovarskih uporabnikov. Ti so na podlagi izkušenj oblikovali določene prakse in strategije slovarskih rab oz. poizvedb, ki v najboljšem primeru pripeljejo do razrešitve jezikovne dileme.

SSKJ in SSKJ^2 sta kot tiskana slovarja digitalno različico dobila s prenosom vsebine na splet, zaradi česar uporabnik do njunih vsebin dostopa lažje (na voljo mora imeti le internetno povezavo in ustrezno napravo) in hitreje (z zapisom besede v iskalnik). Pri iskanju po obeh slovarjih je omogočeno napredno iskanje, kot povezave so v spletni različici SSKJ^2 urejene vodilke, kazalke ter protipomenke in nekatere sopomenke, ki vodijo do ustreznih

[3] <https://fran.si/201/esskj-slovar-slovenskega-knjiznega-jezika/datoteke/eSSKJ_Kako_brati_slovar.pdf >

[4] < https://www.clarin.si/repository/xmlui/>

[5] SSKJ je na spletu dostopen od leta 2000, najprej na portalu *Slovarske in besedilne zbirke*, od 2014 pa na portalu Fran. SSKJ^2 je v tiskani in spletni izdaji izšel leta 2014.

sestavkov.[6] Izboljšala sta se torej dostop do podatkov in deloma navigacija, skoraj nespremenjen pa je ostal prikaz podatkov.

Pri prenosu tiskanih slovarjev v digitalno obliko ter ustvarjanju in oblikovanju izhodiščno digitalnih slovarjev je ob zavedanju prednosti, ki jih ponuja digitalni medij, naraslo zanimanje za slovarskega uporabnika, njegove potrebe oz. načine slovarske uporabe (Lew in de Schryver 2014; Müller-Spitzer 2014; Tarp 2014; Dziemianko 2015; Bergenholtz in Johnsen 2017). Raziskav o uporabniških potrebah in praksah pri uporabi (spletnih) slovarjev, jezikovnih zbirk in portalov za slovenski jezik ni prav veliko (npr. Rozman 2004, 2010; Gliha Komac 2017), pogosto so osredotočene na dvo- ali tujejezične slovarje (npr. Vrbinc 2005, 2020; Farina idr. 2019), prinašajo pa smernice za nadaljnje raziskave (npr. Vrbinc 2005, 2020; Arhar Holdt 2018; Arhar Holdt idr. 2016) ter vrsto poudarkov – zlasti spoznanja o zapletenosti slovarjev, organizaciji podatkov v njih in preferencah v zvezi z ilustrativnim gradivom –, ki jih je pri oblikovanju slovarjev smiselno upoštevati.

Kot opišejo Klemenc idr. (2015), pri nastajanju sodobnih digitalnih slovarjev predpostavljamo njihovo večnivojskost in govorimo o podatkovnem, predstavitvenem in aplikacijskem nivoju slovarja. Podatkovni nivo predstavljajo metajezikovni in jezikovni podatki, kot so zajeti v slovarski bazi, predstavitveni nivo pa njihov prikaz za uporabnika, najpogosteje v obliki slovarskega sestavka, bodisi na spletu bodisi v (e-)knjigi. Med podatkovnim in predstavitvenim nivojem je aplikacijski nivo, ki deluje kot pretvornik uporabniških poizvedb in podatkov v obliko, ki jo potrebuje bodisi podatkovni bodisi predstavitveni nivo (Klemenc idr. 2015: 15–16). Večnivojskost z vidika koncepta prostora, ki je bil v dobi tiskanih slovarjev poudarjen zaradi potrebe po varčevanju z njim, obravnava tudi Lew (2011), ki razlikuje med shranjevalnim prostorom (ang. *storage place*) in predstavitvenim prostorom (ang. *presentation space*). Glavna razlika med shranjevalnim prostorom v tiskanih in digitalnih slovarjih je njegova relativna neomejenost pri digitalnih slovarjih, pri katerih je manj omejen tudi predstavitveni prostor, saj ni več statičen. Lew (2011: 4) navaja še percepcijski prostor (ang. *perceptual space*), ki povezuje slovar z uporabnikom, saj se nanaša na uporabnikovo zmožnost sprejemanja in obdelovanja slovarskih podatkov. Večnivojskost slovarja kaže na interdisciplinarnost leksikografskega procesa, pri katerem sta poleg vsebine ključna tudi tehnološka izvedba in prikaz slovarskih vsebin, kar v ospredje postavlja slovarskega uporabnika, njegove prakse in potrebe.

[6] Več o okoliščinah nastanka SSKJ^2 in vsebinskih spremembah ter novostih gl. Perdih in Snoj 2015.

