

Besedišče slovenskega jezika z oblikoslovnimi podatki (predstavitev)

Ivanka Šircelj-Žnidaršič

Konec leta 1998 je pri založbi ZRC v Ljubljani izšla zajetna knjiga (1009 strani) z naslovom *Besedišče¹ slovenskega jezika z oblikoslovnimi podatki* in podnaslovom *Po gradivu za slovar sodobnega knjižnega jezika zbrane besede, ki niso bile sprejete v Slovar slovenskega knjižnega jezika (SSKJ – njegov zadnji, 5. zvezek je izšel leta 1991)*. Kakor se torej razbere že iz podnaslova, je to zbirka besed, ki je ostala kot neobdelano gradivo za slovar sodobnega knjižnega jezika v kartoteki Inštituta za slovenski jezik, v pričujočem delu pa so te besede onaglašene in oblikoslovno obdelane. Vseh skupaj je 178.457 gesel ali iztočnic. V zadnjem delu knjige pa je še 12 precej povednih tabel ali preglednic, ki prikazujejo statistična razmerja med besednimi vrstami ter med oblikoslovnimi in naglasnimi pojavi.

Inštitut za slovenski jezik je za interno uporabo že leta 1987 izdal v SSKJ nezajeto besedje v dveh knjigah z naslovom *Besedišče slovenskega jezika*.² Tega je bilo kar 182.976 enot, torej precej več kot v SSKJ, ki jih vsebuje približno 107.000 (93.152 gesel in 13.888 podgesel), se pravi, le dobro tretjino vseh v inštitutski kartoteki zbranih besed. *Besedišče* je v javnosti vzbudilo precejšnje zanimanje, ob vse večji možnosti računalniške obdelave pa se je začelo razmišljati tudi o nadaljnjih predstavitev slovenskega besedja kot celote (med drugim še nimamo npr. lematizacijskega slovarja), zato se je Inštitut odločil za novo, izpopolnjeno izdajo, ki naj bi bila uporabnejša in primernejša za nadaljnjo obdelavo. Tako je bilo to gradivo leta 1994 preneseno v računalniško obliko (poskenirano) z domačim programom EVA (Primoža Jakopina), dodano pa mu je bilo še na novo zbrano gradivo iz let 1987–91, žal pa je v končni tehnični obdelavi ponagajal tiskarski škrat in je prav to izpadlo. Gre za približno 6000 besed. Objavljene bodo v nadaljevanju *Besedišča*, ko bo popisano in obdelano novo nabrano gradivo.

Približno 500 besed iz prvotnega *Besedišča* z večjo pogostnostjo rabe je bilo pred izidom zadnjega zvezka SSKJ sprejetih v Dodatke SSKJ, ki so objavljeni na koncu 5. zvezka.

Vse besede je bilo treba ponovno pregledati. Nezanosljive oziroma nejasno izpisane so bile izločene, vendar takih kljub vsemu ni bilo veliko, saj je bil osnovni namen obdelave ohraniti gradivo tako, kot je bilo zapisano.

¹ Izraz besedišče je že uveljavljena beseda – prim. Prešernovo besedišče v Prešernovem slovarju (Gradivo za Prešernov slovar, Stane Suhadolnik, ZRC SAZU, Ljubljana 1985) idr.

² V prvi nakladi je izšlo 250 izvodov, v ponatisu istega leta pa še 300.

Najprej je bilo treba preveriti pravilnost prepisa – vnosa v računalnik. Že pri izpisovanju gradiva za kartoteko je sem in tja ostala kaka napaka – npr. **zanaščen** pri geslu **zamastiti** (prav **zamaščen**), nato še pri prepisovanju oziroma sestavljanju geslovnika. Tam so nastale npr. napake: **zapogub** – prav **zapogib**, **uspulun** za **uspulin**, **uzrezati** za verjetno **urezati**, **vcveteti** najbrž za **vzcveteti** ipd. Če je bila ugotovljena taka napaka, je bila beseda navadno izločena, saj je v pravilni obliki že bila upoštevana v SSKJ, če pa je šlo za očitno tipkarsko napako, smo jo odpravili – npr. pri pridevniku **kroččast**, ki je bil prvotno zapisan kot **krozčast** (te oblike v SP ni, tam je samo zapis z -ž-, torej **kroččast**).