3. 1. Slovarska baza in leksikografska orodja

Kot je zapisano v *Konceptu* (Gliha Komac idr. 2015), je slovarska baza za eSSKJ v tehničnem smislu sestavljena iz več podatkovnih baz, zasnovana in strukturirana pa tako, da je uporabna za jezikoslovne in jezikovnotehnološke namene ter računalniško obdelavo. Jedrna dela sta t. i. lematska baza, ki vsebuje podatke o izrazni podobi slovarskih iztočnic (zapis, izgovor, pregibanje idr.), in baza s pomenskimi, skladenjskimi, besedotvornimi stilističnimi idr. slovarskimi podatki. Slovarska baza zajema več podatkov, kot jih je prikazanih; tako v slovarskem sestavku eSSKJ na *Franu* npr. niso prikazani manj frekventne frazeološke variante, delni sinonimi, dodatna redaktorjeva opažanja, področje idr. Označenost podatkov v slovarskih bazah omogoča različne načine prikaza slovarske vsebine glede na različne uporabniške potrebe, kar smo izkoristili v šolski izdaji eSSKJ na portalu *Franček* (več v poglavju 3. 3.).

Za urejanje slovarske baze eSSKJ se uporablja orodje iLEX (gl. Erlandsen 2010), za analizo gradiva pa SketchEngine,[7] v katerem so generirane skice iz korpusa Gigafida. V dodatno pomoč pri interpretaciji gradiva so še korpusi Gigafida 2.0, metaFida, slWac, KAS idr. Slovaropisna orodja, ki so redaktorjem trenutno na voljo, glede na refleksijo leksikografskega dela (Vrbinc, Farina, in Vrbinc 2021) zahtevam / potrebam redakcijskega dela ne ustrezajo povsem. Zlasti za potrebe zvrstne in stilne analize rabe bi potrebovali zvrstno in stilno uravnotežen korpus (Gliha Komac, Divjak Race 2022) ter generirane skice iz sodobnejših korpusov.

3.2. Vpetost slovarja v portalsko zasnovo in prikaz slovarskih podatkov

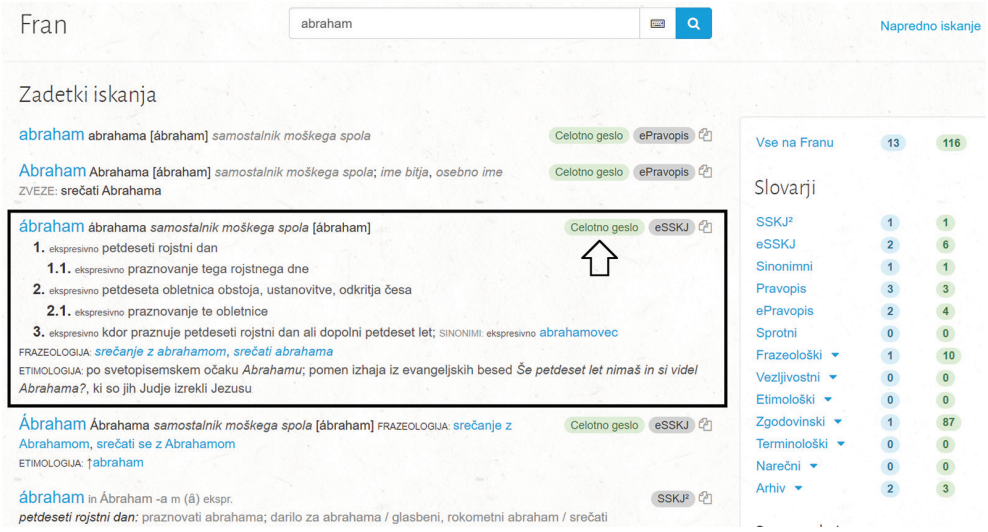
eSSKJ je izhodiščno zasnovan za rabo v spletnem okolju, in sicer na portalu *Fran*,[8] na katerem so dostopni slovarji Inštituta za slovenski jezik Frana Ramovša ZRC SAZU, trenutno 11 splošnih slovarjev (slovarji knjižnega jezika, pravopisna slovarja, *Sinonimni slovar slovenskega jezika* idr.), 2 etimološka slovarja, 6 zgodovinskih, 16 terminoloških in 5 narečnih slovarjev ter *Slovenski lingvistični atlas*, Jezikovna in Terminološka svetovalnica, razne (gradivske) zbirke, npr. *Slovenske slovnice in pravopisi*, inštitutske monografije, *Pleteršnikova zbirka zemljepisnih imen* ipd., korpusi, npr. *Korpus besedil slovenskih protestantskih piscev 16. stoletja*, ter drugo, npr. *Pravopis 8.0: Pravila novega slovenskega pravopisa za javno razpravo*.

[7] <https://www.sketchengine.eu/>
[8] <www.fran.si>

eSSKJ je slovarsko celovita, zaokrožena enota, hkrati pa del portala, kar pomeni, da imajo uporabniki ob uporabi tega slovarja hkrati na voljo podatke iz drugih slovarjev na portalu. Uporabnik tako enostavno in hitro dobi informacije o besedi z različnih vidikov, mogoča pa je primerjava med slovarji. Glede na trenutno zasnovo portala, ki združuje vsebinsko raznovrstne slovarje, je pomembno, da se uporabniki zavedajo razlik med slovarji, sicer lahko pride do neustreznih interpretacij.