Pri zapisovanju besed se je ponujala skušnjava, da bi bile besede prikazane z vseh plati, tudi homonimno, seveda če ne bi upoštevali naglasov – npr.: **visôkoindustrijski** (visôkoindustrijska dežela = izraža stopnjo) proti **visokoindustrijski** (visokoindustrijski izdelki = izdelki visoke industrije)³ – tako kot sta npr. navedena oba izraza **velebitke** – naglašena na i pomeni 'cigareta', kot **vêlebitka** pa 'velika bitka', vendar sta ta dva bila izkazana v gradivu, prva dva pa ne oba. Upoštevano je bilo namreč le, kar je bilo dejansko kje zapisano.

Žal to načelo ni veljalo za besede, ki so formalno že bile registrirane v SSKJ, a so kot homonimi imele v gradivu še čisto drug pomen. Te v **Besedišče** niso bile sprejete, da ne bi vnesle zmede in nezaupanja v prikazano gradivo, jih je pa po svoje škoda. Če bi bila gesla opremljena tudi z razlagami, bi jih bilo mogoče normalno navesti.

Gesla, ki so v enakem pomenu navedena tudi v Slovenskem pravopisu ali v Pleteršnikovem Slovensko-nemškem slovarju, imajo na koncu oznako S oziroma P (ali oboje, če gre za oba).⁴

Zbrano gradivo obsega vsa področja, od splošno rabljenega ali manj pogosto rabljenega sodobnega knjižnega izrazja, narejenega po ustaljenih jezikovnih obrazcih, ki pa zaradi omejenega obsega SSKJ vanj ni bilo zajeto,⁵ do izrazov iz vseh drugih jezikovnih zvrsti, tudi neknjižnih (narečja, žargoni, sleng, ...), zato je včasih kateri izraz videti prav eksotičen. Veliko je malo rabljenih ali znanih poimenovanj stvari ali pojmov, starinskih in zastarelih, pa strokovnih izrazov, ki so se še ohranili ali že povsem opustili, in tudi takih, ki se ponovno oživljajo oziroma dobivajo še nove pomene in širijo svojo družino. Nekaj zelo nenavadnih besed je bilo opuščeni (izločeni) – npr. **antiresanten** za interesanten, **antičeš-kancelar** (enkratna Levstikova šaljiva skovanka). Izpuščeni so bili tudi medmeti, ki so take ali drugačne daljše razširitve tistih iz SSKJ, nekatere že skoraj nesmiselne, npr. **hmhmhmhm, brrrrrrr** ipd.

Lastnih imen, tujih rekel in daljših (tujih) citatov, kratic in simbolov v **Besedišču** ni.

Nekatere besede so le pisne variante navedenih v SSKJ (**forsitija** : **forzicija**) ali samo v **Besedišču**: **levstek** – **levstik**, **persuazija** in **persvazija** ..., druge z

³ Na tem zgledu se tudi nazorno vidi, kako se v besednozvezni razlagi pokaže pomensko-razločevalna vloga naglasov.

⁴ V nadaljevanju je Pleteršnikov slovar tudi v tem zapisu označen s P, Slovenski pravopis (1962) pa s SP.

⁵ Podrobno o tem glej v SSKJ I (A–H), Uvod, str. XI, člen 4, Ljubljana 1970.

nepravilnim zapisom glede na slovnicihna pravila: **laufer** : **lavfar**, **lesbijski** : **lezbij-ski**; **lontavški** – **lontovski**, **čiu** (medmet). (Zanimiv je zapis besede **hřtalec** -lca z l tudi v P, kjer bi pričakovali -vec.)

V starem gradivu (v malo primerkih nabrano od Trubarja dalje – Pohlin, Gutsman ...) je bilo tudi nekaj napak, ki jih je odkrilo samo zelo pozorno branje – npr. **řlahniček** : **řlahtniček**, ki sta bila v kartoteki vložena na napačnem mestu. Sum, da morda imata kaj skupnega, je pokazal povezavo med njima. Ipd.

Po abecedi urejeno preostalo besedje je bilo po razpoložljivih podatkih jakostno onaglašeno⁶ in oblikoslovno sistemizirano, tj. opremljeno z besednovrstnimi in oblikospreminevalnimi podatki po sistemu iz SSKJ: samostalniškimi in pridevniškimi besedam je bil določen spol, sklon (imenovalnik in roditelj ednine oziroma množine pri množinskih samostalnikih) in število, glagolom pa vrsta vida in osnovna osebna oblika. Druge besede oziroma nepregibne besedne vrste pa imajo samo besednovrstno oznako. Nekaj je pri tem tudi odstopanj ali nedoslednosti – npr. pri tipu **alma mater** je nesklonljivi del označen le s tilde, ki je SSKJ ne uporablja: prim. v **Besedišču** **pater familias** -tra ~, **light show** ~ -a (**persona non grata** pa je bila pri korekturah izločena, ostal je samo samostalni **persona**).