Do slovarskih vsebin eSSKJ uporabnik dostopa bodisi preko osnovnega iskanja na portalu bodisi s pomočjo naprednega iskanja, ki omogoča iskanje zelo specifičnih vsebin. Druga možnost je posebej dobrodošla za specializirane uporabnike, npr. prevajalce, jezikoslovce, literate, učitelje ipd., ki se z jezikom profesionalno ukvarjajo.

Portal prinaša slovarske vsebine na dveh ravneh, in sicer na ravni zadetkov iskanja in ravni celotnega gesla. Po vnosu iskane besede v iskalnik rezultat iskanja tvorijo zadetki iz vseh slovarjev, ki obravnavano besedo opisujejo. Npr. če vpišemo besedo *abraham*, nam portal ponudi 13 zadetkov, in sicer iz eSSKJ, *ePravopisa*, *Sinonimnega slovarja slovenskega jezika*, *Slovarja slovenskih frazemov*, *Slovarja stare knjižne prekmurščine* idr., ne pa iz terminoloških slovarjev, saj beseda v njih ne nastopa kot iztočnica. Iz vsakega slovarja so na seznamu zadetkov prikazani le nekateri podatki (npr. iztočnica z najosnovnejšimi podatki oz. podatki, ki so v določenem slovarju osrednji). Celoten slovarski sestavek se odpre s klikom na gumb »Celotno geslo«.



Slika 1: Zadetki iskanja za besedo abraham *na portalu Fran s poudarjenimi podatki iz eSSKJ*

Na ravni seznama zadetkov so iz eSSKJ prikazane osnovne formalne lastnosti besede, pomenska členitev, besedne zveze in frazeološke enote, v katerih beseda nastopa, ter etimološka osvetlitev. Na ravni celotnega gesla oz. slovarskega sestavka je podatkov več, saj prinaša celosten opis rabe posamezne besede in njenih stalnih besednih zvez v slovenskem knjižnem jeziku z avtentičnimi zgledi rabe, skladenjski vzorci, kolokacijami, besedotvornimi, oblikoslovnimi, stilnimi podatki, normativnimi opažanji ipd.

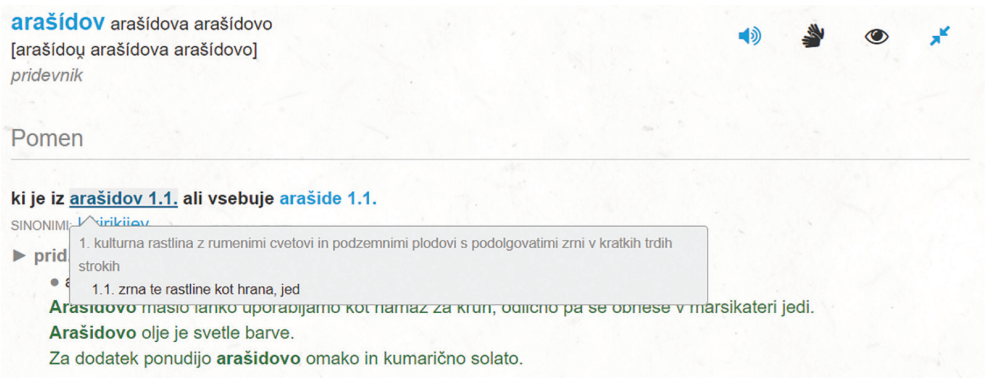
3.2.1. *Slovarski sestavek* eSSKJ

Slovarski sestavki so na portalu organizirani in prikazani tako, da bi bili čim bolj uporabni in razumljivi ciljnemu uporabniku (Gliha Komac idr. 2015). Zaradi možnosti, ki jih ponuja digitalni medij, slovar do določene mere nadgrajuje klasičen način prikazovanja slovarskih podatkov in dostopa do njih. Pri procesu slovarske vizualizacije[9] je prisoten razmislek o razmerju med tradicionalnim prikazom, tj. prikazom slovarskega sestavka, kakršen se pojavlja v tiskanih slovarjih, in sodobnim prikazom, ki se opira na možnosti, ki jih ponuja digitalni medij, ter o značilnostih in pričakovanjih spletnega uporabnika.

Pri eSSKJ so izrabljene lastnosti obeh načinov. Dobršen del slovarskih uporabnikov je še vedno vajen tradicionalne postavitve slovarskega sestavka, zato je pri eSSKJ ohranjena tipična slovarska struktura, ki se začne z onaglašeno iztočnico z zaglavjem, nadaljuje s pomenskim delom itd. K razumevanju in branju take strukture sestavka se učence in dijake navaja tudi pri rednem šolskem pouku. Hkrati pa so izrabljene možnosti digitalnega medija, npr. bogatenje vsebine z avdio-vizualnimi elementi, povezovanje z drugimi slovarji in prehajanje med posameznimi razdelki sestavka s pomočjo izbirnega menija. Slovarske vsebine so tako uporabnikom na voljo tudi preko slušnega prenosnika, saj sta ob zaglavju dodani ikoni za aktivacijo zvočnega posnetka izgovora osnovne in prve stranske oblike iztočnice ter sintetiziranega branja celotnega slovarskega sestavka (zaenkrat poskusno pri iztočnici *kruhovec*). Dodana je povezava do *Slovarja slovenskega znakovnega jezika*, v katerem si uporabnik lahko ogleda video s kretanjem besede, ki jo poišče v slovarju.

[9] Kot slovarsko vizualizacijo razumemo odločitve v zvezi s prenosom podatkov do slovarskega uporabnika s prikazom slovarskih vsebin na spletu, ki zajemajo izbiro načina prikaza podatkov (torej v besedilni / grafični obliki, v obliki interaktivnih shem, pojmovnikov ipd.), osnovne postavitve slovarja, strukturo slovarskega sestavka (mikrostruktura), stopnjo eksplicitnosti / implicitnosti informacij ter možnosti uporabnikovega dostopa do podatkov, načina iskanja in navigiranja po slovarju.

Opisovanje razlik med tiskanimi in digitalnimi slovarji (Dziemianko 2010; Lew 2011, 2013; Lew in de Schryver 2014) je izpostavilo nekaj slabosti tiskanih slovarjev, ki se prenašajo na splet. Oppentocht in Schutz (2004: 215) mdr. navajata šifriran način posredovanja informacij, navzkrižno sklicevanje in strukturiranje sestavka, osredinjenega na iztočnico. Slednje eSSKJ ohranja, ukinjena pa so kazalčna gesla, saj so tudi neprednostne dvojnice (torej tiste, ki so v knjižnem jeziku manj ustrezne) opisane z lastnim slovarskim sestavkom in ne samo z usmeritvijo na opis prednostne dvojnice (uporabnika se na ravni zaglavja še vedno usmerja k ustreznejši dvojnici, npr. *barrique* glej *barík*). V eSSKJ so sklici (kot povezave do drugih sestavkov znotraj slovarja) ohranjeni pri pridevnikih, ki se v razlagi sklicujejo na samostalniško iztočnico, npr. *arašidov* 'ki je iz arašidov ali vsebuje arašide'. Če se uporabnik z miško premakne na moder sklic (gl. podčrtano na sliki 2), se mu v zaslonskem namigu izpiše tisti pomen samostalnika, na katerega se sklicuje razlaga pri pridevniku (v našem primeru gre za pomen 'zrna te rastline kot hrana, jed'). Klik na sklic uporabnika popelje v slovarski sestavek *arašid*. Na tak način se ohranjajo kratke razlage pridevnikov, ali to vpliva na razumljivost razlag, pa bo treba še preveriti.



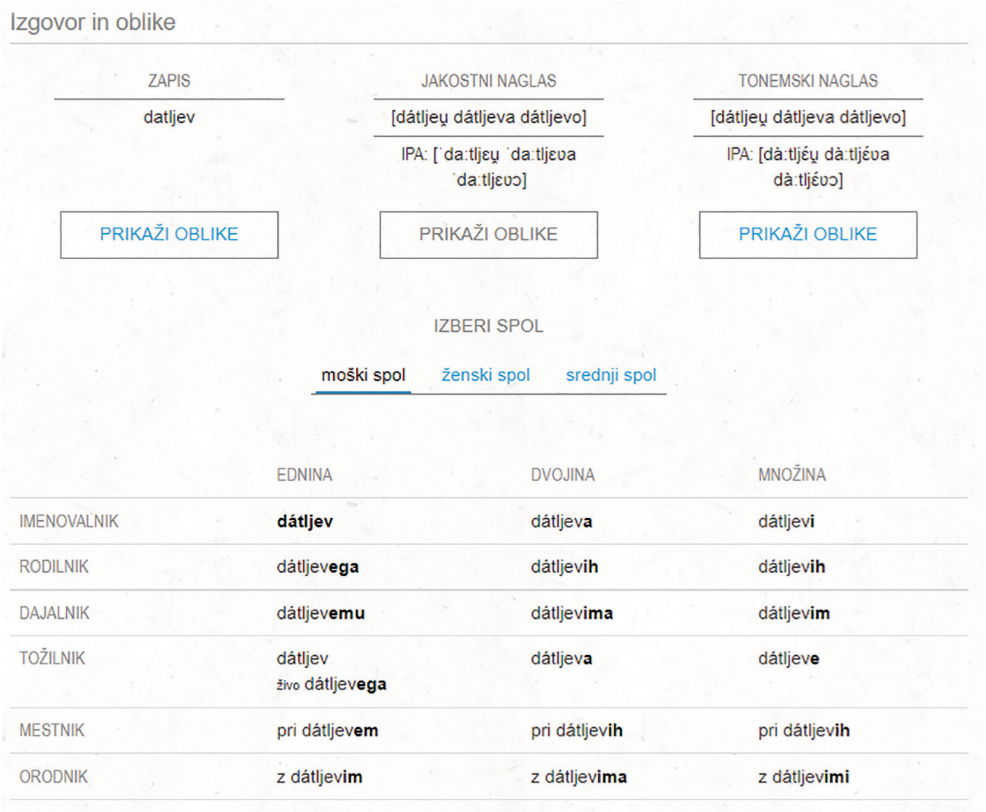
Slika 2: Slovarski sestavek besede arašidov *v eSSKJ*

Da bi bila slovarska vsebina bolj razumljiva in eksplicitna, eSSKJ glede na klasičen prikaz odpravlja določene krajšave, npr. pri kvalifikatorjih (ekspr. > ekspresivno, knj. pog. > knjižno pogovorno) in slovničnih oznakah (m. > samostalnik moškega spola), usmerjen je v eksplicitno navajanje podatkov, dodana so pojasnila v zaslonskih namigih (npr. Izgovor, samostalniška beseda + samostalniška beseda v rodilniku, Stavčni zgled), poimenovani razdelki (npr. Frazeologija – namesto ●, ki v SSKJ$^{(2)}$ uvaja frazeološko gnezdo), različni

tipi podatkov so členjeni v samostojne vrstice idr. Tako naj bi vsebine približali tudi tistim, ki so s slovarskimi principi in konvenciji seznanjeni v manjši meri.