Odstopanje od SSKJ je tudi samo oznaka *mn.* za nesklonljivi samostalni (prim. v SSKJ npr. **shorts** *ž mn. neskl.*, v **Besedišču** pa je bilo npr. pri **slippers** v zadnjih korekturah⁷ *neskl.* črtano, torej je samo *ž mn.*, kar je nenatančno oziroma ni prav.

Podobno je pri samoedninskih tujkah. V **Besedišču** je (poleg besednovrstne oznake) le oznaka *neskl.* (prim. **noesis**), v SSKJ pa, da se rabijo v im. in tož. ed. (prim. **nemesis**).

Večkrat je bilo težko kako besedo že prav zapisati, drugo prav slovnicihno opredeliti in določiti pravilne oblike, kaj šele prav naglasiti (npr. **puba** – po obliki sam. ž. sp. je v resnici oblika za poba, to je m. sp.; prim. še **velečin** – je to vèle čin ali kaj drugega?). Kjer se je dalo opreti na družino, ni bilo večjih težav, težko pa je bilo z osamljenimi besedami, kjer se ni bilo mogoče na kaj opreti (prim. npr. **cetlica** 'planika' /zapis iz l. 1927/ ali **passim**, ki je pravilno latinski prislov, a je v **Besedišču** po analogiji prikazan kot samostalni moškega spola, ker iz prepomanjkljivega gradiva ni bilo mogoče ugotoviti, kaj je, obdelovalci **Besedišča** pa kot sami nelatini na latinski izvor besede očitno nismo pomislili).

Ponekod je gradivo že imelo naglas, čeprav včasih za knjižni jezik nenavaden. V veliki večini je ostal tak, če pa se je ugotovila napaka, je bil popravljen.

Pri določanju naglasnega mesta je bil primerjalno v pomoč SSKJ, dodatno pa še **Slovenski pravopis 1962**, **Pleteršnikov slovar** in po potrebi tudi še drugi slovarji in priročniki. Včasih pa tudi to ni zadostovalo in so bile potrebne še ustne poizvedbe pri strokovnjakih za posamezna področja, čeprav to ob začetku dela ni bilo predvideno, pa tudi čas je bil zelo omejen. (Zanimiv je bil na primer niz arheo-

⁶ **Besedišče** zaznamuje le mesto in kvantiteto naglasa, tonema pa žal ne, saj bi bilo tovrstno preverjanje zelo težko in izredno zamudno, pa tudi le relativno.

⁷ Korekture so bile opravljene ločeno od ostalega dela (in delavcev), zato je prišlo tudi do kake prej nedogovorjene spremembe.

loških izrazov **acheulec**, **acheuléen**, **acheulski** [ašélski] in variante **ašeléen**, ki so se prvotno nahajali v različnih skupinah po abecedi urejenega in po določenih kriterijih razdeljenega besedja in prišli skupaj šele ob dokončni abecedni ureditvi.⁸ Da ti izrazi sodijo skupaj in kako se prav izgovarjajo in naglašajo, smo se prepričali šele po pogovoru z našim starejšim arheologom, saj mladi tega niso poznali.)

Kjer se naglasov ni dalo preveriti, so lahko tudi napačni⁹ – npr. **kroš** ‘ciklama’ je naveden samo v Pleteršnikovem slovarju (Cigale, Medved), a brez naglasa; beseda je morda narečna, v **Besedišču** ima ostrivec).

Tuja naglasna znamenja so ostala (prim. **belle époque**, **prámoustérien**, ...), sicer pa podatkov o izgovorjavi tujih in prevzetih besed v **Besedišču** ni. Če je beseda v **Besedišču** nenaglašena, se izgovarja drugače, kot je zapisana.