Ker eSSKJ prinaša širok nabor jezikovnih podatkov, se pri prikazu celotnega gesla, zlasti pri leksemih z več pomeni in stalnimi zvezami ter pri prikazu pregibnostno-naglasnih vzorcev odpira vprašanje podatkovne preobremenjenosti.[10] Zaradi želje po prikazu čim več podatkov se namreč lahko zgodi, da slovarski sestavek postane nepregleden. Fuertes-Olivera in Tarp (2014: 64) med smernicami za sodobne slovarje opozarjata, da mora slovarski sestavek vsebovati najmanjšo količino podatkov, tj. točno toliko podatkov, kolikor jih uporabnik potrebuje v okviru poizvedbe. To težavo za pomenski razdelek in opis stalnih besednih zvez in frazeoloških enot deloma razrešuje možnost t. i. skrčenega pogleda, ki zmanjša količino ponazarjalnega gradiva, število hkrati prikazanih pregibnostno-naglasnih paradigem pa je zmanjšano tako, da je izbira, katera paradigma se prikaže, prepuščena uporabniku. Ta na prvi ravni izbira med neonaglašeno, jakostno naglašeno ali tonemsko naglašeno pregibnostno paradigmo, na drugi ravni pa so izbire odvisne od posamezne besedne vrste (pri pridevnikih izbira paradigme glede na slovnični spol; pri glagolu lahko pri deležniku izbira prikaz deležnika na *-l* ali na *-č* ipd.). Na primeru pridevnika *datljev* je npr. na tak način število prikazanih tabel zmanjšano z devet na eno, kar izboljša preglednost sestavka.

[10] Na pojav podatkovne preobremenjenosti (ang. *information overload*) pri digitalnih slovarjih opozori Tarp (2014), ki kot eno od možnih rešitev predlaga prilagojen prikaz za različne uporabnike, kar smo uresničili s šolsko izdajo eSSKJ na portalu Franček.



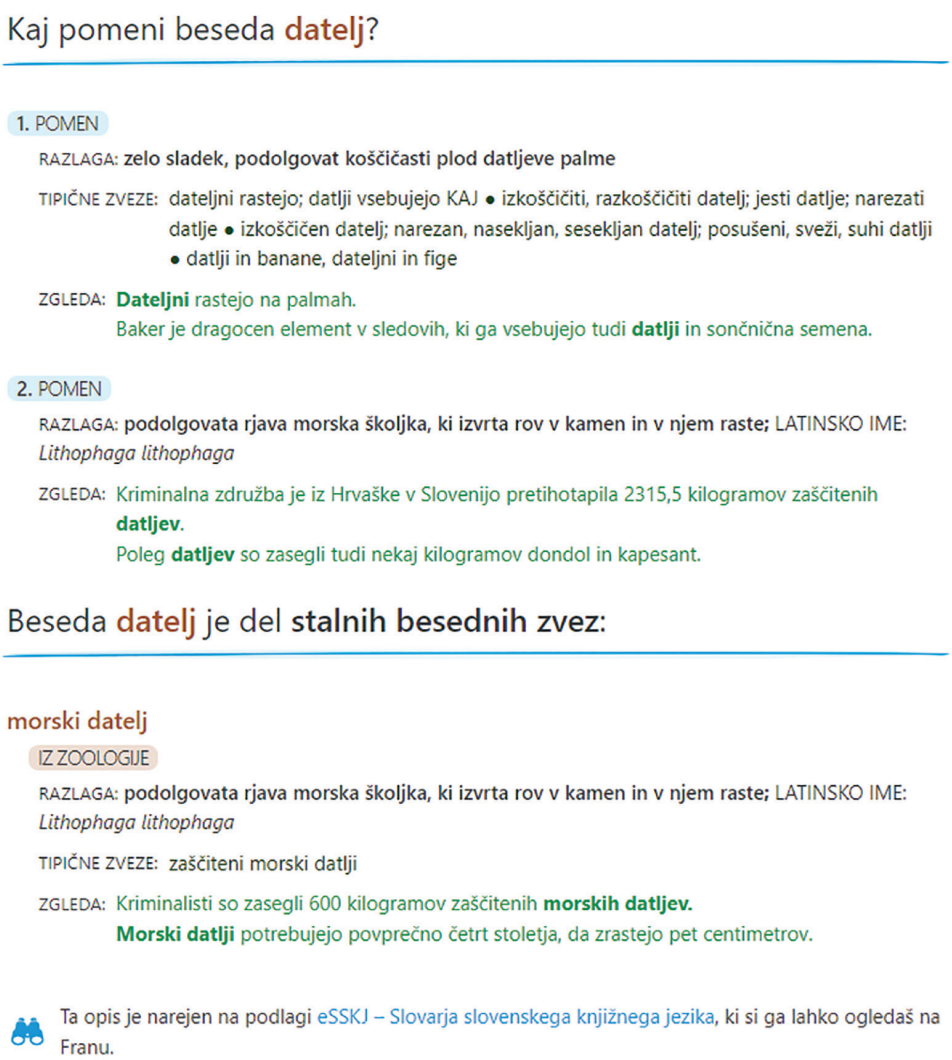
Slika 3: Razdelek Izgovor in oblike pri pridevniku datljev