Pri tem delu se je zelo jasno pokazalo, da je za pravilno naglaševanje treba poznati pomen oziroma sestavo ali izvor besede oziroma da drugačen naglas (z intonacijo vred!) pokaže drug pomen – prim. npr. **sénčece** (P) < **senó** ‘posušena trava’ : **sèncece** < **sencé** ‘del obraza ob strani čela’; **bélostén** (prim. belostno svitanje) : **belostén** (prim. belostena hišica); **avêr** -a *m num.* ‘sprednja stran novca, lice’ in **áver** -vra *m* ‘vrsta letala’ : pilotska kabina trenažnega avra; **gospodíca** ‘gospodična’ : **gospôdica** ‘gospoda’; **pogrâbljati** (< **grabiti**) : **pograbljáti** (< **grabljáti**). Posebno pomembno je poznavanje pomena pri sestavljenkah in zloženkah. Prim. npr. **nadméjen** ‘ki je, se nahaja nad mejo’ (ta v gradivu ni bil potrjen!) in **nâdméjen** ‘ki presega določeno mejo, stopnjo’; **nènapojèn** ‘ki ni napojen’ : **nenapójén** ‘ki se ne da napojiti, nenapojljiv’; **črnoglèd** – *m* ‘pesimist’ : **črnogléd** – *prid.* ‘pesimističen’.

Velika škoda je, da gradivo ni tudi pomensko obdelano. Kakršnakoli pomenska določitev, čeprav skopa, nepopolna ali samo usmerjevalna (npr. v besedno družino ali področje rabe), bi mu dala še posebno sporočilno vrednost in uporabnost, hkrati pa bi s tem postalo dostopno tudi širokemu krogu bralcev. Tako bi se dalo sproti preverjati pravilnost navedenega, sicer se bo lahko marsikdo ob kaki besedi, črkovno enaki ali zelo podobni drugi besedi, spraševal, zakaj tako in ne drugače oziroma, katera beseda je mišljena – ali gre zgolj za napako ali je tako prav? Prim. **žlompati** (voda žlompa v votlinah med skalami) proti **žlampati**¹⁰; **brbljín** *m* (< **brbljati**) ‘brblja, brbljač’, lahko pa bi bil tudi pridevnik **brbljin** (< **brblja**). Razlaga bi tudi takoj pokazala, da npr. pri besedi **oliti** oznake P ne bi smelo biti. V **Besedišču** je namreč ta glagol prikazan kot dovršni z obliko za 1. osebo sed. časa **olijem** (v pomenu ‘obliti’, npr. obliti torto), v Pleteršnikovem slovarju pa je **oliti** -im nedo-

⁸ Dejstvo, da je bilo besedje razdeljeno na več sklopov in se je torej kar nekajkrat začenjalo od začetka abecede ter se združevalo šele v končni fazi obdelave, je zelo otežilo delo – zahtevalo je veliko več popraviljanja in ponovnega preverjanja, kar je zavlačevalo delo. Porabljeni čas bi bilo koristneje porabiti za pisanje razlag.

⁹ Seveda pa se najde tudi kakšen tak, ki ga je zagrešil tiskarski škrat – prim. npr. **módrozêlenočrn**, pri katerem bi morala strešica biti na drugem e, a je z njega preskočila na prvi e.

¹⁰ Povezava je razvidna iz Pleteršnikovega slovarja – prim. **žlompotati** *nedov.* ‘žlapotati’, *Jan.* (op.: Tu je napaka, saj glagola **žlapotati** ni v slovarju, je le **žlampotati** *plätschern* ‘žuboreti’, ‘valovati’, ‘pljuscati’), pri **žlampotati** pa ima tudi kazalko na **žlampati** ‘s šumom in požrešno jesti ali piti’. Ima pa tudi geslo **žlapótek** ‘pokvarjeno jajce’ (*vzh. Št.*).

vršen in brez naglasa, pomeni pa 'postavljati ólihe' (vzh. št.), to je 'postavljati količke, stebričke', npr. ob cesti. Torej je **oliti** iz **Besedišča** nekaj povsem drugega kot tisto iz Pleteršnikovega slovarja in v takih primerih se oznake P v **Besedišču** ni dodajalo (prim. še **žžniti** -em *nedov.* – v gradivu pomeni 'siliti v koga', v Pleteršnikovem slovarju pa 'tleti'). Zaradi tako različnih pomenov bi oznake P v **Besedišču** ne smelo biti, kakor je ni npr. pri **odól** – v Pleteršnikovem slovarju pomeni 'dolina', v **Besedišču** pa 'zobna krema'. In ker je izkazan samo ta pomen, beseda ob sebi nima P-ja.