3.3. Prilagojen prikaz slovarskih podatkov za šolsko rabo

Šolska izdaja eSSKJ je na voljo na portalu Franček, tj. na interaktivnem šolskem slovarsko-slovničnem portalu za slovenščino, ki je namenjen učencem treh starostnih skupin (1.–5. in 6.–9. razred osnovne šole ter srednja šola).[11] Podatki iz eSSKJ so vključeni v različne razdelke na portalu, in sicer v pomenski, frazeološki razdelek in razdelek z oblikami, namenjeni pa so učencem 6.–9. razreda in srednješolcem. Pri prilagajanju prikaza slovarskih podatkov različnim starostnim skupinam so bile upoštevane otrokove zmožnosti in potrebe v skladu z učnim načrtom (Perdih idr. 2021).

[11] Portal Franček <www.franček.si> je zasnovan drugače kot Fran <www.fran.si>, saj podatkov ne prikazuje glede na posamezne slovarje, temveč v različnih tematskih razdelkih (pomen, sopomenke, pregibanje, izgovor, frazemi, narečja, izvor in zgodovina), v katerih so lahko prikazani podatki iz različnih virov (Perdih idr. 2021).

Deli slovarskega sestavka so jasno označeni. Pomene uvajajo modri oblački z napisi *1. pomen*, *1. podpomen* za prvo starostno skupino oz. **številke** za srednješolce. Različni tipi podatkov so členjeni v samostojne vrstice, kar poleg pomenskih oblačkov in podatkov v obliki visečih odstavkov olajša orientacijo po sestavku. Različne tipografske možnosti (ležeči tisk za latinska imena, krepki tisk izpostavitev iztočnice v zgledih ipd.) in barve so uporabljene za poudarjanje osrednje informacije. Struktura sestavka je manj zgoščena, sicer pa sledi uveljavljeni formi slovarskega sestavka.



Slika 4: Beseda datelj *v pomenskem razdelku za učence 6.–9. razreda na portalu Franček*

Prikaz podatkov upošteva, da imajo uporabniki 6.–9. razreda in srednješolci različni stopnji jezikovnega in metajezikovnega znanja ter različne vsebinske potrebe, zato se tudi prikaz za ti dve skupini uporabnikov razlikuje. Prikaz za učence 6–9. razreda ima dele slovarskega sestavka eksplicitno poimenovane (razlaga, tipične zveze, zgledi ipd.), v oblikoslovnem razdelku so v pomoč risane ikone za slovnične kategorije in omejitve (npr. edninski samostalnik je označen s sličico prečrtanih kock, na katerih so dve in tri pike) ter vprašalnice za sklone v zaslonskih namigih. V primerjavi z eSSKJ na Franu je na Frančku vsebina omejena, manjše je število kolokacij, terminoloških in frazeoloških zvez ter njihovih variant. Prikaz za srednješolce predpostavlja, da so uporabniki seznanjeni s slovarsko strukturo, zato so s prikaza umaknjena poimenovanja razdelkov (ta so sicer še vidna v zaslonskih namigih) in pomagala v obliki ikon. Obseg stalnih besednih zvez je večji, prav tako frazeoloških variant.

Za obe stopnji so v prikazu v zaslonskih namigih kratke razlage slovarskih oznak, kvalifikatorjev in slovničnih kategorij ter hitre povezave na obširnejše opise v slovnici.[12] Preimenovani so bili tudi težje razumljivi termini in oznake (npr. *kolokacije* so preimenovane v *tipične zveze*, oznaka *tudi*, ki uvaja neprednostno dvojnico, je preimenovana v *manj običajno*). Tako so ne samo povečane možnosti šolskega uporabnika za razumevanje slovarskih vsebin, do katerih dostopa na *Frančku*, temveč se bodoči uporabniki eSSKJ s seznanjanjem s slovarsko strukturo hkrati uvajajo v delo z različnimi slovarji.