Podobne težave so bile s samostalniki, ki so bili ali ženskega ali moškega spola. Npr.: P ima **vèd** – véda *m* in **véd** -í ž v istem pomenu 'védenje', gradivo pa izkazuje samo ženski spol. Taki zapleti so zahtevali veliko dodatnega preverjanja. Ker ni razlag, se bo marsikaj koristnega in uporabnega tudi izgubilo in velikansko jezikovno-kulturno bogastvo, ki se skriva za temi besedami, bo ostalo prikrito. Kako naj npr. bralec ve, da se pridevnik **virementski** nanaša na francosko besedo **virement** = danes splošno uveljavljeni **virman**, iz katerega je izpeljan in splošno rabljen pridevnik **virnanski** (podatek je iz leta 1922)?

Pri iskanju pravih rešitev so se pokazala tudi odstopanja v rabi med P, SP, SSKJ in sedanjo rabo pri besedah, kjer naglas ni pomensko razločevalen, torej sprememba mesta naglasa ali sprememba kvantitete ne spreminja pomena. Zlasti izstopajoča so pridevniška obrazila na -iv: **piskotljív** -íva : **nezmotljív** v SSKJ. Sicer pa je pri teh pridevnikih v SSKJ dosledno dolžina in nobenkrat dubleta kakor pri nekaterih drugih.¹¹

V okviru samostalnikov in njihovih izpeljav imamo obrazila **-ica**, **-aj**, **-ar** idr., pri katerih naglaševalna stopnja niha od 0–2, tj. od nenaglašenosti do stalne naglašenosti, z različno kvantiteto in kvaliteto naglašenosti. Npr. v P je **pomagačica** (tako je ostalo tudi v **Besedišču**, po **pomagački** iz SSKJ in današnjem občutku pa bi bil naglas na drugem a-ju). Pleteršnikovo naglašanje tipa **pománjkljaj**, **pogóltlaj**, **pojémjlaj** je z današnjega stališča zelo opazno; prav tako **lógar** (tak naglas je tudi še v SP, in to za oba pomena: 1. 'kdor ima hišo v logu', 2. 'gozdni čuvaj'; v SSKJ pa je **logár**); P **cukrovár** je v SSKJ **cukrovár**; P **topničár**, **topničárski**, **topničárstvo** so v SSKJ (prva dva) in **Besedišču** (zadnji) naglašeni na prvem zlogu, prav tako tudi **tópníčarica** (**Besedišče**). Vprašanje je, ali bi se v teh primerih smelo navajati P (s pripombo, da naglas ni identičen). Prim. še **metlíca** < **métla**, **metlár** (SSKJ **métlica** tudi **metlíca**); P **hitromèr** -êra – v SSKJ samo -mér (**tlakomér**, **kislinomér**, **vetromér** idr.).

Ob delu za **Besedišče** so se pokazale tudi nedoslednosti pri naglaševanju izpeljank iz predložnih zvez v samem SSKJ. Prim. npr. **medgórije**, **medósje**, **medpléčje**, **medóčje**, **mednóžje**, pač pa **mèdkróvje**; besede **ilegála**, **ilegálec**, **ilegálstvo** danes izgovarjamo samo z enim naglasom, pri nekaterih drugih besedah pa izgovarjamo dva naglasa, v SSKJ pa je samo eden. Prav tako je npr. pri zloženkah tipa **čudapóln**, **nadepóln**, **trudapóln** proti: **grózepóln**, **múkepóln**, **úpapóln**, **usódepóln**. V **Besedišču** imajo tovrstne besede po dva naglasa: **bogápóln**, **bôjapóln**, **čústvapóln**.

¹¹ Pleteršnikove kračine so v SSKJ lahko tudi na drugem mestu – prim. npr. **dvostràn** – v SSKJ **dvóstràn** in **dvóstràn**; tako tudi pri **belolas**, **dolgokrak**, **kratkodlak**, **zlatoklas** idr.

Ne glede na pomanjkljivosti pa je *Besedišče* kljub temu izredno zanimiv in svojevrsten dokument oblikovanja in obstajanja slovenščine in omogoča dodaten, zanimiv vpogled v slovenski jezikovni laboratorij, zgovoren in pomenljiv za vsakogar, ki razmišlja o jeziku.¹² Skupaj s svojim odzadnjim slovarjem bo dragoceno dopolnilo k SSKJ tudi za nejezikoslovce.

Pričujoča predstavitev nikakor ni popolna. Njen namen je opozoriti na osnovne probleme oziroma zagate, na katere naletimo pri obdelavi besedišča. Iz tega bi se potem lahko izoblikovala natančnejša izhodiščna merila za vse nadaljnje obdelave.