4. Zaključek

eSSKJ se kot temeljni splošni razlagalni slovar uvršča med tiste jezikovne priročnike, po katerih uporabniki slovenskega jezika najpogosteje posegajo pri reševanju jezikovnih vprašanj in težav ter so med najbolj ključnimi viri za razvoj jezika (Gliha Komac 2017: 77−78). Ker prinaša širok nabor raznovrstnih jezikovnih informacij (semantične, fonetične, besedotvorne, oblikoslovne ipd.), služi različnim uporabniškim potrebam in različnim skupinam uporabnikov, ki imajo glede slovarskega orodja različna pričakovanja in preference. Pomemben izziv prihodnosti, ki bo eSSKJ naredil bolj uporaben in ga bolj približal različnim skupinam uporabnikov, je zlasti spremljanje

[12] Za več o prikazu in povezovanju slovničnih podatkov na šolskem jezikovnem portalu Franček gl. Ledinek, Ahačič, Petric Žižić 2022: 152−179.

slovarskih potreb in raziskovanje praks slovarskih uporabnikov ter udejanjanje njihovih pričakovanj in potreb v slovaropisni praksi.

Literatura

Arhar Holdt, Špela (2018), "Odnos jezikovnih uporabnikov do splošnih enojezičnih slovarjev. Slovenski vidik", *Slovenščina 2.0: empirical, applied and interdisciplinary research* 6/1, 1–36, doi: 10.4312/slo2.0.2018.1.1-36.

Arhar Holdt, Špela, Jaka Čibej, Ana Zwitter Vitez, Polona Gantar, Vojko Gorjanc (2016), "Identifying Dictionary Needs of Language Users by Analysing User-Generated Content in Digital Media", u: *The European Dictionary portal: challenges and chances of facing heterogeneity / COST ENeL WG4 meeting Barcelona, Spain, 31 March - 1 April 2016*, 1–6, Barcelona

Bergenholtz, Henning, Mia Johnsen (2017), "Log Files as a Tool for Improving Internet Dictionaries", *HERMES – Journal of Language and Communication in Business* 18, 117–142, doi: 10.7146/hjlcb.v18i34.25802.

Dziemianko, Anna (2015), "Colours in Online Dictionaries. A Case of Functional Labels", *International Journal of Lexicography* 28, 27–61, doi: 10.1093/ijl/ecu028.

Erlandsen, Jens (2010), "iLEX, a general system for traditional dictionaries on paper and adaptive electronic lexical resources", u: A. Dykstra, T. Schoonheim, ur., *Proceedings of the XIV EURALEX International Congress. 6–10 July 2010*, 306, Fryske Akademy – Afûk, Leeuwarden / Ljouwert

Farina, Donna M. T. Cr., Marjeta Vrbinc, Alenka Vrbinc (2019), "The User is King. Advice to Lexicographers of Learner's Dictionaries", u: M. Gürlek, A. N. Çiçekler, Y. Taşdemir, ur., *Proceedings of the 13th International Conference of the Asian Association for Lexicography*, 240–248, Asos, Istanbul

Gliha Komac, Nataša, Nataša Jakop, Janoš Ježovnik, Simona Klemenčič, Domen Krvina, Nina Ledinek, Tanja Mirtič, Andrej Perdih, Špela Petric, Marko Snoj, Andreja Žele (2015), *Koncept novega razlagalnega slovarja slovenskega knjižnega jezika*, Založba ZRC, Ljubljana

Gliha Komac, Nataša, ur. (2017), *Ciljni raziskovalni projekt Jezikovna politika Republike Slovenije in potrebe uporabnikov*. *Raziskovalno poročilo*, ZRC SAZU idr., Ljubljana

Gliha Komac, Nataša (2018), "Jezikovna politika Republike Slovenije in potrebe uporabnikov", *Slavia Centralis* 11/2, 7–15.

Gliha Komac, Nataša, Duša Race (2022), "Tipologija slovarskih oznak za tretjo izdajo Slovarja slovenskega knjižnega jezika (eSSKJ)", *Slavistična revija*

Klemenc, Bojan, Marko Robnik-Šikonja, Luka Fürst, Ciril Bohak, Simon Krek (2015), "Tehnološka izvedba sodobnega digitalnega slovarja", u: V. Gorjanc, P. Gantar, I. Kosem, S. Krek, ur., *Slovar sodobne slovenščine: Problemi in rešitve*, 52–63, Znanstvena založba Filozofske fakultete, Ljubljana

Ledinek, Nina, Kozma Ahačič, Špela Petric Žižić (2022), "Presenting and linking grammatical data on the Franček educational language portal", *Slovenski jezik* / *Slovene Linguistic Studies* 14, 152–179.

Lew, Robert (2011), "Space restrictions in paper and electronic dictionaries and their implications for the design of production dictionaries", *Issues in Modern Lexicography*

Lew, Robert, Gilles-Maurice de Schryver (2014), "Dictionary Users in the Digital Revolution", *International Journal of Lexicography* 27/4, 341–359, doi: 10.1093/ijl/ecu011.

Müller-Spitzer, Carolin, ur. (2014), *Using Online Dictionaries*, De Gruyter

Oppentocht, Lineke, Rik Schutz (2003), "Developments in electronic dictionary design", u: P. van Sterkenburg, ur., *A Practical Guide to Lexicography*, 215–227, Institute for Dutch Lexicology, Leiden

Pastor, Veronica, Amparo Alcina (2010), "Search Techniques in Electronic Dictionaries: A Classification for Translators", *International Journal of Lexicography* 23/3, 307–354, doi: 10.1093/ijl/ecq015.

Perdih, Andrej, Kozma Ahačič, Janoš Ježovnik, Duša Race (2021), "Building an Educational Language Portal Using Existing Dictionary Data", *Journal of Linguistics / Jazykovedný Casopis* 72/2, 568–78, doi: 10.2478/jazcas-2021-0052.

Perdih, Andrej, Marko Snoj (2015), "SSKJ2", *Slavia Centralis* 1, 5–15.

Rozman, Tadeja (2004), "Upoštevanje ciljnih uporabnikov pri izdelavi enojezičnega slovarja za tujce", *Jezik in slovstvo* 49/3–4, 63–75.

Rozman, Tadeja (2010), *Vloga enojezičnega razlagalnega slovarja slovenščine pri razvoju jezikovne zmožnosti*, doktorska disertacija, Ljubljana

Tarp, Sven (2014), "Detecting user needs for new online dictionary projects. Business as usual, user research or …?", u: C. Tiberius in C. Müller-Spitzer, ur., *Research into dictionary use: Wörterbuchbenutzungsforschung. 5. Arbeitsbericht des wissenschaftlichen Netzwerks* "*Internetlexikografie*", 16–26, Institut für Deutsche Sprache, Mannheim, OPAL – Online publizierte Arbeiten zur Linguistik

Vrbinc, Marjeta (2005), "An Empirical Study of Dictionary Use. The Case of Slovenia", *ELOPE: English Language Overseas Perspectives and Enquiries* 2/97, 97–106, doi: 10.4312/elope.2.1-2.

Vrbinc, Marjeta, Alenka Vrbinc (2020), "User Guides in Monolingual Slovenian Dictionaries. Users' perspective", *Lexicography* 7, 123–134, doi: https://doi.org/10.1007/s40607-020-00067-w.

Vrbinc, Alenka, Donna M. T. Cr Farina, Marjeta Vrbinc (2021), "The Yin and Yang of Dictionary Making. Slovenian Lexicographers Reflect", *Lexikos* 31, 159–76, doi: 10.5788/31-1-1634.

THE eSSKJ GROWING WEB DICTIONARY OF THE SLOVENIAN STANDARD LANGUAGE – ORGANISATION AND DISPLAY OF LINGUISTIC DATA

Abstract

This article presents the third edition of the *Dictionary of the Slovenian Standard Language* (*eSSKJ*), i.e. a growing dictionary of the contemporary Slovenian standard language that is produced at the ZRC SAZU, Fran Ramovš Institute of the Slovenian Language since 2016. The dictionary is updated yearly on *Fran*, the institute's dictionary portal. Since September 2021,

an adapted school edition of the dictionary has been available on the *Franček* portal for primary- and secondary-school students.

eSSKJ provides semantic information, basic formal and grammatical properties, e.g. orthographic properties, pronunciation and part-of-speech classification. In addition, it also includes information on pragmatic peculiarities of single- and multi-word lexis, typical collocations grouped according to syntactic patterns, full inflectional paradigms (complete with information on dynamic and pitch accents), phraseological units and their variants, etymological highlights and, where relevant, normative explanations and information on word formation. The tagging of data in the dictionary database enables different ways of displaying dictionary content for different users and user needs.

Due to the possibilities offered by the web medium, *eSSKJ* exceeds the classical mode of displaying and accessing linguistic information and provides links to external resources, such as the *Dictionary of the Slovenian Sign Language* (*Slovar slovenskega znakovnega jezika*). Each dictionary entry is also equipped with an audio recording of the pronunciation of the headword and its secondary form.

The content is displayed on two levels, which are dictated by the portal design. The first level is the list of search results provided to the user when they enter the desired word in the search box. Here, the *eSSKJ* search results include the basic information on the searched word (e.g. formal properties of the word, breakdown of meanings, set phrases and phraseological units), which, alongside information from other dictionaries (orthographic, historical, dialectal, etc.) provides the user with a comprehensive view of the word from different perspectives. The second level is the display of the entire dictionary entry, which the user gets by selecting the desired dictionary in the list of search results. In the case of *eSSKJ*, this provides a comprehensive description of the usage of the given word in contemporary Slovenian standard language.

Besides usefulness, the dictionary aims for understandability. The latter is pursued through dispensing with abbreviations, which are characteristic of printed dictionaries, a tendency towards explicitly listing information, the addition of explanations in tooltips, the introduction of dictionary metatext to label sections, etc.

New solutions in the interpretation of dictionary content can also be found in the school edition of eSSKJ on the Franček portal, which includes more helping hints, explicitly named microstructural elements (e.g. explanation, example, variant), brief explanations of dictionary labels and grammatical categories in tooltips, quick links to appropriate chapters in the school grammar, renamed terms and labels that are harder to understand (e.g. *collocations* to *typical phrases*, the label *also* to *less usual* in the context of doublets etc.). Students can thus understand the dictionary content more easily but also familiarise themselves with the dictionary structure in general and with the use of different dictionaries.

Key words: growing web dictionary, Slovenian, display of linguistic information, dictionary user